Submitted to the Annals of Statistics

PCA FOR POINT PROCESSES

BY FRANCK PICARD[1,a] VINCENT RIVOIRARD[2,b]
ANGELINA ROCHE[2,c] AND VICTOR PANARETOS[3,d]

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

[2]CEREMADE, Université Paris Dauphine, , [b]rivoirard@ceremade.dauphine.fr; [c]roche@ceremade.dauphine.fr

[3]Institut de Mathématiques, Ecole Polytechnique Fédérale de Lausanne, , [d]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 [40], 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 [32, 33] 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 [6, 17, 42] to genomics [12, 30], ecology [56], epidemiology [15], seismology [48], social sciences [20, 28, 46] and stock prices moves [26] to name but a few.

Statistical inference for point processes has a long history [1, 39], with non-parametric approaches being favored for analyzing Poisson [4, 41, 52, 57] and Hawkes processes [3, 13, 31, 53]. Maximum likelihood [5, 17, 48] and the Bayesian framework have also shown

MSC2020 subject classifications: Primary 62M, 62P; secondary 60G55, 62H25.
Keywords and phrases: Point processes, Karhunen-Loève expansion, PCA.

good performance [25, 50, 55]. 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 [48], 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 [38]. In single-cell genomics, it is now possible to individually characterize genomes based on epigenetic marks organized along chromosomes in a one-dimensional array [60]. 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 developping a new framework for dimension reduction and visualisation 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 [43]. It enables the visualization of functional data, seen as random curves over compact sets (as described in [54]), enables dimensionality reduction and regularisation in regression/testing/classification [24, 34, 51], and provides a bridge to extend classical multivariate procedures (like model based clustering) to functional data [27, 37]. The theoretical underpinnings of fPCA rely on two fundamental results: the Karhunen-Loève Theorem and Mercer's Theorem [35]. 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: [58] have considered PCA based on kernel estimators of the intensities of the point processes and they obtained a Karhunen-Loève decomposition for the corresponding intensities. A similar approach, based on the estimation of L-functions, has been adopted in [36] in the case of spatial processes. 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 [44] who proposed a convenient Hilbertian setting for PCA of measures for the analysis of distribution functions appearing in the study of grain-size curve. Closest in spirit to our approach, [11] recently obtained Karhunen-Loève 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 $(N_1, N_2, \dots, N_n)$, 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 (10). 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], that 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 introducess 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 Section 8.

Notation. We introduce some notation that will be useful throughout the article. For a set I, we denote

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

where $\|f\|$ stands for the $\mathbb{L}_2$-norm of f: $\|f\| = \left(\int_I |f|^2\right)^{1/2}$. The associated scalar product is denoted $\langle f, g\rangle = \int_I f(t)g(t)dt$. Finally, for $f \in \mathbb{L}^2(I)$ and μ a measure on I, we also use the notation $\langle f, \mu\rangle = \int_I f(t)d\mu(t)$.

[1] https://github.com/franckpicard/pppca

2. A signed-measure model for dimension reduction. Consider n independent and identically distributed (i.i.d.) temporal point processes $(N_1, N_2, \ldots, N_n)$, 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 $(\Pi_1, \ldots, \Pi_n)$ so that:

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

where $\mathcal{B}$ is the Borel σ-field of subsets of $[0,1]$. Hence, $\Pi_1, \ldots, \Pi_n$ 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 N the point process associated with Π. Assuming that $\mathbb{E}[\Pi([0,1])] < +\infty$, we denote the first moment of Π by W:

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

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

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

The Δ_i's have the same distribution as $\Delta = \Pi - W$. Then, we define of a second-order moment for Δ, inspired from the formalism proposed by [49] and [11]. For this purpose, we define the signed measure of covariance: for all $B \times B' \in \mathcal{B} \otimes \mathcal{B}$

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

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

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

We assume throughout that

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

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

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

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

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

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 non-negative 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

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

with uniform and absolute convergence of the series (see Theorem 4.6.5 of [35]). 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 $\lambda_j \to 0$ when $j \to +\infty$ (see Chapter 6 of [10]). In the sequel, we assume that in (3), all the λ_j's are strictly positive. 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 Assumption (1) is satisfied and that K_Δ is continuous. Then, for all $j \geq 1$, the derivative in the distributionalsense 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 $\varphi \in \mathcal{H}_0^1 = \Big\{ f \in \mathbb{L}^2(I) : f' \in \mathbb{L}^2(I)$ and $f(t) = 0$ for all $t \notin I \}$,

$$\langle \eta_j, \varphi' \rangle = \langle F_{\mu_j}, \varphi' \rangle = -\int_0^1 \varphi(s) d\mu_j(s) = -\langle \varphi, \mu_j \rangle. \tag{4}$$

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 $[0,1]$,

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

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:

$$\Pi(B) = W(B) + \sum_{j\geq 1} \sqrt{\lambda_j} \xi_j \mu_j(B), \qquad B \in \mathcal{B}. \tag{5}$$

The convergence of the series holds then pointwise, i.e. 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, unifromly 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 $I = (0,1)^2$

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

where, for a bivariate function $f : (s,t) \mapsto f(s,t)$, and a bi-index $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}^2$, we set $|\alpha| = \alpha_1 + \alpha_2$ and $\partial_\alpha f = \frac{\partial^{|\alpha|} f}{\partial s^{\alpha_1} \partial s^{\alpha_2}}$; when f is a univariate function $\partial_\alpha f = f^{(\alpha)}$. 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. Sobolev norms of negative order are also introduced as the norm of the dual space of $\mathcal{H}_0^k$, namely

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

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

THEOREM 3.2 (Karhunen-Loève Theorem for point processes). *Suppose that Assumption (1) is satisfied and that K_Δ is continuous. Then, there exists a sequence $\{\xi_j\}_{j\ge1}$ of uncorrelated real random variables of mean zero and variance one such that*

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

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

$$\lim_{J\to+\infty} \sup_{\varphi\in\mathcal{H}_0^1, \|\varphi\|+\|\varphi'\|\le1} \mathbb{E}\left[\left(\sum_{T\in N} \varphi(T) - \mathbb{E}\left[\sum_{T\in N} \varphi(T) \right] - \sum_{j=1}^J \sqrt{\lambda_j} \xi_j \langle \mu_j, \varphi \rangle \right)^2 \right] = 0.$$

By analogy with fPCA, the elements of the sequence $\{\xi_j\}_{j\ge1}$ 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 6). Indeed, for each individual i, it can be deduced from the Karhunen-Loève decomposition that for J large,

$$\Pi_i(B) \approx W(B) + \sum_{j=1}^J \sqrt{\lambda_j} \xi_{i,j} \mu_j(B), \qquad B \in \mathcal{B} \tag{6}$$

or

$$\Pi_i([0,t]) \approx W([0,t]) + \sum_{j=1}^{J} \sqrt{\lambda_j}\xi_{i,j} F_{\mu_j}(t), \qquad t \geq 0 \tag{7}$$

with

$$\xi_{i,j} = \frac{\langle F_{\mu_j}, F_{\Delta_i}\rangle}{\sqrt{\lambda_j}}. \tag{8}$$

As to what concerns the covariance measure C_Δ, we establish the following Mercer theorem:

THEOREM 3.4 (Mercer's Theorem for C_Δ). *We suppose Assumption* (1) *is satisfied and that K_Δ is continuous. Then*

$$\left\| \sum_{j=1}^{J} \lambda_j \mu_j \otimes \mu_j - C_\Delta \right\|_{\mathcal{H}^{-2}} \xrightarrow[J \to +\infty]{} 0. \tag{9}$$

In the next sections, we focus on the study of the eigenelements $(\lambda_j, F_{\mu_j})_{j\geq 1}$. Recall that we have an easy correspondence between the measures μ_j and the mass functions F_{μ_j} thanks to Eq. (2). In Section 4, we study Poisson and specific classes of Hawkes processes analytically. For these central examples of points 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 $(N_1, N_2, ..., N_n)$.

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 behave, at least asymptotically. For the two classical point process models, Poisson and Hawkes with exponential self-exciting functions, we show that determining their eigenelements involves solving second-order differential equations of the form

$$\begin{cases} -\lambda y''(t) = w(t)y(t) + \int_0^1 M(s,t)y(s)ds, t \in (0,1), \\ y(1) = 0, \quad y'(0) = 0. \end{cases} \tag{10}$$

More precisely,

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 Eq. (18)).

Observe that for any constant $c > 0$, (λ, y) is a solution of (10) if and only if $(\lambda, c \times y)$ is also a solution. For $j \geq 1$, we introduce

$$H_j(t) = \int_t^1 F_{\mu_j}(s)ds, \quad t \in [0,1], \tag{11}$$

and 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 (10), once H_j is renormalized so that $\|F_{\mu_j}\|_2 = 1$.

Relationships that 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 (10) 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 straighfroward proof of Theorem 4.1, below. 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 $t \in [0,1] \longmapsto w(t)$ 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

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

and

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

The following result, which involves the functions H_j introduced in (11), provides the PCA dimension reduction for Poisson processes.

THEOREM 4.1. *Assuming w is continuous and positive on $(0,1)$, then $(\lambda_j, F_{\mu_j})_{j\ge1}$ are the eigenelements of the operator Γ_Δ if and only if $(\lambda_j, H_j)_{j\ge1}$ are solutions of the following Sturm-Liouville problem:*

$$\begin{cases} -\lambda y''(t) = w(t)y(t), & t \in (0,1), \\ y(1) = 0, \quad y'(0) = 0. \end{cases} \tag{12}$$

Mercer's and Karhunen-Loève expansions for Poisson processes are then provided by solutions of System (12) 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 (12) with $\lambda = 0$ and then $F_{\mu_0} = 0$.*

Unfortunately, solutions of System (12) 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 (12) can be obtained explicitly

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

which implies, taking into account the constraint $\|F_{\mu_j}\| = 1$,

$$F_{\mu_j}(t) = \sqrt{2}\sin(\pi(2j-1)t/2), \quad t \in [0,1], \quad j \geq 1. \tag{13}$$

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

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

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

with $\{\xi_j\}_{j\geq 1}$ uncorrelated centred 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 the Brownian motion, so that the Karhunen-Loève decomposition of Π can be deduced from the one of the Brownian motion (see e.g. [2], pp. 41–42). The main difference with the Brownian motion is that the scores $\{\xi_j\}_{j\geq 1}$ are not Gaussian.*

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 (12), in full generality. But we prove in this section that solutions of (12) exist and some useful qualitative properties of them can be obtained. Actually, system (12) 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 [59] 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 [59].

THEOREM 4.4 (Theorems 4.3.1 and 4.6.2 of [59]). *We consider the solutions $(\lambda_j, H_j)_j$ of the Sturm-Liouville problem (12) 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*

$$\lambda_j \overset{j\to+\infty}{\sim} \frac{\left(\int_0^1 \sqrt{w(u)}du\right)^2}{\pi^2 j^2}.$$

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* $\mathbb{L}_2([0,1],w)$*, i.e.*

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

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 $j \to +\infty$. 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 (13)). 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 Surm-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 [32, 33] 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

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

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 occurence at time t given the past.

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

$$w(t) = w_0 + \int_{-\infty}^{t^-} h(t-s)dN_s = w_0 + \sum_{T \in N,\, T<t} h(t-T). \tag{14}$$

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, i.e.

$$h(t) = \alpha \exp(-\beta t), \quad t > 0, \tag{15}$$

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 [47]. In the sequel, we assume that $\|h\|_1 < 1$. For the case of exponential self-exciting function h given in (15), this condition is equivalent to $\alpha < \beta$. If this assumption is violated, that is, in the case $\|h\|_1 > 1$, we would eventually get some clusters that produce offspring indefinitely, causing the process to explode. Furthermore, assuming $\|h\|_1 < 1$ ensures that there exists a unique stationary version of the Hawkes process N with intensity given by (14). Under this assumption, we can derive the mean and the covariance kernel of N. 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,*

$$F_W(t) = W([0,t]) = \frac{w_0 \beta t}{\beta - \alpha}, \quad t > 0, \tag{16}$$

and

$$K_\Delta(s,t) = \frac{\beta^3 w_0}{(\beta-\alpha)^3} s + \frac{w_0 \alpha \beta (2\beta - \alpha)}{2(\beta-\alpha)^4} \left(-1 - e^{(\alpha-\beta)(t-s)} + e^{(\alpha-\beta)t} + e^{(\alpha-\beta)s} \right), \quad 0 < s \le t. \tag{17}$$

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 (11), the following result provides the PCA decompositin for Hawkes processes with exponential self-exciting function.

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

$$\begin{cases} -\lambda y''(t) = \dfrac{\beta w_0}{\beta - \alpha} y(t) + \dfrac{w_0 \alpha \beta (2\beta - \alpha)}{2(\beta-\alpha)^2} \displaystyle\int_0^1 e^{-(\beta-\alpha)|t-s|} y(s) ds, \, t \in (0,1), \\ y(1) = 0, \quad y'(0) = 0. \end{cases} \tag{18}$$

REMARK 4.9. *Similarly 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 if α is small, the solutions of the ODE (18) 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*

(19) $$\frac{\alpha(2\beta-\alpha)}{\beta-\alpha}\Big(2+\frac{3-2e^{-(\beta-\alpha)/2}-e^{-(\beta-\alpha)}}{\beta-\alpha}\Big)<1.$$

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 (18) such that $\|F_{\mu_j}\|_2=1$ and we have

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

and

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

for C_1 and C_2 two constants not depending on j.

Let us discuss Assumption (19). Observe that the term

$$\Big(2+\frac{3-2e^{-(\beta-\alpha)/2}-e^{-(\beta-\alpha)}}{\beta-\alpha}\Big)$$

belongs to the interval $[2,4]$ whatever the values $0<\alpha<\beta<\infty$ may be. Therefore, setting

$$r=\frac{\alpha}{\beta}$$

assumption (19) mainly depends on the term

$$\frac{\alpha(2\beta-\alpha)}{\beta-\alpha}=\beta\times\frac{r(2-r)}{1-r}.$$

Since the function $r\longmapsto\frac{r(2-r)}{1-r}$ is increasing on $(0,1)$ from 0 to $+\infty$, Assumption (19) is satisfied if both r and β take small values. Observe that r measures the strength of the dependence (the larger r, the larger the number of occurrences, see (16)) and β its range (the larger β the closer to the present past occurrences are).

Surprisingly, we observe that the eigenvalues have an asymptotic behavior similar to the homogeneous Poisson case, except that it involves the constant w_1 instead of w_0. Since the upper bound (20) is polynomial in j, approximations are sharp. Of course, $w_1=w_0$ if $\alpha=0$ and w_1 is close to w_0 when β takes large values with fixed α. This is expected since both cases correspond to situations where the past has small influence on new occurrences. The Poisson analogy for eigenfunctions F_{μ_j} is valid as well (see (13)). These results are confirmed by the numerical study of Section 6.

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 (6)). 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, 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. 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

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

which provides the definition of the empirical covariance kernel:

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

along with the empirical integral operator:

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

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

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

Then, $\mathrm{Im}(\widehat{\Gamma}_{\widehat{\Delta}}) \in \mathrm{span}\{F_{\widehat{\Delta}_1}, \dots, F_{\widehat{\Delta}_n}\}$, which implies that $\widehat{\Gamma}_{\widehat{\Delta}}$ 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 $\widehat{\Gamma}_{\widehat{\Delta}}$. We denote by $(\widehat{\lambda}_j)_{j\geq 1}$ the associated eigenvalue sequence, sorted in non-increasing order, so that:

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

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

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

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

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

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

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

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

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

LEMMA 5.1.

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

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

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

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

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

$$\mathrm{Sp}(\widehat{\Gamma}_{\widehat{\Delta}}) \backslash \{0\} \subset \mathrm{Sp}(\widehat{G}_{\widehat{\Delta}}) \subset \mathrm{Sp}(\widehat{\Gamma}_{\widehat{\Delta}}).$$

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

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

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

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

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

REMARK 5.4. *The estimated scores can be written*

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

where the last equality comes from Equation (49).

