Computational Statistics and Data Analysis 187 (2023) 107813

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Statistical depth for point process via the isometric log-ratio transformation

Xinyu Zhou *, Yijia Ma, Wei Wu

Department of Statistics, Florida State University, Tallahassee, FL 32306-4330, USA

ARTICLE INFO

Article history:
Received 1 March 2022
Received in revised form 8 May 2023
Accepted 23 June 2023
Available online 28 June 2023

Dataset link: https://github.com/zhoujackson94/ILR_depth_all

Dataset link: https://www.kaggle.com/sobhanmoosavi/us-accidents

Keywords:
Statistical depth
Point process
Isometric log-ratio transformation
Poisson process
Time rescaling

ABSTRACT

Statistical depth, a useful tool to measure the center-outward rank of multivariate and functional data, is still under-explored in temporal point processes. Recent studies on point process depth proposed a weighted product of two terms - one indicates the depth of the cardinality of the process, and the other characterizes the conditional depth of the temporal events given the cardinality. The second term is of great challenge because of the apparent nonlinear structure of event times, and so far only parametric representations such as Gaussian and Dirichlet densities have been adopted in the definitions. However, these parametric forms ignore the underlying distribution of the process events and are difficult to apply to complicated patterns. To deal with these problems, a novel distribution-based approach is proposed to the conditional depth via the well-known Isometric Log-Ratio (ILR) transformation on the inter-event times. Motivated by a uniform distribution on simplex, the new method, called the ILR depth, is formally defined for general point process via a Time Rescaling. The mathematical properties of the ILR depth are thoroughly examined and the new method is well illustrated by using Poisson and non-Poisson processes to demonstrate its superiority over previous methods. Finally, the ILR depth is applied in a real dataset and the result clearly shows its effectiveness in ranking.

1. Introduction

In this paper, we study the center-outward rank in a set of temporal point process observations. Temporal point process (point process for short) is an important area in stochastic process, which has been extensively studied in both theory and applications. A point process is basically a list of event times in increasing order. Data of many real phenomena can be represented as a point pattern. For instance, traffic accident times in a specific intersection, scoring times in a soccer match, or earthquake happening times in a geographical region. In this study, we focus on orderly point process, where the process can be fully characterized with a conditional intensity function.

Statistical depth has been a powerful tool to measure the center-outward rank of multivariate data and functional data. It was first studied on multivariate data by Tukey (1975) of a given data group in Euclidean space. From then on, various new depths on multivariate data were developed, which include Oja depth (Oja, 1983), simplicial depth (Liu, 1990), Mahalanobis depth (Liu and Singh, 1993), projection depth (Zuo, 2000), zonoid depth (Dyckerhoff et al., 1996), and likelihood depth (Fraiman et al., 1999). Recent investigations on depth have focused more on functional observations. López-Pintado and Romo (2009) developed the notion of functional depth for the first time. Then, Nieto-Reyes (2011) and

* Corresponding author at: 117 N Woodward Ave, Tallahassee, FL, 32306-4330, USA.
E-mail address: xz19c@fsu.edu (X. Zhou).

https://doi.org/10.1016/j.csda.2023.107813
0167-9473/

Mosler and Polyakova (2012) conducted studies on the general concepts and properties of functional depth. More applications of functional depth have focused on outlier detection (Narisetty and Nair, 2016) and data classification (Makinde, 2019).

The notion of depth has recently been introduced to non-Euclidean metric space to deal with more complicated data structures (Geenens et al., 2021; Dai et al., 2021). The methods were illustrated with examples in non-Euclidean function space, Riemannian manifold, and geodesic space. It is clear that the point process is also a non-Euclidean metric space (Wu and Srivastava, 2011) and the metric-based depth methods can be directly applied. Moreover, the notion of depth for point process data has apparent benefits: 1. The center-outward ranking process can enhance the understanding of the temporal variability in the point process, and provide a new and useful tool in addition to the traditional time-based frameworks. 2. This ranking result has broad applications such as analyzing traffic accident patterns (see Section 6 in this paper) and modeling neuronal spike trains (Kass and Ventura, 2001).

Although important, the notion of depth on point process remains under-explored. The only two previous approaches on point process data are the generalized Mahalanobis depth (Liu and Wu, 2017) and the Dirichlet depth (Qi et al., 2021). However, due to the constrained time domain, these methods have clear limitations such as ignoring ordered property of all events and lacking proper symmetry for the deepest point. In this paper, we aim to remove those constraints and define a new depth in an unconstrained Euclidean space. Our approach is based on the well-known Isometric Log-Ratio (ILR) transformation on the inter-event time (IET), and a new symmetry will be introduced to formally define a center with the largest depth value. The ILR transformation is a Log-Ratio analysis method and commonly used in compositional data analysis (Aitchison et al., 2000). It was formally defined by Egozcue et al. (2003) as an isometric isomorphism between the simplex $\mathcal{S}^D$ and Euclidean space $\mathbb{R}^{D-1}$, where $\mathcal{S}^D = \{(x_1, \cdots, x_D) \in \mathbb{R}^D \mid \sum_{i=1}^D x_i = C, x_i > 0, i = 1, \cdots, D\}$, C is a positive constant, and $D > 1$ is an integer. The ILR transformation provides an isometric bijection between the constrained space $\mathcal{S}^D$ and the unconstrained space $\mathbb{R}^{D-1}$. For any point process with $D-1$ events in a finite time domain, its IETs can be equivalently transformed to a vector in $\mathbb{R}^{D-1}$.