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

THEOREM 5.5. *We assume that*

$$\mathbb{E}[\|F_\Delta\|^4] < +\infty. \tag{22}$$

We have:

$$\mathbb{E}\Big[\sup_{j\geq 1} |\widehat{\lambda}_j - \lambda_j|^2\Big] \leq 4\frac{\mathbb{E}[\|F_\Delta\|^4]}{n}. \tag{23}$$

If we further assume that the eigenvalues of Γ_Δ are simple, then, by introducing the eigengaps

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

and

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

with $\widetilde{\eta}_j = \operatorname{sign}(\langle\widehat{\eta}_j, \eta_j\rangle)\eta_j$, we have

$$\mathbb{E}[\|\widehat{\eta}_j - \widetilde{\eta}_j\|^2] \leq 32\delta_j^{-2}\frac{\mathbb{E}[\|F_\Delta\|^4]}{n} \tag{24}$$

and

$$\mathbb{E}\left[\|\widehat{\mu}_j - \widetilde{\mu}_j\|^2_{\mathcal{H}^{-1}}\right] \leq 32\delta_j^{-2}\frac{\mathbb{E}[\|F_\Delta\|^4]}{n}, \tag{25}$$

with $\widetilde{\mu}_j = \operatorname{sign}(\langle\widehat{\eta}_j, \eta_j\rangle)\mu_j$.

The moment Assumption (22) 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 (19) (see Theorem 4.10). Even if our framework is nonparametric in nature, 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 non-asymptotic setting? For Hawkes processes, some restrictions due to Assumption (19) 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 $n = 1{,}000$ Hawkes processes on $[0, 100]$, with $w_0 = 100$, and self-exciting function $h(t) = \alpha e^{-\beta t}$. Data

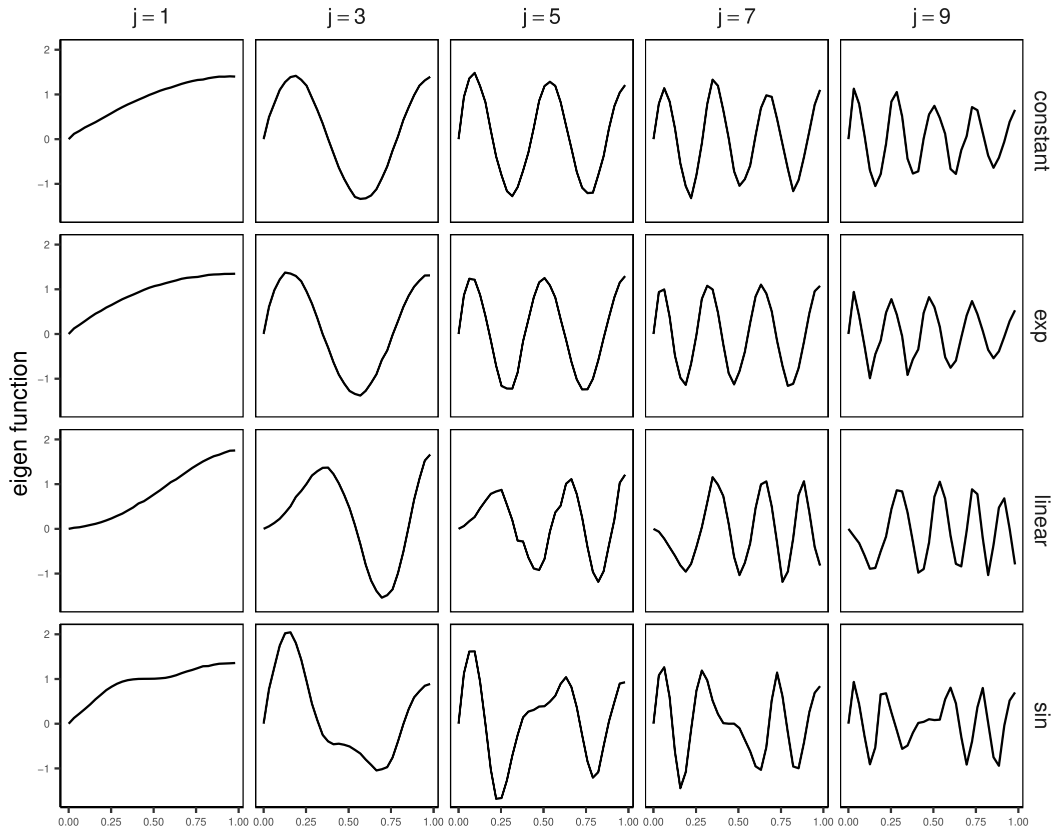

FIG 1. *Average eigenfunctions for Poisson processes with different intensity functions over 50 replicates*

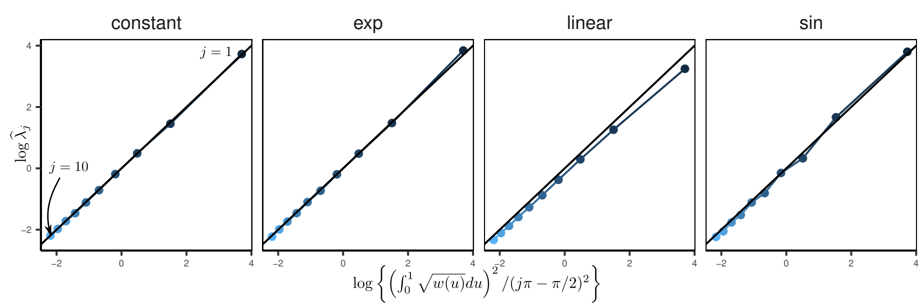

FIG 2. *Average eigenvalues (log-scale) for Poisson Processes over 50 replicated. Each dot corresponds to a value of* $j \in \{1, \ldots, 10\}$. *The empirical average is plotted vs the expected theoretical asymptotic regime of eigenvalues in* $(\int_0^1 \sqrt{w(u)}du)^2/(j\pi - \pi/2)^2$, *as expected 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.*

were finally rescaled in $[0, 1]$. To explain briefly the simulation method, the package considers a reproduction function h that is decomposed such as $h = mh^*$, with $m = \int h(t)dt$ and h^* a true density function such that $\int h^*(t)dt = 1$. In our setting, this means that $h^*(t) = \beta \exp(-\beta t)$, and that $h(t) = r \times h^*(t)$, with $r = \alpha/\beta$. Since Theorem 4.10 exhibits the central role of r and β, in our simulation we consider $\beta \in \{0.01, 0.1, 1, 10\}$ and

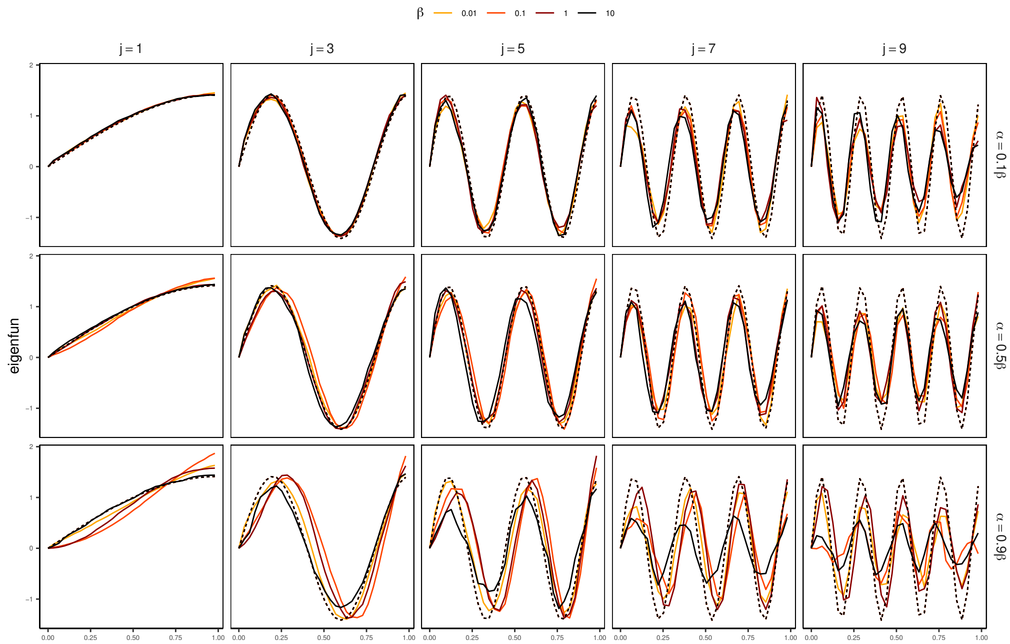

FIG 3. *Average eigenfunctions for Hawkes Processes with different transfert functions over 50 replicates. Dotted lines: asymptotic eigenfunctions* $t \longmapsto \sqrt{2}\sin(\pi(2j-1)t/2)$ *(see Theorem 4.10).*

$r \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$. 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 do not depend on the parameters (α, β) and agree with the values predicted by the theory. However, the quality of eigenfunctions estimation deteriorates for large j's, and when r or β increase. These results are not surprising in view of Assumption (19). As for eigenvalues, their asymptotics in j differs from the expected theoretical regime for large values of r and β (Figure 4). The empirical results on eigenvalues illustrate the critical importance of r on convergence, but also that of β, which confirms the central role of Assumption (19) in our theoretical results.

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 r and β 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 (Fig. 5-A). In recent years, the Gulf of Gökova in Southwest Turkey has witnessed two significant seismic events: the Bodrum earthquake on July 20, 2017 (magnitude 6.6, also felt on the Greek island of Kos) [38],

[2] http://www.koeri.boun.edu.tr/sismo/2/earthquake-catalog/

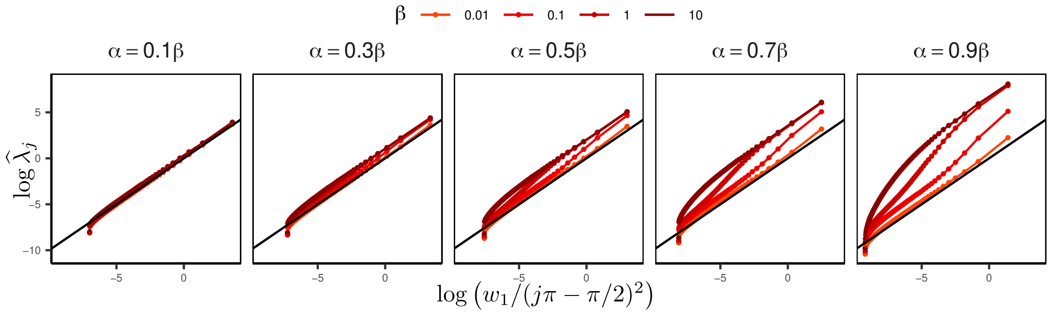

FIG 4. *Average eigenvalues (log-scale) for Hawkes Processes over 50 replicated. Each dot corresponds to a value of* $j \in \{1, \ldots, 50\}$*. The empirical average is plotted vs the expected theoretical asymptotic regime of eigenvalues in* $w_1/(j\pi - \pi/2)^2$*, as expected from Theorem 4.10. The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

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.

As in classical PCA, we first inspect the percentage of variance carried by each axis (Fig. 5-B), the first axis carrying $\widehat{\lambda}_1/\sum_j \widehat{\lambda}_j = 89\%$ of the variability. This implies 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 (as shown by $\widehat{\eta}_1$, Fig. 5-C). Also, scores on Axis 1 ($(\widehat{\xi}_{i1})_i$) directly correspond to the total number of earthquakes in each city over the specified period (Fig. 5-D). We identify Akdeniz as an outlier with an unusually high number of earthquakes compared to the regional average. This city will be excluded from subsequent analyses that investitage 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 (6). Consequently, we focus on the variations of the estimated eigenfunctions $(\widehat{\eta}_j)_j$, and represent the positions of cities according to their scores $(\widehat{\xi}_j)_j$ on successive axes as in any PCA analysis (Fig. 6). These representations allow us to identify cities (like Gokova Korfezi, Onika Adalar) that have typical accumulations of earthquakes over the period.

Interestingly, the second axis, $\widehat{\eta}_2$ (Fig. 6-C) reveals a distinct change in seismic activity between the two main earthquakes (2017 and 2020). Cities exhibiting positive scores on this axis (Fig. 6-A) 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 $\widehat{\eta}_3$ (Fig. 6-A) 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, captures sharp and fine-scale variations in this dynamics, as shown by the distinct peaks in $\widehat{\eta}_4$ and $\widehat{\eta}_5$ (Fig. 6-E-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 molec-

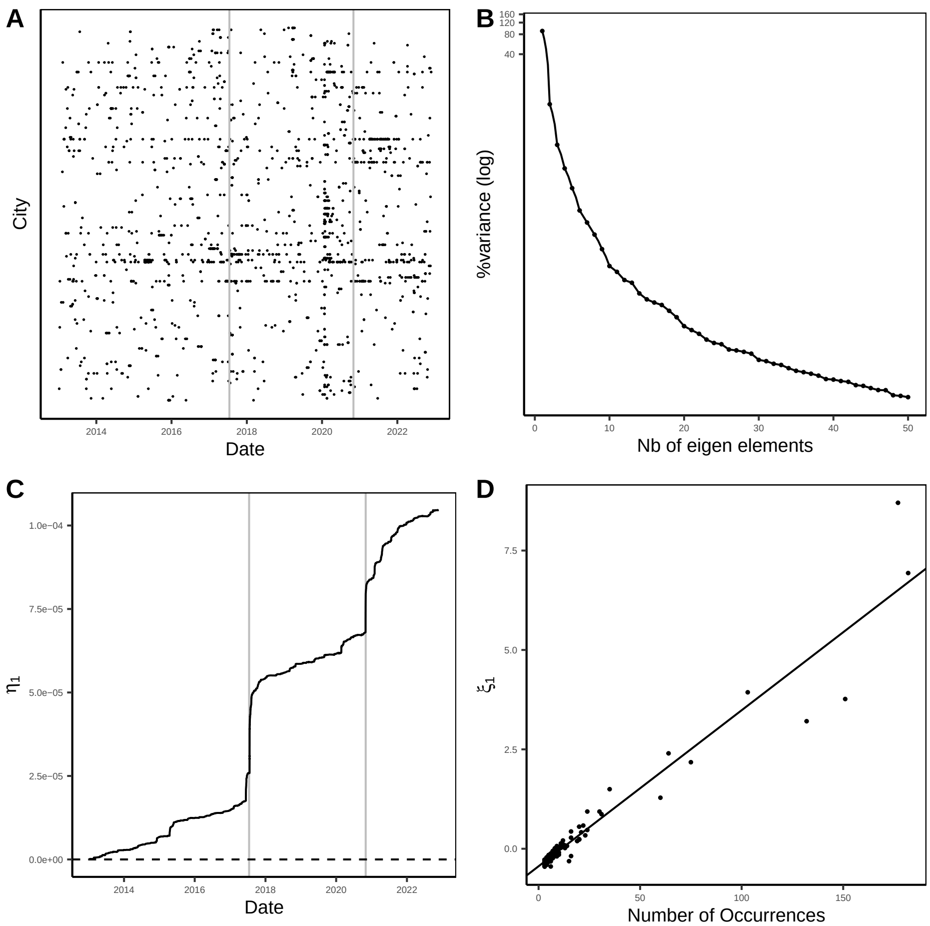

FIG 5. *A: Raster plot of earthquakes during the period 2013-2023. Each line corresponds to a city and each dot to an earthquake occurrence. Breakpoint dates (grey vertical lines) correspond to 2017.07.20 and 2020.11.01 B: Percentage of variance according to the number of eigenelements (log scale). C: First eigenfunction $\widehat{\eta}_1$ 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 $\widehat{\xi}_{i1}$ according to the number of occurrences $W([0,t])$.*

ular portraits of cell populations, providing valuable insights into the complexity of living tissues [60]. 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 [45] 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.

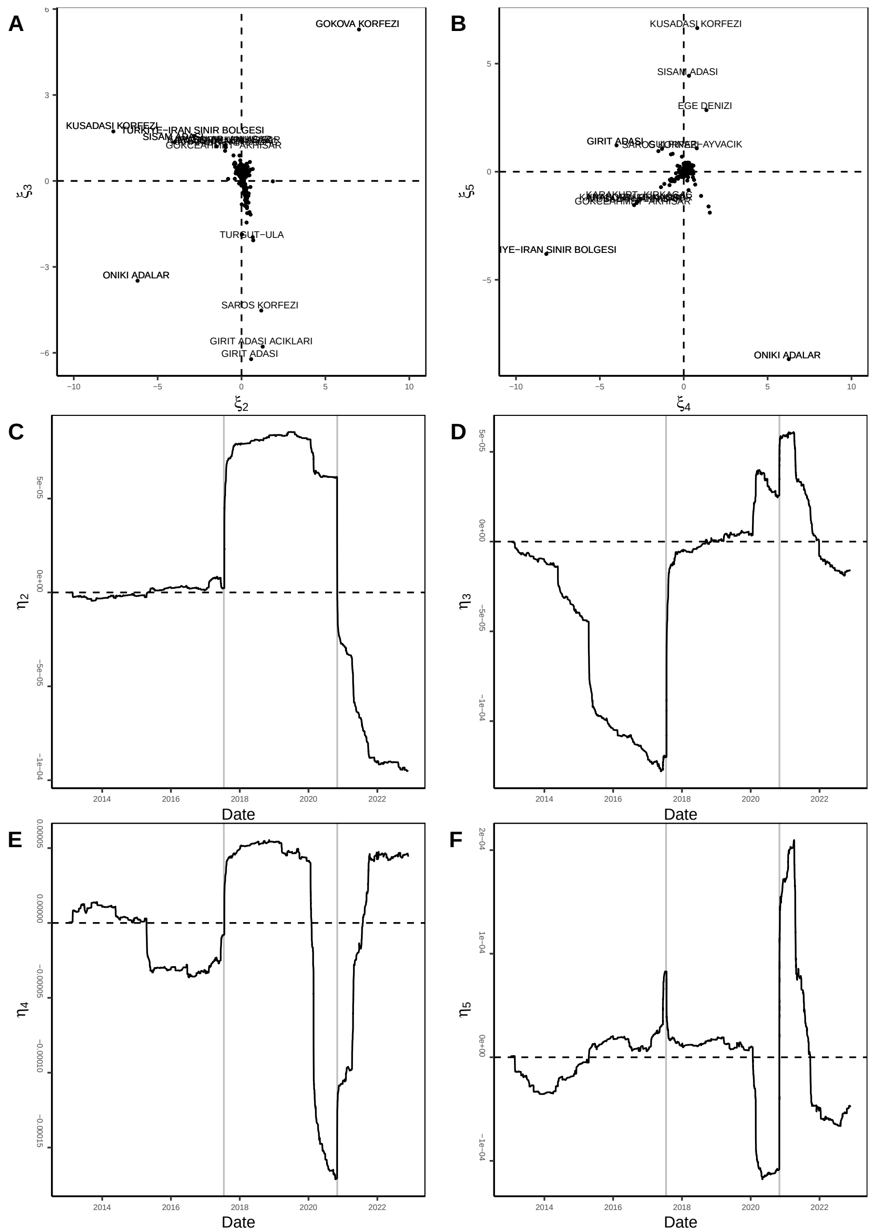

FIG 6. *A: Estimated scores of cities* $\widehat{\xi}_{i2}$ *vs.* $\widehat{\xi}_{i3}$*. B: Estimated scores of cities* $\widehat{\xi}_{i4}$ *vs.* $\widehat{\xi}_{i5}$*. C-D-E-F: Estimated eigenfunctions* $\widehat{\eta}_j$ *over time for* $j = 2, 3, 4, 5$.

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 (Fig. 7 A-B). Our PCA framework then provides principal scores $\widehat{\xi}_{ij}$, 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 ($\widehat{Z}_{ik} = 1$ if cell i is in group k), we represent the estimated cluster-wise probability mass function ($\widehat{\Pi}_k([0,t])$, Fig. 7 C-D) and intensity ($w_k(t)$, Fig. 7 E-F):

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

This analysis reveals the spatial variability of histone modifications among cancer cells, offering a new perspective on epigenomic heterogeneity by revealing three sub-populations 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 multi-electrodes 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 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 co-vary, 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 non-human 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-Electrods-Array (MEA, $n = 256$ electrods) grid. Theses 256 electrodes form a 16 x 16 grid onto a device that is positionned on the retina of macaques. The activity of each electrode can be related to a particular retina region, so that the data are typically spatio-temporal, 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 (Fig 8 A), as well as the temporal dynamics that explain this variability in spike train occurrences (Fig 8 B). These principal components can be interpreted as latent spike trains that compose the dynamics of the population. For instance, principal measure 3 ($\widehat{\eta}_3$) shows a change in dynamics following the stimulus. These latent processes could be of interest for understanding the neuronal coordination underlying the population response. Additionally,

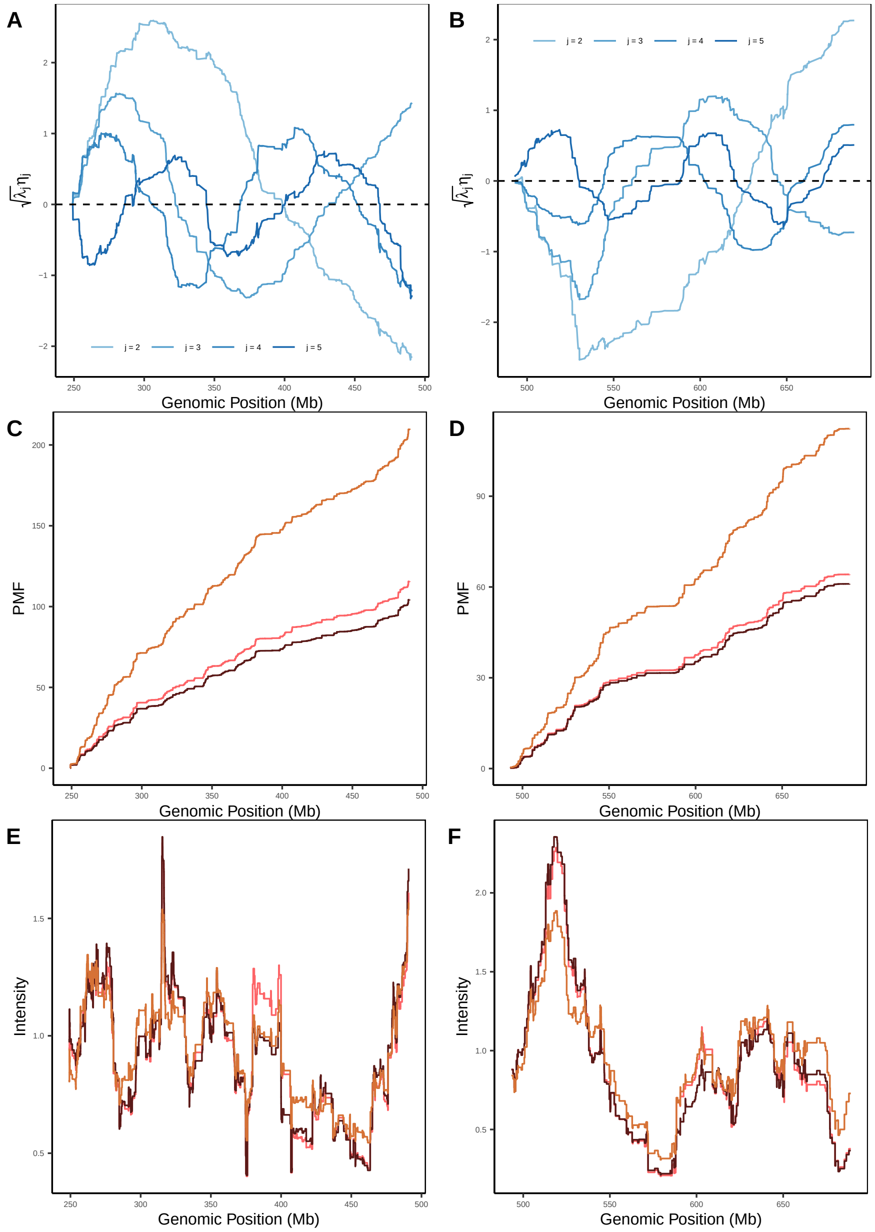

FIG 7. *A-B: Eigenfunctions weighted by their respective eigenvalues* $\sqrt{\widehat{\lambda}_j}\widehat{\eta}_j$ *according to the position along the genome. Chromosome 2 (left) and Chromosome 3 (right). C-D: Cluster-wise probability mass functions. E-F: scaled cluster-wise intensity functions.*

the score of each electrode can be represented with respect to the position of the electrodes on the device (Fig 8 C,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 (Fig 8 E-F). In this representation, each time point t corresponds to a single point in the latent firing rate space $\widehat{\eta}_2(t), \widehat{\eta}_3(t)$. 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

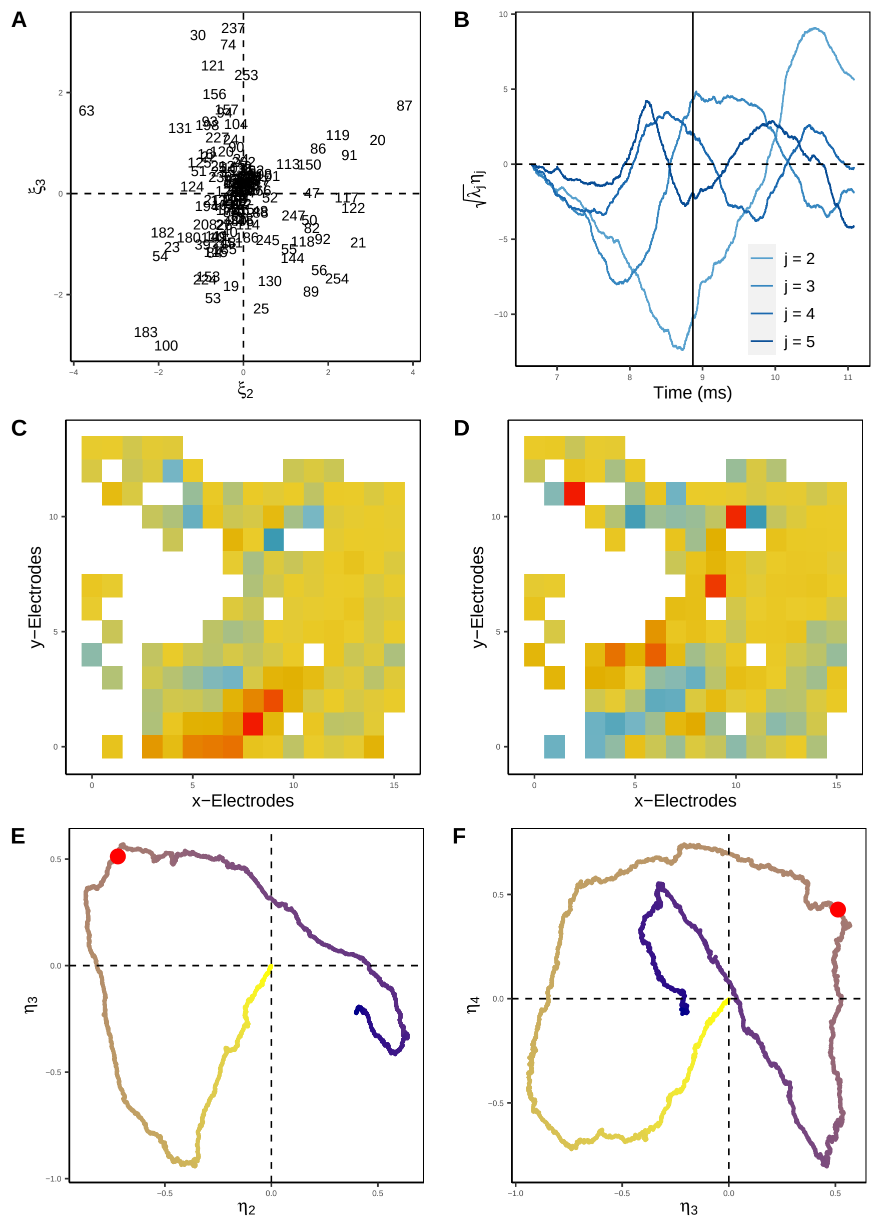

FIG 8. *A: scores* $\widehat{\xi}_2$ *vs* $\widehat{\xi}_3$ *for electrodes* $n = 1, \ldots, 256$. *B: Eigenfunctions weighted by their respective eigenvalues* $\widehat{\lambda}_j^{1/2}\widehat{\eta}_j$ *according to time (in ms). The vertical line corresponds to the stimulus at* $t = 8.86$*ms. C-D: visualisation of scores* $\widehat{\xi}_2$*, (C) and* $\widehat{\xi}_3$ *(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,* $\widehat{\eta_3}$ *vs* $\widehat{\eta_2}$*, and F,* $\widehat{\eta_4}$ *vs* $\widehat{\eta_3}$*). Time is represented by colors (yellow: begining of the assay, purple: end of the assay). Red dot: stimulus time.*

population-based spike train analysis by providing simple decompositions that help unravel the diversity and variability of neuronal population spiking activities.

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.

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.

8. Proofs.

8.1. *Proofs of Section 3.*

8.1.1. *Proof of Proposition 3.1.* Let $j \geq 1$. We observe that, since $\lambda_j > 0$,

$$\eta_j(t) = \lambda_j^{-1} \int_0^1 K_\Delta(t,s)\eta_j(s)ds, \qquad t \in [0,1]. \tag{26}$$

We prove that η_j is of finite variation on $[0,1]$. We recall that a function f is a function of finite variation on an interval $[a,b]$ if the quantity

$$V_f := \sup_{a \leq t_0 < t_1 < \ldots < t_p \leq b} \left\{ \sum_{\ell=1}^{p} |f(t_\ell) - f(t_{\ell-1})| \right\}$$

is finite. For any grid $(t_0, \ldots, t_p)$ of $[0,1]$, we have:

$$\begin{aligned}
\sum_{\ell=1}^{p} |\eta_j(t_\ell) - \eta_j(t_{\ell-1})| &= \lambda_j^{-1} \sum_{\ell=1}^{p} \left| \int_0^1 (K_\Delta(s,t_\ell) - K_\Delta(s,t_{\ell-1}))\eta_j(s)ds \right| \\
&= \lambda_j^{-1} \sum_{\ell=1}^{p} \left| \int_0^1 \left(\mathbb{E}[F_\Delta(s)F_\Delta(t_\ell)] - \mathbb{E}[F_\Delta(s)F_\Delta(t_{\ell-1})]\right) \eta_j(s)ds \right| \\
&= \lambda_j^{-1} \sum_{\ell=1}^{p} \left| \int_0^1 \mathbb{E}[F_\Delta(s)(F_\Delta(t_\ell) - F_\Delta(t_{\ell-1}))]\eta_j(s)ds \right| \\
&\leq \lambda_j^{-1} \int_0^1 \mathbb{E}\left[|F_\Delta(s)| \sum_{\ell=1}^{p} |F_\Delta(t_\ell) - F_\Delta(t_{\ell-1})| \right] |\eta_j(s)|\, ds \\
&\leq \lambda_j^{-1} \int_0^1 \mathbb{E}\left[|F_\Delta(s)| \sum_{\ell=1}^{p} |\Delta(]t_{\ell-1}, t_\ell])| \right] |\eta_j(s)|\, ds \\
&\leq \lambda_j^{-1} \int_0^1 \mathbb{E}\left[|F_\Delta(s)| \sum_{\ell=1}^{p} (\Pi(]t_{\ell-1}, t_\ell]) + W(]t_{\ell-1}, t_\ell])) \right] |\eta_j(s)|\, ds \\
&\leq \lambda_j^{-1} \int_0^1 \mathbb{E}\left[|F_\Delta(s)| \left(\Pi([0,1]) + W([0,1])\right)\right] |\eta_j(s)|\, ds,
\end{aligned}$$

which is independent of the grid. We observe that

$$\begin{aligned}
\mathbb{E}\left[|F_\Delta(s)| \left(\Pi([0,1]) + W([0,1])\right)\right] &\leq \sqrt{\mathbb{E}\left[|F_\Delta(s)|^2\right] \times \mathbb{E}\left[(\Pi([0,1]) + W([0,1]))^2\right]} \\
&\leq 2\sqrt{\mathbb{E}\left[|F_\Delta(s)|^2\right] \mathbb{E}\left[\Pi^2([0,1])\right]}
\end{aligned}$$

and

$$\int_0^1 \mathbb{E}\left[|F_\Delta(s)| \left(\Pi([0,1]) + W([0,1])\right)\right] |\eta_j(s)|\, ds$$

$$\leq 2\|\eta_j\|_2\sqrt{\mathbb{E}\left[\Pi^2([0,1])\right]}\left(\int_0^1 \mathbb{E}\left[|F_\Delta(s)|^2\right]ds\right)^{1/2}$$
$$\leq 4\|\eta_j\|_2\mathbb{E}\left[\Pi^2([0,1])\right] < \infty.$$

Then, η_j is a function of finite variation. Since K_Δ is continuous, η_j is also continuous (and then right-continuous). We also have that $\eta_j(0) = 0$. Then, applying Proposition 4.4.3 of [18], we obtain that there exists a signed measure μ_j such that

$$\eta_j(t) = \mu_j([0,t]), \quad t \in [0,1].$$

From Fubini's Theorem we have that for any function $\varphi \in \mathcal{H}_0^1$

$$\int_0^1 \varphi'(t)\eta_j(t)dt = \int_0^1 \varphi'(t)\left(\int_0^t d\mu_j(s)\right)dt = \int_0^1 \left(\int_s^1 \varphi'(t)dt\right)d\mu_j(s)$$
$$= -\int_0^1 \varphi(s)d\mu_j(s), \tag{27}$$

since $\varphi(1) = 0$.

8.1.2. *Proof of Theorem 3.2.* By the Karhunen-Loève theorem (see e.g. [35], Theorem 7.3.5) applied to the stochastic process F_Π,

$$\lim_{J\to+\infty}\sup_{t\in[0,1]}\mathbb{E}\left[\left(F_\Pi(t) - F_W(t) - \sum_{j=1}^J \sqrt{\lambda_j}\xi_j\eta_j(t)\right)^2\right] = 0. \tag{28}$$

Furthermore, similarly to (27), we have, for all $\varphi \in \mathcal{H}_0^1$,

$$\langle \varphi, \Pi - W\rangle = -\langle F_\Pi - F_W, \varphi'\rangle.$$

Then, using (4),

$$\mathbb{E}\left[\left\|\Pi - W - \sum_{j=1}^J \sqrt{\lambda_j}\xi_j\mu_j\right\|_{\mathcal{H}^{-1}}^2\right] \leq \mathbb{E}\left[\sup_{\varphi\in\mathcal{H}_0^1, \|\varphi'\|\leq 1}\left(\langle\varphi, \Pi - W\rangle - \sum_{j=1}^J \sqrt{\lambda_j}\xi_j\langle\varphi,\mu_j\rangle\right)^2\right]$$
$$\leq \mathbb{E}\left[\sup_{\varphi\in\mathcal{H}_0^1, \|\varphi'\|\leq 1}\left\langle F_\Pi - F_W - \sum_{j=1}^J \sqrt{\lambda_j}\xi_j\eta_j, \varphi'\right\rangle^2\right]$$
$$\leq \mathbb{E}\left[\sup_{\varphi\in\mathcal{H}_0^1, \|\varphi'\|\leq 1}\|\varphi'\|^2\left\|F_\Pi - F_W - \sum_{j=1}^J \sqrt{\lambda_j}\xi_j\eta_j\right\|^2\right]$$
$$\leq \mathbb{E}\left[\int_0^1\left(F_\Pi(t) - F_W(t) - \sum_{j=1}^J \sqrt{\lambda_j}\xi_j\eta_j(t)\right)^2 dt\right]$$

and the result comes from (28).

8.1.3. *Proof of Theorem 3.4.* A direct consequence of Mercer's theorem is that

$$\left\langle K_\Delta, \frac{\partial^2\varphi}{\partial t\partial s}\right\rangle = \left\langle \sum_{j\geq 1}\lambda_j\eta_j\otimes\eta_j, \frac{\partial^2\varphi}{\partial t\partial s}\right\rangle = \sum_{j\geq 1}\lambda_j\left\langle \eta_j\otimes\eta_j, \frac{\partial^2\varphi}{\partial t\partial s}\right\rangle = \sum_{j\geq 1}\lambda_j\left\langle\varphi, \mu_j\otimes\mu_j\right\rangle.$$

Indeed, Fubini's theorem and the definition of μ_j implies

$$\begin{aligned}\left\langle \eta_j \otimes \eta_j, \frac{\partial^2 \varphi}{\partial t \partial s} \right\rangle &= \int_{[0,1]^2} \eta_j(s)\eta_j(t) \frac{\partial^2 \varphi}{\partial t \partial s}(s,t) ds dt \\ &= \int_0^1 \eta_j(s) \left(\int_0^1 \eta_j(t) \frac{\partial^2 \varphi}{\partial t \partial s}(s,t) dt \right) ds \\ &= \int_0^1 \eta_j(s) \left\langle \eta_j, \frac{\partial^2 \varphi}{\partial t \partial s}(s,\cdot) \right\rangle ds \\ &= -\int_0^1 \eta_j(s) \left\langle \frac{\partial \varphi}{\partial s}(s,\cdot), \mu_j \right\rangle ds \\ &= -\int_0^1 \eta_j(s) \int_0^1 \frac{\partial \varphi}{\partial s}(s,t) d\mu_j(t) ds \\ &= \langle \varphi, \mu_j \otimes \mu_j \rangle,\end{aligned}$$

using Fubini's theorem and the definition of μ_j a second time. Hence we can write

$$\left\langle \frac{\partial^2 K_\Delta}{\partial t \partial s}, \varphi \right\rangle = \sum_{j \geq 1} \lambda_j \langle \varphi, \mu_j \otimes \mu_j \rangle.$$

Now remark also that,

$$C_\Delta = \frac{\partial^2 K_\Delta}{\partial t \partial s}.$$

Indeed, for a function $\varphi \in \mathcal{H}_0^1$, we get

$$\begin{aligned}\int_0^1 F_\Pi(t)\varphi'(t)dt &= \int_0^1 \sum_{T \in N} \mathbf{1}_{T \leq t} \varphi'(t) dt = \sum_{T \in N} \int_T^1 \varphi'(t) dt \\ &= \sum_{T \in N} (\varphi(1) - \varphi(T)) = -\langle \Pi, \varphi \rangle.\end{aligned}$$

In particular

$$\mathbb{E}\left[\int_0^1 F_\Pi(t)\varphi'(t)dt \right] = -\mathbb{E}[\langle \Pi, \varphi \rangle] = -\langle W, \varphi \rangle.$$

This implies

$$\begin{aligned}\int_{[0,1]^2} K_\Delta(s,t) \frac{\partial^2 \varphi}{\partial t \partial s}(s,t) ds dt &= \int_{[0,1]^2} \mathbb{E}[F_\Delta(s) F_\Delta(t)] \frac{\partial^2 \varphi}{\partial t \partial s}(s,t) ds dt \\ &= \mathbb{E}\left[\int_{[0,1]^2} F_\Pi(s) F_\Pi(t) \frac{\partial^2 \varphi}{\partial t \partial s}(s,t) ds dt \right] \\ &\quad - \int_{[0,1]^2} F_W(s) F_W(t) \frac{\partial^2 \varphi}{\partial t \partial s}(s,t) ds dt \\ &= \mathbb{E}\left[\int_0^1 F_\Pi(s) \left(\int_0^1 F_\Pi(t) \frac{\partial^2 \varphi}{\partial t \partial s}(s,t) dt \right) ds \right] \\ &\quad - \int_0^1 F_W(s) \mathbb{E}\left[\int_0^1 F_\Pi(t) \frac{\partial^2 \varphi}{\partial t \partial s}(s,t) dt \right] ds\end{aligned}$$

$$= -\mathbb{E}\left[\int_0^1 F_\Pi(s) \sum_{T\in N} \frac{\partial \varphi}{\partial s}(s,T)ds\right] + \int_0^1 F_W(s) \int_0^1 \frac{\partial \varphi}{\partial s}(s,t)W(t)dtds$$

$$= \mathbb{E}\left[\sum_{T,T'\in N} \varphi(T',T)\right] - \langle \varphi, W\otimes W\rangle$$

$$= \langle C_\Delta, \varphi\rangle.$$

Then, for all $\varphi \in \mathcal{H}_0^2$, from Cauchy-Schwarz's inequality we deduce

$$\left|\left\langle \sum_{j=1}^J \lambda_j \mu_j \otimes \mu_j - C_\Delta, \varphi\right\rangle\right| = \left|\left\langle \sum_{j=1}^J \lambda_j \eta_j \otimes \eta_j - K_\Delta, \frac{\partial^2 \varphi}{\partial s \partial t}\right\rangle\right| \le \left\| K_\Delta - \sum_{j=1}^J \lambda_j \eta_j \otimes \eta_j\right\| \left\| \frac{\partial^2 \varphi}{\partial s\partial t}\right\|,$$

which implies

$$\left\| \sum_{j=1}^J \lambda_j \mu_j \otimes \mu_j - C_\Delta\right\|_{\mathcal{H}^{-2}} \le \left\| \sum_{j=1}^J \lambda_j \eta_j \otimes \eta_j - K_\Delta\right\|,$$

and Mercer theorem implies the expected result.

8.2. *Proofs of Section 4.*

8.2.1. *Proof of Theorem 4.1.* First, assume that F_{μ_j} is an eigenfunction of the operator Γ_Δ associated to the eigenvalue λ_j. We have:

$$\lambda_j F_{\mu_j}(t) = \int_0^1 K_\Delta(t,s)F_{\mu_j}(s)ds, \quad t\in[0,1]. \tag{29}$$

We have for any $(t,t')\in[0,1]^2$, since $\lambda_j > 0$,

$$|F_{\mu_j}(t) - F_{\mu_j}(t')| \le \|F_{\mu_j}\|_2 \times \lambda_j^{-1}\left(\int_0^1 \big(K_\Delta(t,s) - K_\Delta(t',s)\big)^2 ds\right)^{1/2}.$$

And since K_Δ is continuous, $|F_{\mu_j}(t) - F_{\mu_j}(t')|$ goes to 0 when $t'\to t$. So F_{μ_j} is continuous. Equation (29) gives

$$\lambda_j F_{\mu_j}(t) = \int_0^1 \Big(\int_0^{\min\{s;t\}} w(u)du\Big)F_{\mu_j}(s)ds$$

$$= \int_0^t F_{\mu_j}(s)\Big(\int_0^s w(u)du\Big)ds + \int_t^1 F_{\mu_j}(s)\Big(\int_0^t w(u)du\Big)ds. \tag{30}$$

Now, consider

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

Since F_{μ_j} is continuous, H_j is a differentiable function and

$$H_j'(t) = -F_{\mu_j}(t).$$

Actually, previous computations show that F_{μ_j} is an infinitely-differentiable function and differentiating Equation (30) gives

$$\lambda_j F'_{\mu_j}(t) = w(t)\int_t^1 F_{\mu_j}(u)du \tag{31}$$

and then H_j is solution of a second-order differential equation,

$$-\lambda_j H''_j(t) = H_j(t)w(t), \tag{32}$$

with boundary conditions:

$$H_j(1) = 0, \quad H'_j(0) = -F_{\mu_j}(0) = 0.$$

Therefore, (λ_j, H_j) is solution of the system (12).

Conversely, assume that (λ_j, H_j) is a solution of (12). Then, setting

$$F_{\mu_j}(s) = -H'_j(s), \quad s \in [0,1],$$

we obtain, by using integration by part, for any $t \in [0,1]$,

$$\begin{aligned}\int_0^1 K_\Delta(t,s)F_{\mu_j}(s)ds &= \int_0^t F_{\mu_j}(s)\Big(\int_0^s w(u)du\Big)ds + \int_t^1 F_{\mu_j}(s)\Big(\int_0^t w(u)du\Big)ds\\ &= -\int_0^t H'_j(s)\Big(\int_0^s w(u)du\Big)ds - \Big(\int_0^t w(u)du\Big)\int_t^1 H'_j(s)ds\\ &= -H_j(t)\int_0^t w(u)du + \int_0^t H_j(s)w(s)ds - \Big(\int_0^t w(u)du\Big)\Big(H_j(1) - H_j(t)\Big)\\ &= \int_0^t H_j(s)w(s)ds,\end{aligned}$$

since $H_j(1) = 0$. Then, since $H'_j(0) = 0$, for any $t \in [0,1]$,

$$\begin{aligned}\int_0^1 K_\Delta(t,s)F_{\mu_j}(s)ds &= -\lambda_j \int_0^t H''_j(s)ds\\ &= -\lambda_j H'_j(t) = \lambda_j F_{\mu_j}(t).\end{aligned}$$

Then, (λ_j, F_{μ_j}) are eigenelements of the operator of kernel K_Δ.

8.2.2. *Proofs of Theorem 4.8 and of the result of Remark 4.9.* In the sequel, we set

$$\theta = \beta - \alpha, \quad C_1 = \frac{\beta^3 w_0}{(\beta-\alpha)^3}, \quad C_2 = -\frac{w_0\alpha\beta(2\beta-\alpha)}{2(\beta-\alpha)^4}.$$

We observe

$$\begin{aligned}(\alpha-\beta)^{-2}(2C_2(\alpha-\beta) - C_1) &= (\alpha-\beta)^{-2}\Big(\frac{w_0\alpha\beta(2\beta-\alpha)}{(\beta-\alpha)^3} - \frac{\beta^3 w_0}{(\beta-\alpha)^3}\Big)\\ &= w_0\beta(\beta-\alpha)^{-5}\big(\alpha(2\beta-\alpha) - \beta^2\big)\\ &= -\frac{w_0\beta}{(\beta-\alpha)^3} = -C_1\beta^{-2}\end{aligned}$$

and

$$2C_2\theta + C_1 = C_1\beta^{-2}\theta^2. \tag{33}$$

Using Equation (17), we have, by symmetry of the bivariate kernel

$$K_\Delta(s,t) = C_1 \min\{s;t\} - C_2\Big(-1 - e^{-\theta|t-s|} + e^{-\theta t} + e^{-\theta s}\Big), \quad s,t \in [0,1].$$

Furthermore, using definitions of C_1 and C_2, (λ_j, H_j) is solution of (18) if and only if we have:

$$\begin{cases} -\lambda_j H_j''(s) = C_1\beta^{-2}\theta^2 H_j(s) - C_2\theta^2 \displaystyle\int_0^1 e^{-\theta|t-s|} H_j(t)dt, \, s \in (0,1), \\ H_j(1) = 0, \quad H_j'(0) = 0 \end{cases}$$

Then, let us express $\Gamma_\Delta F_{\mu_j}$. For any $s \in [0,1]$,

$$\begin{aligned}
\Gamma_\Delta F_{\mu_j}(s) &= \int_0^1 K_\Delta(s,t) F_{\mu_j}(t)dt \\
&= \int_0^1 \left(C_1 \min\{s,t\} - C_2 \left(-1 - e^{-\theta|t-s|} + e^{-\theta t} + e^{-\theta s}\right)\right) F_{\mu_j}(t)dt \\
&= C_1 \int_0^s tF_{\mu_j}(t)dt + C_1 s \int_s^1 F_{\mu_j}(t)dt + C_2(1 - e^{-\theta s}) \int_0^1 F_{\mu_j}(t)dt \\
&\quad - C_2 \int_0^1 e^{-\theta t} F_{\mu_j}(t)dt + C_2 e^{\theta s} \int_s^1 e^{-\theta t} F_{\mu_j}(t)dt + C_2 e^{-\theta s} \int_0^s e^{\theta t} F_{\mu_j}(t)dt.
\end{aligned}$$

We have

$$H_j(s) = \int_s^1 F_{\mu_j}(t)dt, \quad s \in [0,1].$$

Since K_Δ is continuous, the eigenfunctions are continuous as well (eigenvalues are positive). Since F_{μ_j} is continuous, we have

$$F_{\mu_j}(s) = -H_j'(s), \quad s \in [0,1]$$

with boundary conditions

$$H_j(1) = 0, \quad H_j'(0) = 0.$$

Then, we can write

$$\Gamma_\Delta F_{\mu_j}(s) = (I) + (II) - (III) + (IV) + (V)$$

with

$$\begin{aligned}
(I) &= C_1 \int_0^s tF_{\mu_j}(t)dt + C_1 s \int_s^1 F_{\mu_j}(t)dt \\
&= C_1 \int_0^s t(-H_j'(t))dt + C_1 s H_j(s) \\
&= C_1 \int_0^s H_j(t)dt,
\end{aligned}$$

$$\begin{aligned}
(II) &= C_2(1 - e^{-\theta s}) \int_0^1 F_{\mu_j}(t)dt \\
&= C_2(1 - e^{-\theta s}) H_j(0),
\end{aligned}$$

$$
\begin{aligned}
(III) &= C_2 \int_0^1 e^{-\theta t} F_{\mu_j}(t)dt \\
&= C_2 \int_0^1 e^{-\theta t}(-H_j'(t))dt \\
&= C_2 H_j(0) - C_2\theta \int_0^1 e^{-\theta t} H_j(t)dt
\end{aligned}
$$

and

$$
-(III) = -C_2 H_j(0) + C_2\theta \int_0^1 e^{-\theta t} H_j(t)dt.
$$

Now,

$$
\begin{aligned}
(IV) &= C_2 e^{\theta s} \int_s^1 e^{-\theta t} F_{\mu_j}(t)dt \\
&= C_2 e^{\theta s} \int_s^1 e^{-\theta t}(-H_j'(t))dt \\
&= C_2 e^{\theta s}\Big(e^{-\theta s} H_j(s) - \theta \int_s^1 e^{-\theta t} H_j(t)dt\Big) \\
&= C_2 H_j(s) - C_2\theta e^{\theta s} \int_s^1 e^{-\theta t} H_j(t)dt
\end{aligned}
$$

and, finally,

$$
\begin{aligned}
(V) &= C_2 e^{-\theta s} \int_0^s e^{\theta t} F_{\mu_j}(t)dt \\
&= C_2 e^{-\theta s} \int_0^s e^{\theta t}(-H_j'(t))dt \\
&= C_2 e^{-\theta s}\big(H_j(0) - e^{\theta s} H_j(s)\big) + C_2\theta e^{-\theta s} \int_0^s e^{\theta t} H_j(t)dt \\
&= C_2 e^{-\theta s} H_j(0) - C_2 H_j(s) + C_2\theta e^{-\theta s} \int_0^s e^{\theta t} H_j(t)dt.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\Gamma_\Delta F_{\mu_j}(s) &= (I) + (II) - (III) + (IV) + (V) \\
&= C_1 \int_0^s H_j(t)dt + C_2(1 - e^{-\theta s}) H_j(0) \\
&\quad - C_2 H_j(0) + C_2\theta \int_0^1 e^{-\theta t} H_j(t)dt + C_2 H_j(s) - C_2\theta e^{\theta s} \int_s^1 e^{-\theta t} H_j(t)dt \\
&\quad + C_2 e^{-\theta s} H_j(0) - C_2 H_j(s) + C_2\theta e^{-\theta s} \int_0^s e^{\theta t} H_j(t)dt \\
&= C_1 \int_0^s H_j(t)dt + C_2\theta \int_0^1 e^{-\theta t} H_j(t)dt - C_2\theta e^{\theta s} \int_s^1 e^{-\theta t} H_j(t)dt \\
&\quad + C_2\theta e^{-\theta s} \int_0^s e^{\theta t} H_j(t)dt.
\end{aligned}
$$

Setting

$$U(s) = e^{-\theta s}\int_0^s e^{\theta t}H_j(t)dt - e^{\theta s}\int_s^1 e^{-\theta t}H_j(t)dt, \quad s\in[0,1],$$

we finally obtain that for any $s\in[0,1]$,

$$\Gamma_\Delta F_{\mu_j}(s) = C_1\int_0^s H_j(t)dt + C_2\theta\int_0^1 e^{-\theta t}H_j(t)dt + C_2\theta U(s). \tag{34}$$

We then obtain that

$$\Gamma_\Delta F_{\mu_j}(s) = \lambda_j F_{\mu_j}(s), \quad s\in[0,1] \tag{35}$$

if and only if

$$\lambda_j F_{\mu_j}(s) = C_1\int_0^s H_j(t)dt + C_2\theta\int_0^1 e^{-\theta t}H_j(t)dt + C_2\theta U(s), \quad s\in[0,1].$$

In the sequel, we use the following lemma.

LEMMA 8.1. *Setting*

$$Z(s) = \int_0^1 e^{-\theta|t-s|}H_j(t)dt, \tag{36}$$

we have for any $s\in[0,1]$*,*

$$U(s) = U(0) + 2\int_0^s H_j(t)dt - \theta\int_0^s Z(t)dt.$$

PROOF OF THE LEMMA. We have for any $s\in[0,1]$,

$$\begin{aligned} U'(s) &= -\theta e^{-\theta s}\int_0^s e^{\theta t}H_j(t)dt + H_j(s) - \theta e^{\theta s}\int_s^1 e^{-\theta t}H_j(t)dt + H_j(s)\\ &= 2H_j(s) - \theta Z(s). \end{aligned}$$

This provides the result of the lemma. □

Now, we assume that (35) is true. We have

$$\begin{aligned} \lambda_j F'_{\mu_j}(s) &= C_1H_j(s) + C_2\theta U'(s)\\ &= C_1H_j(s) + C_2\theta\Big(2H_j(s) - \theta Z(s)\Big)\\ &= \frac{C_1\theta^2}{\beta^2}H_j(s) - C_2\theta^2 Z(s), \end{aligned}$$

by using (33) and the lemma. This gives

$$-\lambda_j H''_j(s) = \frac{C_1\theta^2}{\beta^2}H_j(s) - C_2\theta^2 Z(s), \quad s\in[0,1], \tag{37}$$

meaning that (λ_j, H_j) is solution of (18).

Conversely, assume that (λ_j, H_j) is solution of (18). We prove that (35) is satisfied. Indeed, using (34), the result of the lemma and

$$U(0) = -\int_0^1 e^{-\theta t}H_j(t)dt,$$

we have for any $s \in [0,1]$,

$$\begin{aligned}\Gamma_\Delta F_{\mu_j}(s) &= C_1 \int_0^s H_j(t)dt + C_2\theta \int_0^1 e^{-\theta t} H_j(t)dt + C_2\theta U(s) \\ &= C_1 \int_0^s H_j(t)dt + C_2\theta \int_0^1 e^{-\theta t} H_j(t)dt + C_2\theta\Big(U(0) + 2\int_0^s H_j(t)dt - \theta \int_0^s Z(t)dt\Big) \\ &= \frac{C_1\theta^2}{\beta^2} \int_0^s H_j(t)dt - C_2\theta^2 \int_0^s Z(t)dt.\end{aligned}$$

by using (33). Since (λ_j, H_j) is solution of (18), we finally obtain for any $s \in [0,1]$,

$$\begin{aligned}\Gamma_\Delta F_{\mu_j}(s) &= -\lambda_j \int_0^s H_j''(t)dt \\ &= \lambda_j \big(H_j'(0) - H_j'(s)\big) \\ &= \lambda_j F_{\mu_j}(s)\end{aligned}$$

and the result is proved.

Now, we prove the result of Remark 4.9. Using Equation (37), we obtain

$$-\lambda_j H_j^{(4)}(s) = \frac{C_1\theta^2}{\beta^2} H_j''(s) - C_2\theta^2 Z''(s), \quad s \in [0,1].$$

By definition of Z given in (36), we have for any $s \in [0,1]$,

$$Z'(s) = -\theta e^{-\theta s} \int_0^s e^{\theta u} H_j(u)du + \theta e^{\theta s} \int_s^1 e^{-\theta u} H_j(u)du$$

and

$$\begin{aligned}Z''(s) &= \theta^2 e^{-\theta s} \int_0^s e^{\theta u} H_j(u)du + \theta^2 e^{\theta s} \int_s^1 e^{-\theta u} H_j(u)du - 2\theta H_j(s) \\ &= \theta^2 Z(s) - 2\theta H_j(s).\end{aligned}$$

Now, we obtain, still using Equations (37) and (33),

$$\begin{aligned}-\lambda_j H_j^{(4)}(s) &= \frac{C_1\theta^2}{\beta^2} H_j''(s) - C_2\theta^4 Z(s) + 2C_2\theta^3 H_j(s) \\ &= \frac{C_1\theta^2}{\beta^2} H_j''(s) - \lambda_j\theta^2 H_j''(s) + \Big(2C_2\theta^3 - \frac{C_1\theta^4}{\beta^2}\Big) H_j(s) \\ &= \frac{C_1\theta^2}{\beta^2} H_j''(s) - \lambda_j\theta^2 H_j''(s) - C_1\theta^2 H_j(s).\end{aligned}$$

Finally, we obtain for any $s \in [0,1]$,

$$-\lambda_j H_j^{(4)}(s) + \lambda_j\theta^2 H_j''(s) = \frac{C_1\theta^2}{\beta^2}\Big(H_j''(s) - \beta^2 H_j(s)\Big). \tag{38}$$

In particular, if $\lambda_j = 0$, then

$$H_j(s) = c_1 e^{\beta s} + c_2 e^{-\beta s}, \quad s \in [0,1],$$

for c_1 and c_2 two constants. The boundary condition $H_j'(0) = 0$ gives $c_1 = c_2$ and $H_j(1) = 0$ gives $c_1 = c_2 = 0$. Then $\lambda_j = 0$ entails $H_j = 0$ and then $F_{\mu_j} = 0$. We have assumed that F_{μ_j} is continuous, which allows to use previous computations.

8.2.3. *Proof of Theorem 4.10.* We set

$$w_1 = \frac{\beta w_0}{\beta - \alpha}, \quad c = \frac{w_0 \alpha \beta (2\beta - \alpha)}{2(\beta - \alpha)^2}, \quad \theta = \beta - \alpha$$

and

$$I(y)(t) = \frac{c}{w_1} \int_0^1 e^{-\theta |t-s|} y(s) ds, \quad t \in [0,1].$$

In the sequel, we assume that (19) is satisfied, which is equivalent to

$$\frac{2c}{w_1}\Big(2 + \frac{3 - 2e^{-\theta/2} - e^{-\theta}}{\theta}\Big) < 1. \tag{39}$$

Now (18) writes

$$\begin{cases} -\lambda y''(t) = w_1 y(t) + w_1 I(y(t)), t \in (0,1), \\ y(1) = 0, \quad y'(0) = 0. \end{cases}$$

Let H_j a solution of this ODE as expressed in Theorem 4.8. We have

$$\begin{cases} H_j''(t) + \rho_j^2 H_j(t) = -\rho_j^2 I(H_j)(t), t \in (0,1), \\ H_j(1) = 0, \quad H_j'(0) = 0 \end{cases}$$

with

$$\rho_j := \sqrt{\frac{w_1}{\lambda_j}} \overset{j \to +\infty}{\longrightarrow} +\infty. \tag{40}$$

We write H_j and H_j' as

$$\begin{cases} H_j(t) = \quad a(t)\cos(\rho_j t) + b(t)\sin(\rho_j t), \\ H_j'(t) = -\rho_j a(t)\sin(\rho_j t) + \rho_j b(t)\cos(\rho_j t), \end{cases}$$

with a and b C^2-functions that must satisfy

$$\begin{cases} a'(t)\cos(\rho_j t) + b'(t)\sin(\rho_j t) = 0 \\ -a'(t)\rho_j \sin(\rho_j t) + b'(t)\rho_j \cos(\rho_j t) = -\rho_j^2 I(H_j)(t). \end{cases}$$

This means that

$$\begin{pmatrix} \cos(\rho_j t) & \sin(\rho_j t) \\ -\sin(\rho_j t) & \cos(\rho_j t) \end{pmatrix} \begin{pmatrix} a'(t) \\ b'(t) \end{pmatrix} = \begin{pmatrix} 0 \\ -\rho_j I(H_j)(t) \end{pmatrix},$$

or equivalently

$$\begin{pmatrix} a'(t) \\ b'(t) \end{pmatrix} = \begin{pmatrix} \cos(\rho_j t) & -\sin(\rho_j t) \\ \sin(\rho_j t) & \cos(\rho_j t) \end{pmatrix} \begin{pmatrix} 0 \\ -\rho_j I(H_j)(t) \end{pmatrix}.$$

Therefore

$$\begin{cases} a(t) = a + \rho_j \int_0^t \sin(\rho_j s) I(H_j)(s) ds \\ b(t) = b - \rho_j \int_0^t \cos(\rho_j s) I(H_j)(s) ds, \end{cases}$$

with a and b two constants. Finally, for any $t \in [0,1]$,

$$H_j(t) = a\cos(\rho_j t) + b\sin(\rho_j t) + \rho_j \cos(\rho_j t) \int_0^t \sin(\rho_j s) I(H_j)(s) ds - \rho_j \sin(\rho_j t) \int_0^t \cos(\rho_j s) I(H_j)(s) ds,$$

and

$$H_j'(t) = -a\rho_j \sin(\rho_j t) + b\rho_j \cos(\rho_j t) - \rho_j^2 \sin(\rho_j t) \int_0^t \sin(\rho_j s) I(H_j)(s) ds - \rho_j^2 \cos(\rho_j t) \int_0^t \cos(\rho_j s) I(H_j)(s) ds,$$

with a and b two constants. The boundary conditions $H_j'(0)=0$ implies that $b=0$. We then obtain that H_j must write

(41)

$$H_j(t)=a\cos(\rho_j t)+\rho_j\cos(\rho_j t)\int_0^t \sin(\rho_j s)I(H_j)(s)ds-\rho_j\sin(\rho_j t)\int_0^t \cos(\rho_j s)I(H_j)(s)ds,$$

and

$$H_j'(t)=-a\rho_j\sin(\rho_j t)-\rho_j^2\sin(\rho_j t)\int_0^t \sin(\rho_j s)I(H_j)(s)ds-\rho_j^2\cos(\rho_j t)\int_0^t \cos(\rho_j s)I(H_j)(s)ds.$$

In particular, we have

$$\begin{aligned}H_j(t)&=a\cos(\rho_j t)-\rho_j\int_0^t \sin(\rho_j(t-s))I(H_j)(s)ds\\&=a\cos(\rho_j t)-\int_0^{\rho_j t}\sin(u)I(H_j)\Big(t-\frac{u}{\rho_j}\Big)du.\end{aligned}$$

We denote $J_j:\mathbb{L}_\infty[0,1]\mapsto\mathbb{L}_\infty[0,1]$ the linear operator defined for $f\in\mathbb{L}_\infty[0,1]$, for $t\in[0,1]$, by

$$\begin{aligned}J_j(f)(t)&:=\rho_j\cos(\rho_j t)\int_0^t \sin(\rho_j s)I(f)(s)ds-\rho_j\sin(\rho_j t)\int_0^t\cos(\rho_j s)I(f)(s)ds\\&=-\rho_j\int_0^t\sin(\rho_j(t-s))I(f)(s)ds\\&=-\int_0^{\rho_j t}\sin(u)I(f)\Big(t-\frac{u}{\rho_j}\Big)du.\end{aligned}$$

Finally, if H_j satisfies Equation (18), H_j satisfies

(42)
$$H_j(t)=a\cos(\rho_j t)+J_j(H_j)(t),\quad t\in[0,1].$$

Conversely, if H_j satisfies Equation (42), for $t\in[0,1]$,

$$\begin{aligned}H_j''(t)&=-a\rho_j^2\cos(\rho_j t)+J_j''(H_j)(t)\\&=-a\rho_j^2\cos(\rho_j t)-\rho_j^3\cos(\rho_j t)\int_0^t\sin(\rho_j s)I(H_j)(s)ds-\rho_j^2\sin^2(\rho_j t)I(H_j)(t)\\&\qquad+\rho_j^3\sin(\rho_j t)\int_0^t\cos(\rho_j s)I(H_j)(s)ds-\rho_j^2\cos^2(\rho_j t)I(H_j)(t)\\&=-a\rho_j^2\cos(\rho_j t)-\rho_j^2J_j(H_j)(t)-\rho_j^2I(H_j)(t)\\&=-\rho_j^2H_j(t)-\rho_j^2I(H_j)(t).\end{aligned}$$

It means that if H_j satisfies Equation (42), then it satisfies (18). We have proved:

(43)
$$H_j \text{ satisfies Equation (18)} \iff H_j \text{ satisfies Equation (42)}.$$

Now, let us deal with $F_{\mu_j}(t)=-H_j'(t)$, for $t\in[0,1]$. We have:

$$\begin{aligned}F_{\mu_j}(t)&=-H_j'(t)\\&=a\rho_j\sin(\rho_j t)+\rho_j^2\sin(\rho_j t)\int_0^t\sin(\rho_j s)I(H_j)(s)ds+\rho_j^2\cos(\rho_j t)\int_0^t\cos(\rho_j s)I(H_j)(s)ds\end{aligned}$$

$$= a\rho_j \sin(\rho_j t) + \rho_j^2 \int_0^t \cos(\rho_j(t-s)) I(H_j)(s) ds$$
$$= a\rho_j \sin(\rho_j t) + \rho_j \int_0^{\rho_j t} \cos(u) I\big(H_j\big)\Big(t - \frac{u}{\rho_j}\Big) du$$
$$= a\rho_j \sin(\rho_j t) + \frac{c\rho_j}{w_1} \int_0^{\rho_j t} \cos(u) \Big(\int_0^1 e^{-\theta\left|t-\rho_j^{-1}u-s\right|} H_j(s) ds\Big) du.$$

It gives

$$F_{\mu_j}(t) = a\rho_j \sin(\rho_j t) + \rho_j K_j(H_j)(t), \quad t \in [0,1], \tag{44}$$

with for any bounded function f,

$$K_j(f)(t) = \frac{c}{w_1} \int_0^{\rho_j t} \cos(u) \Big(\int_0^1 e^{-\theta\left|t-\rho_j^{-1}u-s\right|} f(s) ds\Big) du.$$

LEMMA 8.2. *We set*

$$|||J_j|||_\infty := \sup_{f:\ \|f\|_\infty = 1} \|J_j(f)\|_\infty.$$

If (39) *is satisfied, we have*

$$|||J_j|||_\infty < \frac{1}{2}.$$

Furthermore, there exist c_1, c_2 and c_3 three positive constants only depending on c, θ and w_1 such that

$$\|J_j(\cos(\rho_j \cdot))\|_\infty \le c_1 \rho_j^{-1}, \quad \|K_j(\cos(\rho_j \cdot))\|_\infty \le c_2 \rho_j^{-1} \quad \textit{and} \quad |||K_j|||_\infty \le c_3.$$

PROOF OF LEMMA 8.2. We have

$$-J_j(f)(t) = \int_0^{\rho_j t} \sin(u) I\big(f\big)\Big(t - \frac{u}{\rho_j}\Big) du$$
$$= \frac{c}{w_1} \int_0^{\rho_j t} \sin(u) \Big(\int_0^1 e^{-\theta\left|t-\rho_j^{-1}u-s\right|} f(s) ds\Big) du$$
$$=: \mathcal{I}m\bigg(J^{(1)}(f)(t) + J^{(2)}(f)(t) + J^{(3)}(f)(t)\bigg),$$

with

$$J^{(1)}(f)(t) = \frac{c}{w_1} \int_0^{\rho_j t} e^{iu} \Big(\int_0^{t-\rho_j^{-1}u} e^{-\theta\left(t-\rho_j^{-1}u-s\right)} f(s) ds\Big) du$$
$$= \frac{c}{w_1} \int_0^t e^{-\theta(t-s)} f(s) \Big(\int_0^{\rho_j(t-s)} e^{(i+\theta\rho_j^{-1})u} du\Big) ds$$
$$= \frac{c}{w_1(i+\theta\rho_j^{-1})} \int_0^t e^{-\theta(t-s)} f(s) \Big(e^{i\rho_j(t-s)+\theta(t-s)} - 1\Big) ds$$
$$= \frac{c(-i+\theta\rho_j^{-1})}{w_1\sqrt{1+\theta^2\rho_j^{-2}}} \int_0^t f(s) \Big(e^{i\rho_j(t-s)} - e^{-\theta(t-s)}\Big) ds,$$

$$J^{(2)}(f)(t) = \frac{c}{w_1}\int_0^{\rho_j t} e^{iu}\Big(\int_{t-\rho_j^{-1}u}^{t} e^{\theta\left(t-\rho_j^{-1}u-s\right)} f(s)ds\Big)du$$
$$= \frac{c}{w_1}\int_0^{t} e^{\theta(t-s)} f(s)\Big(\int_{\rho_j(t-s)}^{\rho_j t} e^{(i-\theta\rho_j^{-1})u}du\Big)ds$$
$$= \frac{c}{w_1(i-\theta\rho_j^{-1})}\int_0^{t} e^{\theta(t-s)} f(s)\Big(e^{i\rho_j t-\theta t} - e^{i\rho_j(t-s)-\theta(t-s)}\Big)ds$$
$$= \frac{c(-i-\theta\rho_j^{-1})}{w_1\sqrt{1+\theta^2\rho_j^{-2}}}\int_0^{t} f(s)\Big(e^{i\rho_j t-\theta s} - e^{i\rho_j(t-s)}\Big)ds$$

and

$$J^{(3)}(f)(t) = \frac{c}{w_1}\int_0^{\rho_j t} e^{iu}\Big(\int_{t}^{1} e^{\theta\left(t-\rho_j^{-1}u-s\right)} f(s)ds\Big)du$$
$$= \frac{c}{w_1}\int_t^{1} e^{\theta(t-s)} f(s)\Big(\int_{0}^{\rho_j t} e^{(i-\theta\rho_j^{-1})u}du\Big)ds$$
$$= \frac{c}{w_1(i-\theta\rho_j^{-1})}\int_t^{1} e^{\theta(t-s)} f(s)\Big(e^{i\rho_j t-\theta t} - 1\Big)ds$$
$$= \frac{c(-i-\theta\rho_j^{-1})}{w_1\sqrt{1+\theta^2\rho_j^{-2}}}\int_t^{1} f(s)\Big(e^{i\rho_j t-\theta s} - e^{\theta(t-s)}\Big)ds.$$

This yields

(45)

$$-J_j(f)(t) = \frac{2c\theta\rho_j^{-1}}{w_1\sqrt{1+\theta^2\rho_j^{-2}}}\mathcal{I}m\left(\int_0^{t} f(s)e^{i\rho_j(t-s)}ds\right) + \frac{c}{w_1\sqrt{1+\theta^2\rho_j^{-2}}}\int_0^{1} f(s)e^{-\theta|t-s|}ds$$
$$+\mathcal{I}m\left(\frac{c(-i-\theta\rho_j^{-1})}{w_1\sqrt{1+\theta^2\rho_j^{-2}}}\int_0^{1} f(s)e^{i\rho_j t-\theta s}ds\right).$$

We obtain for any $t\in[0,1]$,

$$|J_j(f)(t)| \le \Big(\frac{2c}{w_1} + \frac{c}{w_1}\int_0^1 e^{-\theta|t-s|}ds + \frac{c}{w_1}\int_0^1 e^{-\theta s}ds\Big)\|f\|_\infty$$
$$\le \Big(\frac{2c}{w_1} + \frac{ce^{-\theta t}}{w_1}\int_0^t e^{\theta s}ds + \frac{ce^{\theta t}}{w_1}\int_t^1 e^{-\theta s}ds + \frac{c}{w_1}\int_0^1 e^{-\theta s}ds\Big)\|f\|_\infty$$
$$\le \frac{c}{w_1}\Big(2 + \frac{3-e^{-\theta t}-e^{\theta(t-1)}-e^{-\theta}}{\theta}\Big)\|f\|_\infty$$
$$\le \frac{c}{w_1}\Big(2 + \frac{3-2e^{-\theta/2}-e^{-\theta}}{\theta}\Big)\|f\|_\infty$$

and

$$|||J_j|||_\infty \le \frac{c}{w_1}\Big(2 + \frac{3-2e^{-\theta/2}-e^{-\theta}}{\theta}\Big).$$

This yields the first result.

For the second result, observe that with $f = \cos(\rho_j \cdot)$, we obtain from (45),

$$|J_j(\cos(\rho_j\cdot))(t)| \leq \frac{2c\theta}{w_1\rho_j} + \frac{c}{w_1}\Big|\int_0^1 e^{i\rho_j s}e^{-\theta|t-s|}ds\Big| + \frac{c}{w_1}\Big|\int_0^1 e^{i\rho_j s}e^{-\theta s}ds\Big|.$$

But

$$\begin{aligned}\int_0^1 e^{i\rho_j s}e^{-\theta|t-s|}ds &= e^{-\theta t}\int_0^t e^{(i\rho_j+\theta)s}ds + e^{\theta t}\int_t^1 e^{(i\rho_j-\theta)s}ds \\ &= e^{-\theta t}\Big(\frac{e^{(i\rho_j+\theta)t}-1}{i\rho_j+\theta}\Big) + e^{\theta t}\Big(\frac{e^{i\rho_j-\theta}-e^{(i\rho_j-\theta)t}}{i\rho_j-\theta}\Big)\end{aligned}$$

and there exists c_1 only depending on c, θ and w_1 such that

$$\|J_j(\cos(\rho_j\cdot))\|_\infty \leq c_1\rho_j^{-1}.$$

For the third result, observe that

$$\begin{aligned}K_j(f)(t) &= \mathcal{R}e\Bigg(\frac{c}{w_1}\int_0^{\rho_j t} e^{iu}\Big(\int_0^1 e^{-\theta|t-\rho_j^{-1}u-s|}f(s)ds\Big)du\Bigg) \\ &= \mathcal{R}e\Bigg(J^{(1)}(f)(t) + J^{(2)}(f)(t) + J^{(3)}(f)(t)\Bigg) \\ &= \frac{2c\theta\rho_j^{-1}}{w_1\sqrt{1+\theta^2\rho_j^{-2}}}\mathcal{R}e\Bigg(\int_0^t f(s)e^{i\rho_j(t-s)}ds\Bigg) + \frac{c\theta\rho_j^{-1}}{w_1\sqrt{1+\theta^2\rho_j^{-2}}}\int_0^1 f(s)e^{-\theta|t-s|}ds \\ &\quad + \mathcal{R}e\Bigg(\frac{c(-i-\theta\rho_j^{-1})}{w_1\sqrt{1+\theta^2\rho_j^{-2}}}\int_0^1 f(s)e^{i\rho_j t-\theta s}ds\Bigg).\end{aligned}$$

Therefore, as previously, there exists c_2 only depending on c, θ and w_1 such that

$$\|K_j(\cos(\rho_j\cdot))\|_\infty \leq c_2\rho_j^{-1}.$$

and there exists c_3 only depending on c, θ and w_1 such that

$$|||K_j|||_\infty \leq c_3.$$

□

Now, we end the proof of the theorem and use Equivalence (43). Since J_j is a linear operator, for any integer N, we have, for $t \in [0,1]$,

$$\begin{aligned}H_j(t) &= a\cos(\rho_j t) + J_j(H_j)(t) \\ &= a\cos(\rho_j t) + aJ_j(\cos(\rho_j\cdot))(t) + (J_j\circ J_j)(H_j)(t) \\ &= a\sum_{k=0}^{N-1} J_j^{k\circ}(\cos(\rho_j\cdot))(t) + (J_j^{N\circ})(H_j)(t),\end{aligned}$$

with for any bounded function f,

$$J_j^{k\circ}(f) = (\underbrace{J_j\circ\cdots\circ J_j}_{k \text{ times}})(f).$$

Since H_j is bounded and $|||J_j|||_\infty < 1/2$ (see Lemma 8.2), we have

$$\limsup_{N\to+\infty} \|(J_j^{N\circ})(H_j)\|_\infty \le \limsup_{N\to+\infty} |||J_j|||_\infty^N \|H_j\|_\infty = 0.$$

Therefore

$$H_j(t) = a\sum_{k=0}^{+\infty} J_j^{k\circ}(\cos(\rho_j\cdot))(t), \quad t\in[0,1]. \tag{46}$$

This implies that, up to renormalization, we have at most one solution to the problem. Now, we obtain a solution of the system (18) if and only if the boundary condition $H_j(1) = 0$ is satisfied. Using (46), this is equivalent to

$$\cos(\rho_j) = -\sum_{k=1}^{+\infty} J_j^{k\circ}(\cos(\rho_j\cdot))(1).$$

We can write $\rho_j = (k_j - 1/2)\pi + u_j\pi$, with $k_j \in \mathbb{N}^*$ and $-1/2 < u_j \le 1/2$ (k_j and u_j are uniquely defined). Then, $H_j(1) = 0$ writes

$$(-1)^{k_j}\sin(u_j\pi) + \sum_{k=1}^{+\infty} J_j^{k\circ}(\cos(((k_j - 1/2)\pi + u_j\pi)\cdot))(1) = 0. \tag{47}$$

Since $|||J_j|||_\infty < 1/2$ (see Lemma 8.2), for k_j fixed, the left hand side is a continuous fonction of u_j taking positive and negative values when u_j describes the interval $(-1/2;1/2]$. Therefore, for any positive integer k_j, there exists $u_j \in (-1/2;1/2]$ such that Equation (47) is satisfied. Now, let us give a control of u_j. We have:

$$2|u_j| \le |\sin(u_j\pi)| \le \sum_{k=1}^{+\infty} |||J_j|||_\infty^{k-1} \times \|J_j(\cos(\rho_j\cdot))\|_\infty \le \sum_{k=1}^{+\infty}\left(\frac{1}{2}\right)^{k-1} \times \frac{c_1}{\rho_j} = \frac{2c_1}{\rho_j}.$$

leading to

$$|u_j| \le \frac{c_1}{\rho_j}.$$

These computations hold for any integer k_j and in particular, we can consider $k_j = j$. In this case, we obtain

$$\rho_j = (j-1/2)\pi + u_j\pi$$

and, using (40),

$$\lambda_j = \frac{w_1}{\rho_j^2} = \frac{w_1}{\big((j-1/2)\pi + u_j\pi\big)^2} = \frac{w_1}{\big(j\pi - \pi/2\big)^2}\Big(1 + \frac{u_j}{(j-1/2)}\Big)^{-2}.$$

We have finally proved that for any $j \ge 1$, System (18) has a solution (λ_j, H_j) and there exists a constant C_1 depending only on c, θ and w_1 such that for any $j \ge 1$,

$$\Big|\lambda_j - \frac{w_1}{\big(j\pi - \pi/2\big)^2}\Big| \le C_1 j^{-4}.$$

Now, it remains to determine F_{μ_j}. Using (44), we have for any $t \in [0,1]$,

$$\begin{aligned} F_{\mu_j}(t) &= a\rho_j\sin(\rho_j t) + \rho_j K_j(H_j)(t) \\ &= a\rho_j\sin(\rho_j t) + a\rho_j K_j(\cos(\rho_j\cdot))(t) + a\rho_j\sum_{k=1}^{+\infty} K_j\big(J_j^{k\circ}(\cos(\rho_j\cdot))\big)(t). \end{aligned}$$

and we have to determine the value of a by using the property $\|F_{\mu_j}\|_2 = 1$. Setting

$$v_j(t) := \rho_j K_j(\cos(\rho_j\cdot))(t) + \rho_j \sum_{k=1}^{+\infty} K_j\big(J_j^{k\circ}(\cos(\rho_j\cdot))\big)(t), \quad t\in[0,1],$$

so that

$$F_{\mu_j}(t) = a\rho_j \sin(\rho_j t) + a v_j(t), \quad t\in[0,1], \tag{48}$$

we have, using Lemma 8.2,

$$\|v_j\|_\infty \le c_2 + c_3\rho_j \sum_{k=1}^{+\infty} 2^{1-k}\|J_j(\cos(\rho_j\cdot))\|_\infty \le c_2 + 2c_1c_3$$

and, since $\rho_j^{-1} = O(j^{-1})$,

$$1 = \int F_{\mu_j}^2(t)dt = a^2\rho_j^2 \int_0^1 \Big(\sin(\rho_j t) + \frac{v_j(t)}{\rho_j}\Big)^2 dt = a^2\rho_j^2 \times \Big(\frac{1}{2} + O(j^{-1})\Big).$$

This implies $a\rho_j = \sqrt{2} + O(j^{-1})$ and $a = \frac{\sqrt{2}}{\rho_j} + O(j^{-2})$. Using again Equation (48) yields the result stated in (21). Equation (46) shows that the eigenvalues $(\lambda_j)_{j\ge1}$ are simple.

8.3. *Proofs of Section 5.*

8.3.1. *Proof of Lemma 5.1.* We first prove some intermediate results. First remark that given $i = 1,\dots,n$, if $T\in N_i$, there exists $k\in\{0,\dots,|\mathcal{T}|\}$ such that $T = T_k$, then for all $\ell\in\{1,\dots,|\mathcal{T}|\}$,

$$\int_{T_{\ell-1}}^{T_\ell} \mathbf{1}_{T\le t}dt = \int_{T_{\ell-1}}^{T_\ell} \mathbf{1}_{T_k\le t}dt = (T_\ell - T_{\ell-1})\mathbf{1}_{T_k\le T_{\ell-1}} = (T_\ell - T_{\ell-1})\mathbf{1}_{T\le T_{\ell-1}}.$$

This implies

$$\begin{aligned}\int_{T_{\ell-1}}^{T_\ell} F_{\Pi_i}(t)dt &= \int_{T_{\ell-1}}^{T_\ell} \sum_{T\in N_i} \mathbf{1}_{T\le t}dt = \sum_{T\in N_i}\int_{T_{\ell-1}}^{T_\ell} \mathbf{1}_{T\le t}dt \\ &= \sum_{T\in N_i}(T_\ell - T_{\ell-1})\mathbf{1}_{T\le T_{\ell-1}} = (T_\ell - T_{\ell-1})F_{\Pi_i}(T_{\ell-1}),\end{aligned}$$

and consequently

$$\int_{T_{\ell-1}}^{T_\ell} F_{\widehat{W}}(t)dt = (T_\ell - T_{\ell-1})\frac{1}{n}\sum_{i=1}^n F_{\Pi_i}(T_{\ell-1}) = (T_\ell - T_{\ell-1})F_{\widehat{W}}(T_{\ell-1}).$$

and also

$$\int_{T_{\ell-1}}^{T_\ell} F_{\widehat{\Delta}_i}(t)dt = (T_\ell - T_{\ell-1})F_{\widehat{\Delta}_i}(T_{\ell-1}). \tag{49}$$

We turn now to the proof of the first point. Let $\ell\in\{1,\dots,|\mathcal{T}|\}$,

$$\begin{aligned}\widehat{\Gamma}_{\widehat{\Delta}}e_\ell(s) &= \frac{1}{\sqrt{T_\ell - T_{\ell-1}}}\int_{T_{\ell-1}}^{T_\ell} \widehat{K}_{\widehat{\Delta}}(s,t)dt \\ &= \frac{1}{\sqrt{T_\ell - T_{\ell-1}}}\int_{T_{\ell-1}}^{T_\ell}\left(\frac{1}{n}\sum_{i=1}^n\sum_{T,T'\in N_i}\mathbf{1}_{T\le s}\mathbf{1}_{T'\le t} - F_{\widehat{W}}(s)F_{\widehat{W}}(t)\right)dt\end{aligned}$$

$$= \frac{1}{\sqrt{T_\ell - T_{\ell-1}}} \frac{1}{n} \sum_{i=1}^{n} \left(\sum_{T,T' \in N_i} \mathbf{1}_{T \le s} \int_{T_{\ell-1}}^{T_\ell} \mathbf{1}_{T' \le t} dt - F_{\widehat{W}}(s) \int_{T_{\ell-1}}^{T_\ell} F_{\widehat{W}}(t) dt \right)$$

$$= \frac{1}{\sqrt{T_\ell - T_{\ell-1}}} \frac{1}{n} \sum_{i=1}^{n} \left(F_{\Pi_i}(s) \sum_{T' \in N_i} (T_\ell - T_{\ell-1}) \mathbf{1}_{T' \le T_{\ell-1}} - F_{\widehat{W}}(s) F_{\widehat{W}}(T_{\ell-1})(T_\ell - T_{\ell-1}) \right).$$

$$= \frac{\sqrt{T_\ell - T_{\ell-1}}}{n} \sum_{i=1}^{n} \left(F_{\Pi_i}(s) F_{\Pi_i}(T_{\ell-1}) - F_{\widehat{W}}(s) F_{\widehat{W}}(T_{\ell-1}) \right).$$

Following the same lines, for all $\ell' \in \{1, \ldots, |\mathcal{T}|\}$,

$$\langle \widehat{\Gamma}_{\widehat{\Delta}} e_\ell, e_{\ell'} \rangle = \frac{1}{\sqrt{T_{\ell'} - T_{\ell'-1}}} \int_{T_{\ell'-1}}^{T_{\ell'}} \widehat{\Gamma}_{\widehat{\Delta}} e_\ell(s) ds$$

$$= \frac{\sqrt{T_\ell - T_{\ell-1}}}{n\sqrt{T_{\ell'} - T_{\ell'-1}}} \sum_{i=1}^{n} \int_{T_{\ell'-1}}^{T_{\ell'}} \left(F_{\Pi_i}(s) F_{\Pi_i}(T_{\ell-1}) - F_{\widehat{W}}(s) F_{\widehat{W}}(T_{\ell-1}) \right) ds$$

$$= \frac{\sqrt{(T_\ell - T_{\ell-1})(T_{\ell'} - T_{\ell'-1})}}{n} \sum_{i=1}^{n} \left(F_{\Pi_i}(T_{\ell'-1}) F_{\Pi_i}(T_{\ell-1}) - F_{\widehat{W}}(T_{\ell'-1}) F_{\widehat{W}}(T_{\ell-1}) \right).$$

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

$$\int_{T_{\ell-1}}^{T_\ell} F_{\widehat{W}}(t) dt = \frac{\ell-1}{n}(T_\ell - T_{\ell-1}),$$

which concludes the proof of the first point.

Now, we start the proof of the second point by proving that

$$\text{(50)} \qquad \operatorname{Im}(\widehat{\Gamma}_{\widehat{\Delta}}) \subset \operatorname{span}\{e_1, \ldots, e_{|\mathcal{T}|}\}.$$

Since $\widehat{\Gamma}_{\widehat{\Delta}}$ is a finite rank operator, both spaces $\operatorname{Im}(\widehat{\Gamma}_{\widehat{\Delta}})$ and $\operatorname{span}\{e_1, \ldots, e_{|\mathcal{T}|}\}$ are finite-dimensional vector spaces and it is sufficient to prove

$$\operatorname{span}\{e_1, \ldots, e_{|\mathcal{T}|}\}^\perp \subset \operatorname{Im}(\widehat{\Gamma}_{\widehat{\Delta}})^\perp.$$

Moroever, since $\widehat{\Gamma}_{\widehat{\Delta}}$ is a self-adjoint operator,

$$\operatorname{Im}(\widehat{\Gamma}_{\widehat{\Delta}})^\perp = \operatorname{Ker}(\widehat{\Gamma}_{\widehat{\Delta}}).$$

Then, to prove (50), it is sufficient to prove that

$$\operatorname{span}\{e_1, \ldots, e_{|\mathcal{T}|}\}^\perp \subset \operatorname{Ker}(\widehat{\Gamma}_{\widehat{\Delta}}).$$

Let $f \in \operatorname{span}\{e_1, \ldots, e_{|\mathcal{T}|}\}^\perp$, we have for any $\ell \in \{1, \ldots, |\mathcal{T}|\}$,

$$\langle f, e_\ell \rangle = 0,$$

then, by definition of e_ℓ,

$$\int_{T_{\ell-1}}^{T_\ell} f(t) dt = 0.$$

This implies that, for all $i=1,\dots,n$, for all $T\in N_i$, since, by construction of the grid $\mathcal{T}$, there exists ℓ such that $T=T_{\ell-1}$,

$$\int_T^1 f(t)dt=\sum_{\ell'=\ell}^{|\mathcal{T}|}\int_{T_{\ell'}-1}^{T_{\ell'}} f(t)dt=0.$$

Hence

$$\langle F_{\widehat{W}},f\rangle=\int_0^1\widehat{W}([0,t])f(t)dt=\frac{1}{n}\sum_{i=1}^n\sum_{T\in N_i}\int_0^1\mathbf{1}_{\{T\le t\}}f(t)dt=\frac{1}{n}\sum_{i=1}^n\sum_{T\in N_i}\int_T^1 f(t)dt=0$$

and

$$\begin{aligned}\widehat{\Gamma}_{\widehat{\Delta}}f(t)&=\frac{1}{n}\sum_{i=1}^n\sum_{T,T'\in N_i}\int_0^1\mathbf{1}_{\{T\le s,T'\le t\}}f(s)ds-\widehat{W}([0,t])\langle F_{\widehat{W}},f\rangle,\\&=\frac{1}{n}\sum_{i=1}^n\sum_{T,T'\in N_i}\mathbf{1}_{\{T'\le t\}}\int_T^1 f(s)ds-\widehat{W}([0,t])\langle F_{\widehat{W}},f\rangle=0.\end{aligned}$$

This concludes the proof of (50).

Now we identify the eigenfunctions of $\widehat{\Gamma}_{\widehat{\Delta}}$ with the eigenvectors of the matrix $\widehat{G}_{\widehat{\Delta}}$.

- Let $j\ge 1$. We consider $\widehat{\lambda}_j$ a non-zero eigenvalue of $\widehat{\Gamma}_{\widehat{\Delta}}$. Let $\widehat{\eta}_j$ an associated eigenfunction such that $\|\widehat{\eta}_j\|=1$. We set

$$\widehat{v}_j=(\langle\widehat{\eta}_j,e_\ell\rangle)_{\ell=1,\dots,|\mathcal{T}|}.$$

We have

$$\begin{aligned}\widehat{G}_{\widehat{\Delta}}\widehat{v}_j&=\left(\sum_{\ell'=1}^{|\mathcal{T}|}\langle\widehat{\Gamma}_{\widehat{\Delta}}e_\ell,e_{\ell'}\rangle\langle\widehat{\eta}_j,e_{\ell'}\rangle\right)_{\ell=1,\dots,|\mathcal{T}|}\\&=\left(\left\langle\widehat{\Gamma}_{\widehat{\Delta}}e_\ell,\sum_{\ell'=1}^{|\mathcal{T}|}\langle\widehat{\eta}_j,e_{\ell'}\rangle e_{\ell'}\right\rangle\right)_{\ell=1,\dots,|\mathcal{T}|}.\end{aligned}\tag{51}$$

Now remark that $\sum_{\ell'=1}^{|\mathcal{T}|}\langle\widehat{\eta}_j,e_{\ell'}\rangle e_{\ell'}$ is the orthogonal projection of the function $\widehat{\eta}_j$ on the space $\operatorname{span}\{e_1,\dots,e_{|\mathcal{T}|}\}$. Then, since $\widehat{\Gamma}_{\widehat{\Delta}}\widehat{\eta}_j=\widehat{\lambda}_j\widehat{\eta}_j$ and $\widehat{\lambda}_j>0$, from (50), we deduce

$$\widehat{\eta}_j\in\operatorname{Im}(\widehat{\Gamma}_{\widehat{\Delta}})\subset\operatorname{span}\{e_1,\dots,e_{|\mathcal{T}|}\},$$

meaning that

$$\sum_{\ell'=1}^{|\mathcal{T}|}\langle\widehat{\eta}_j,e_{\ell'}\rangle e_{\ell'}=\widehat{\eta}_j$$

and $\|\widehat{v}_j\|=1$. Then, since $\widehat{\Gamma}_{\widehat{\Delta}}$ is a self-adjoint operator, (51) can be written

$$\widehat{G}_{\widehat{\Delta}}\widehat{v}_j=\left(\left\langle\widehat{\Gamma}_{\widehat{\Delta}}e_\ell,\widehat{\eta}_j\right\rangle\right)_{\ell=1,\dots,|\mathcal{T}|}=\left(\left\langle e_\ell,\widehat{\Gamma}_{\widehat{\Delta}}\widehat{\eta}_j\right\rangle\right)_{\ell=1,\dots,|\mathcal{T}|}=\widehat{\lambda}_j\left(\langle e_\ell,\widehat{\eta}_j\rangle\right)_{\ell=1,\dots,|\mathcal{T}|}=\widehat{\lambda}_j\widehat{v}_j.$$

This means that $\widehat{\lambda}_j\in\operatorname{Sp}(\widehat{G}_{\widehat{\Delta}})$ and, consequently

$$\operatorname{Sp}(\widehat{\Gamma}_{\widehat{\Delta}})\backslash\{0\}\subset\operatorname{Sp}(\widehat{G}_{\widehat{\Delta}}).$$

- Conversely, let $\widehat{v}_j = (\widehat{v}_j^1, \ldots, \widehat{v}_j^{|\mathcal{T}|})^t$ a unit-norm eigenvector of $\widehat{G}_{\widehat{\Delta}}$ associated with the eigenvalue $\widehat{\lambda}_j$ and let

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

We have:

$$\widehat{\Gamma}_{\widehat{\Delta}} \widehat{\eta}_j = \sum_{\ell'=1}^{|\mathcal{T}|} \langle \widehat{\Gamma}_{\widehat{\Delta}} \widehat{\eta}_j, e_{\ell'} \rangle e_{\ell'} = \sum_{\ell,\ell'=1}^{|\mathcal{T}|} \widehat{v}_j^\ell \langle \widehat{\Gamma}_{\widehat{\Delta}} e_\ell, e_{\ell'} \rangle e_{\ell'} = \sum_{\ell'=1}^{|\mathcal{T}|} \left[\widehat{G}_{\widehat{\Delta}} \widehat{v}_j \right]_{\ell'} e_{\ell'}$$

$$= \widehat{\lambda}_j \sum_{\ell'=1}^{|\mathcal{T}|} \widehat{v}_j^{\ell'} e_{\ell'} = \widehat{\lambda}_j \widehat{\eta}_j.$$

Then, $\widehat{\eta}_j$ is a unit-norm eigenfunction of the operator $\widehat{\Gamma}_{\widehat{\Delta}}$ associated with the eigenvalue $\widehat{\lambda}_j$ and

$$\mathrm{Sp}(\widehat{G}_{\widehat{\Delta}}) \subset \mathrm{Sp}(\widehat{\Gamma}_{\widehat{\Delta}}).$$

To end the proof of the third point, it remains to remark that, since the functions e_ℓ's are right continuous functions, then $\widehat{\eta}_j$ also is a right-continuous function. Moreover, following the same lines than the proof of Proposition 3.1, $\widehat{\eta}_j$ is a function with bounded variations and $\widehat{\eta}_j(t) = 0$ for $t < 0$. Then, applying Proposition 4.4.3 of [18], we obtain the expected result: there exists a unique measure $\widehat{\mu}_j$ such that $F_{\widehat{\mu}_j} = \widehat{\eta}_j$.

8.3.2. *Proof of Theorem 5.5.* We first prove the first point. From [7], Lemma 4.2, p. 103, we can deduce that

$$\sup_{j \geq 1} |\widehat{\lambda}_j - \lambda_j| \leq |||\Gamma_\Delta - \widehat{\Gamma}_{\widehat{\Delta}}|||, \tag{52}$$

where, for an operator S on $\mathbb{L}^2([0,1])$,

$$|||S||| = \sup_{x \in \mathbb{L}^2([0,1])} \frac{\|Sx\|}{\|x\|}$$

denotes the usual operator norm. Let $\widehat{\Gamma}_\Delta = \frac{1}{n}\sum_{i=1}^n F_{\Delta_i} \otimes F_{\Delta_i}$ we get

$$|||\Gamma_\Delta - \widehat{\Gamma}_{\widehat{\Delta}}||| \leq |||\widehat{\Gamma}_\Delta - \Gamma_\Delta||| + |||\widehat{\Gamma}_{\widehat{\Delta}} - \widehat{\Gamma}_\Delta|||. \tag{53}$$

We deal with the first term. We get classically,

$$\widehat{\Gamma}_\Delta - \Gamma_\Delta = \frac{1}{n} \sum_{i=1}^n F_{\Delta_i} \otimes F_{\Delta_i} - \mathbb{E}[F_\Delta \otimes F_\Delta].$$

We remark that $F_{\Delta_i} \otimes F_{\Delta_i} - \mathbb{E}[F_\Delta \otimes F_\Delta]$, $i = 1, \ldots, n$ is an i.i.d. sequence of centered random variables taking values in the space of Hilbert-Schmidt operators equipped with its usual norm $\| \cdot \|_{HS}$ defined by $\|S\|_{HS}^2 = \sum_{j \geq 1} \|Se_j\|^2$ for $(e_j)_{j \geq 1}$ an orthonormal (hilbertian) basis of $\mathbb{L}^2([0,1])$ (recall that the definition of the norm does not depend on the basis). Furthermore, we have

$$|||\widehat{\Gamma}_\Delta - \Gamma_\Delta||| \leq \|\widehat{\Gamma}_\Delta - \Gamma_\Delta\|_{HS}$$

and

$$\|\widehat{\Gamma}_\Delta - \Gamma_\Delta\|_{HS}^2 = \frac{1}{n^2}\sum_{i,j=1}^{n} \langle F_{\Delta_i}\otimes F_{\Delta_i} - \mathbb{E}[F_\Delta \otimes F_\Delta], F_{\Delta_j}\otimes F_{\Delta_j} - \mathbb{E}[F_\Delta\otimes F_\Delta]\rangle_{HS},$$

where, for two Hilbert-Schmidt operators S and T, the Hilbert-Schmidt scalar product is defined by $\langle S_1, S_2\rangle_{HS} = \sum_{j\geq 1}\langle S_1 e_j, S_2 e_j\rangle$ and does not depend on the considered basis. Then, taking $(e_j)_j = (\eta_j)_j$, we remark that

$$\begin{aligned}
&\mathbb{E}[\langle F_{\Delta_i}\otimes F_{\Delta_i} - \mathbb{E}[F_\Delta\otimes F_\Delta], F_{\Delta_j}\otimes F_{\Delta_j} - \mathbb{E}[F_\Delta \otimes F_\Delta]\rangle_{HS}]\\
&\quad = \sum_{\ell\geq 1}\mathbb{E}\left[\langle (F_{\Delta_i}\otimes F_{\Delta_i} - \mathbb{E}[F_\Delta\otimes F_\Delta])\eta_\ell, (F_{\Delta_j}\otimes F_{\Delta_j} - \mathbb{E}[F_\Delta\otimes F_\Delta])\eta_\ell\rangle\right]\\
&\quad = \sum_{\ell\geq 1}\int_0^1 \mathrm{Cov}\left(F_{\Delta_i}\otimes F_{\Delta_i}\eta_\ell(t), F_{\Delta_j}\otimes F_{\Delta_j}\eta_\ell(t)\right)dt\\
&\quad = \mathbf{1}_{i=j}\sum_{\ell\geq 1}\int_0^1 \mathrm{Var}\left(F_\Delta\otimes F_\Delta \eta_\ell(t)\right)dt\\
&\quad = \mathbf{1}_{i=j}\sum_{\ell\geq 1}\int_0^1 \mathrm{Var}\left(\langle F_\Delta, \eta_\ell\rangle F_\Delta(t)\right)dt\\
&\quad \leq \mathbf{1}_{i=j}\sum_{\ell\geq 1}\mathbb{E}\left[\langle F_\Delta, \eta_\ell\rangle^2\int_0^1 F_\Delta^2(t)dt\right] = \mathbf{1}_{i=j}\mathbb{E}\left[\|F_\Delta\|^4\right].
\end{aligned}$$

Then, finally, we get

$$\mathbb{E}\left[|||\widehat{\Gamma}_\Delta - \Gamma_\Delta|||^2\right] \leq \frac{\mathbb{E}\left[\|F_\Delta\|^4\right]}{n}. \tag{54}$$

Then for the second term,

$$|||\widehat{\Gamma}_\Delta - \widehat{\Gamma}_{\widehat{\Delta}}||| = |||(F_{\widehat{W}} - F_W)\otimes(F_{\widehat{W}} - F_W)||| = \|F_{\widehat{W}} - F_W\|^2.$$

Then, since $F_\Delta(t) = F_\Pi(t) - \mathbb{E}[F_\Pi(t)]$,

$$\begin{aligned}
\mathbb{E}[|||\widehat{\Gamma}_\Delta - \widehat{\Gamma}_{\widehat{\Delta}}|||] &= \mathbb{E}[\|F_{\widehat{W}} - F_W\|^2]\\
&= \mathbb{E}\left[\int_0^1 \left(F_{\widehat{W}}(t) - F_W(t)\right)^2 dt\right]\\
&= \int_0^1 \mathrm{Var}\left(F_{\widehat{W}}(t)\right)dt\\
&= \int_0^1 \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n F_{\Pi_i}(t)\right)dt\\
&= \frac{1}{n}\int_0^1 \mathrm{Var}\left(F_\Pi(t)\right)dt = \frac{1}{n}\int_0^1 \mathbb{E}[F_\Delta(t)^2]dt = \frac{1}{n}\mathbb{E}[\|F_\Delta\|^2]. \qquad (55)
\end{aligned}$$

Finally, combining eq. (53), (54) and (55), we get

$$\mathbb{E}[|||\Gamma_\Delta - \widehat{\Gamma}_{\widehat{\Delta}}|||^2] \leq 4\frac{\mathbb{E}\left[\|F_\Delta\|^4\right]}{n}, \tag{56}$$

and Inequality (23) follows from Bosq inequality (52).

From [7], Lemma 4.3, p.104, we have the upper-bound

$$\|\widehat{\eta}_j - \widetilde{\eta}_j\| \leq 2\sqrt{2}\delta_j^{-1} \||\Gamma_\Delta - \widehat{\Gamma}_{\widehat{\Delta}}\||,$$

and Inequality (24) follows from (56). Now, we also have, by Cauchy-Schwarz's inequality, for all $\varphi \in \mathcal{H}_0^1$,

$$|\langle \widehat{\mu}_j - \widetilde{\mu}_j, \varphi \rangle|^2 = \left|\langle \widehat{\eta}_j - \widetilde{\eta}_j, \varphi' \rangle\right|^2 \leq \|\widehat{\eta}_j - \widetilde{\eta}_j\|^2 \|\varphi'\|^2.$$

Then

$$\|\widehat{\mu}_j - \widetilde{\mu}_j\|_{\mathcal{H}^{-1}}^2 \leq \|\widehat{\eta}_j - \widetilde{\eta}_j\|^2$$

and Inequality (25) follows.

REFERENCES

[1] Andersen, P., Borgan, O., Gill, R. and Keiding, N. (1993). *Statistical models based on counting processes.* New York: Springer-Verlag.

[2] Ash, R. B. and Gardner, M. F. (1975). *Topics in stochastic processes*. Academic Press [Harcourt Brace Jovanovich Publishers], New York. Probability and Mathematical Statistics, Vol. 27. MR0448463 (56 ##6769)

[3] Bacry, E., Bompaire, M., Gaïffas, S. and Muzy, J.-F. (2020). Sparse and low-rank multivariate Hawkes processes. *J. Mach. Learn. Res.* **21** Paper No. 50, 32. MR4095329

[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. https://doi.org/10.1016/j.jspi.2014.03.009 MR3390131

[5] 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. https://doi.org/10.1007/s11222-023-10264-w MR4606276

[6] Bonnet, A., Dion-Blanc, C., Gindraud, F. and Lemler, S. (2022). Neuronal network inference and membrane potential model using multivariate Hawkes processes. *Journal of Neuroscience Methods* **372** 109550. https://doi.org/10.1016/j.jneumeth.2022.109550

[7] Bosq, D. (2000). *Linear Processes in Function Spaces: Theory and Applications*. Springer New York.

[8] Brémaud, P. and Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *Ann. Probab.* **24** 1563–1588. https://doi.org/10.1214/aop/1065725193 MR1411506

[9] Brémaud, P. and Massoulié, L. (2001). Hawkes branching point processes without ancestors. *J. Appl. Probab.* **38** 122–135. https://doi.org/10.1017/s0021900200018556 MR1816118

[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. 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 bioinformatics* **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. https://doi.org/10.1214/17-EJS1251 MR3634334

[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. *International Journal of Forecasting* **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*. John Wiley & Sons.

[17] Chornoboy, E., Schramm, L. and Karr, A. (1988). Maximum likelihood identification of neural point process systems. *Biological cybernetics* **59** 265–275.

[18] Cohn, D. L. (1993). *Measure theory*. Birkhäuser Boston, Inc., Boston, MA Reprint of the 1980 original. MR1454121

[19] Corlay, S. and Pagès, G. (2015). Functional quantization-based stratified sampling methods. *Monte Carlo Methods Appl.* **21** 1–32. https://doi.org/10.1515/mcma-2014-0010 MR3318550

[20] CRANE, R. and SORNETTE, D. (2008). Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of Sciences* **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*, second ed. *Probability and its Applications (New York)*. Springer-Verlag, New York Elementary theory and methods. MR1950431
[23] DALEY, D. J. and VERE-JONES, D. (2008). *An introduction to the theory of point processes. Vol. II*, second ed. *Probability and its Applications (New York)*. Springer, New York General theory and structure. https://doi.org/10.1007/978-0-387-49835-5 MR2371524
[24] DELAIGLE, A., HALL, P. and BATHIA, N. (2012). Componentwise classification and clustering of functional data. *Biometrika* **99** 299–313.
[25] DONNET, S., RIVOIRARD, V. and ROUSSEAU, J. (2020). Nonparametric Bayesian estimation for multivariate Hawkes processes. *Ann. Statist.* **48** 2698–2727. https://doi.org/10.1214/19-AOS1903 MR4152118
[26] EMBRECHTS, P., LINIGER, T. and LIN, L. (2011). Multivariate Hawkes processes: an application to financial data. *J. Appl. Probab.* **48A** 367–378. https://doi.org/10.1239/jap/1318940477 MR2865638
[27] ESCABIAS, M., AGUILERA, A. M. and VALDERRAMA, M. J. (2004). Principal component estimation of functional logistic regression: discussion of two different approaches. *Journal of Nonparametric Statistics* **16** 365–384. https://doi.org/10.1080/10485250310001624738
[28] FARAJTABAR, M., WANG, Y., RODRIGUEZ, M. G., LI, S., ZHA, H. and SONG, L. (2016). COEVOLVE: A Joint Point Process Model for Information Diffusion and Network Co-evolution.
[29] GAO, X. and ZHU, L. (2018). A functional central limit theorem for stationary Hawkes processes and its application to infinite-server queues. *Queueing Systems* **90**. https://doi.org/10.1007/s11134-018-9570-5
[30] GUSTO, G. and SCHBATH, S. S. (2005). FADO: A statistical method to detect favored or avoided distances between occurrences of motifs using the Hawkes model. *Statistical Applications in Genetics and Molecular Biology* **4** n.p.
[31] HANSEN, N. R., REYNAUD-BOURET, P. and RIVOIRARD, V. (2015). Lasso and probabilistic inequalities for multivariate point processes. *Bernoulli* **21** 83–143. https://doi.org/10.3150/13-BEJ562 MR3322314
[32] HAWKES, A. G. (1971a). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58** 83–90. https://doi.org/10.1093/biomet/58.1.83 MR278410
[33] HAWKES, A. G. (1971b). Point spectra of some mutually exciting point processes. *J. Roy. Statist. Soc. Ser. B* **33** 438–443. MR358976
[34] HILGERT, N., MAS, A. and VERZELEN, N. (2013). Minimax adaptive tests for the functional linear model. *The Annals of Statistics* **41** 838 – 869. https://doi.org/10.1214/13-AOS1093
[35] HSING, T. and EUBANK, R. (2015). *Theoretical foundations of functional data analysis, with an introduction to linear operators* **997**. John Wiley & Sons.
[36] 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*. Springer New York, New York, NY. https://doi.org/10.1007/0-387-31144-0_7
[37] JACQUES, J. and PREDA, C. (2014). Model-based clustering for multivariate functional data. *Computational Statistics & Data Analysis* **71** 92-106.
[38] KARASÖZEN, E., NISSEN, E., BÜYÜKAKPINAR, P., CAMBAZ, M. D., KAHRAMAN, M., KALKAN ERTAN, E., ABGARMI, B., BERGMAN, E., GHODS, A. and ÖZACAR, A. A. (2018). The 2017 July 20 Mw 6.6 Bodrum–Kos earthquake illuminates active faulting in the Gulf of Gökova, SW Turkey. *Geophysical Journal International* **214** 185-199.
[39] KARR, A. F. (1991). *Point processes and their statistical inference*, second ed. *Probability: Pure and Applied* **7**. Marcel Dekker, Inc., New York. MR1113698
[40] KINGMAN, J. F. C. (1993). *Poisson processes*. *Oxford Studies in Probability* **3**. The Clarendon Press, Oxford University Press, New York Oxford Science Publications. MR1207584
[41] KOLACZYK, E. D. (1999). Wavelet shrinkage estimation of certain Poisson intensity signals using corrected thresholds. *Statist. Sinica* **9** 119–135. MR1678884
[42] 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. *Journal of Neuroscience Methods* **297**. https://doi.org/10.1016/j.jneumeth.2017.12.026
[43] LI, Y., WANG, N. and CARROLL, R. J. (2013). Selecting the Number of Principal Components in Functional Data. *Journal of the American Statistical Association* **108** 1284–1294.
[44] MANTÉ, C., YAO, A.-F. and DEGIOVANNI, C. (2007). Principal component analysis of measures, with special emphasis on grain-size curves. *Computational Statistics & Data Analysis* **51** 4969-4983.

[45] MARSOLIER, J., PROMPSY, P. . . . and VALLOT, C. (2022). H3K27me3 conditions chemotolerance in triple-negative breast cancer. *Nature Genetics* **54** 459–468.

[46] MOHLER, G. O., SHORT, M. B., BRANTINGHAM, P. J., SCHOENBERG, F. P. and TITA, G. E. (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association* **106** 100–108.

[47] OAKES, D. (1975). The Markovian self-exciting process. *J. Appl. Probability* **12** 69–77. https://doi.org/10.1017/s0021900200033106 MR362522

[48] OGATA, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical association* **83** 9–27.

[49] PANARETOS, V. M. and ZEMEL, Y. (2016). Amplitude and phase variation of point processes. *The Annals of Statistics* **44**. https://doi.org/10.1214/15-aos1387

[50] RASMUSSEN, J. G. (2013). Bayesian Inference for Hawkes Processes. *Methodology and Computing in Applied Probability* **15** 623–642. https://doi.org/10.1007/s11009-011-9272-5

[51] REISS, P. T. and OGDEN, R. T. (2007). Functional Principal Component Regression and Functional Partial Least Squares. *Journal of the American Statistical Association* **102** 984-996.

[52] REYNAUD-BOURET, P. (2003). Adaptive estimation of the intensity of inhomogeneous Poisson processes via concentration inequalities. *Probab. Theory Related Fields* **126** 103–153. https://doi.org/10.1007/s00440-003-0259-1 MR1981635

[53] REYNAUD-BOURET, P. and SCHBATH, S. (2010). Adaptive estimation for Hawkes processes; application to genome analysis. *The Annals of Statistics* **38** 2781?2822. https://doi.org/10.1214/10-aos806

[54] SILVERMAN, B. and RAMSAY, J. (2002). *Applied functional data analysis: methods and case studies*. Springer, New York, NY.

[55] SULEM, D., RIVOIRARD, V. and ROUSSEAU, J. (2024). Bayesian estimation of nonlinear Hawkes processes. *Bernoulli* **30** 1257–1286. https://doi.org/10.3150/23-bej1631 MR4699552

[56] 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. *Journal of the Royal Statistical Society Series C: Applied Statistics* **71** 1402–1426.

[57] WILLETT, R. M. and NOWAK, R. D. (2007). Multiscale Poisson intensity and density estimation. *IEEE Trans. Inform. Theory* **53** 3171–3187. https://doi.org/10.1109/TIT.2007.903139 MR2417680

[58] WU, S., MÜLLER, H.-G. and ZHANG, Z. (2013). Functional data analysis for point processes with rare events. *Statist. Sinica* **23** 1–23. MR3076156

[59] ZETTL, A. (2005). *Sturm-Liouville theory*. *Mathematical Surveys and Monographs* **121**. American Mathematical Society, Providence, RI. https://doi.org/10.1090/surv/121 MR2170950

[60] ZHENG, G. X. Y., TERRY, J. M. . . . and BIELAS, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications* **8** 14049.