In this paper, we will define a density-based depth in the unconstrained space $\mathbb{R}^{D-1}$ after the IET and ILR transformations. As both transformations are invertible, we can map the depth values to the original point process observations. This procedure can eliminate all limitations in the previous methods. For homogeneous Poisson process, we will show that the density of the ILR transformed data is given in a closed form in the unconstrained Euclidean space. This density function is similar to a Gaussian density that it has one global maximum in the center and the function value consistently decreases from the center to boundary. Inspired by the conventional Mahalanobis depth (Liu and Singh, 1993), a new depth can be defined by this density function.

We will show in Section 3 that our proposed new depth method has the following benefits: 1) It is built under a rigorous mathematical framework and can maximally exploit the distribution pattern in the given data. 2) For homogeneous Poisson process, the density after the ILR transformation is given in a closed form. 3) This density has a center-outward pattern with a clear center under orthogonal transformations. 4) This density naturally leads to a new definition of depth, which satisfies all important properties for depth functions. 5) The new depth definition can be easily extended to any general point process with the well-known *Time Rescaling Method* (Brown et al., 2002).

The rest of this paper is organized as follows. In Section 2, we will provide the background of the depth framework that we will adopt in this study. In Section 3, we will conduct the ILR transformation on the inter-event times of a homogeneous Poisson process and derive the density function of the transformed vector in Euclidean space. The formal depth definition will be given on general point process via a Time Rescaling. The mathematical properties of the new depth will be thoroughly examined. In Section 4, we will discuss the computational algorithms for the ILR depth in Poisson and non-Poisson processes. The simulation results and the comparison with previous methods will be shown in Section 5. In Section 6, we will apply the new depth to a real world dataset to demonstrate its effectiveness in characterizing typical patterns. Finally, we will summarize our study and provide future work in Section 7. All mathematical details are shown in Appendices as supplementary materials. The MATLAB code of the implementation of the ILR depth is available at the link https://github.com/zhoujackson94/ILR_depth_all.

2. Background

In this section, we will provide the preliminary background of the depth methods for point process. Let $\mathbb{S}$ denote the set of all point processes in the time domain $[T_1, T_2]$ and $\mathbb{S}_k$ denote the set of all point processes with cardinality k in the time domain $[T_1, T_2]$, e.g. $\mathbb{S}_k = \{(s_1, s_2, \ldots, s_k)^T \in \mathbb{R}^k | T_1 \le s_1 \le s_2 \le \cdots \le s_k \le T_2\}$ and k is any non-negative integer. Therefore, $\mathbb{S} = \bigcup_{k=0}^{\infty} \mathbb{S}_k$. For any point process $\boldsymbol{s} \in \mathbb{S}$, a depth function maps from $\mathbb{S}$ to $\mathbb{R}^+$. The boundary set for point process with cardinality k is denoted as $\mathbb{B}_k = \{(s_1, s_2, \ldots, s_k)^T \in \mathbb{S}_k | \text{at least one equality holds in: } T_1 \le s_1 \le s_2 \le \cdots \le s_k \le T_2\}$.

In this paper, we adopt the overall depth framework in Qi et al. (2021), where the depth is defined as the product of two terms. For each process, the first term is a normalized one dimensional depth of the number of time events, and the second term is a conditional depth of point process given its cardinality. The formal depth definition is given below:

Definition 1. For point process $S \in \mathbb{S}$ defined on $[T_1, T_2]$ with probability measure P, denote $P_{|S|}$ as a probability measure on the cardinality $|S|$ and $P_{S||S|}$ as the probability measure on the ordered events S given $|S|$. For a realization $\boldsymbol{s} \in \mathbb{S}$, denote $\boldsymbol{u}$ as the inter-event time of $\boldsymbol{s}$, the depth $D(\boldsymbol{s}; P)$ is defined as:

$$D(\boldsymbol{s}; P) = w(|\boldsymbol{s}|; P_{|S|})^r D_c(\boldsymbol{s}; P_{S||S|}), \tag{2.1}$$

where $w(|\boldsymbol{s}|; P_{|S|}) = \frac{D_1(|\boldsymbol{s}|; P_{|S|})}{\max_k D_1(|\boldsymbol{s}|=k; P_{|S|})}$ is the normalized one dimensional depth on the cardinality $|\boldsymbol{s}|$, $D_1(|\boldsymbol{s}|; P_{|S|}) = \min\{P_{|S|}(|S| \le |\boldsymbol{s}|), P_{|S|}(|S| \ge |\boldsymbol{s}|)\}$, $r > 0$ is a hyper-parameter and $D_c(\boldsymbol{s}; P_{S||S|})$ is the depth of $\boldsymbol{s}$ conditioned on $|\boldsymbol{s}|$.

Remark 2.1. There are many methods to estimate the one dimensional depth $D_1(|\boldsymbol{s}|; P_{|S|})$; in this paper we adopt the same approach as in Qi et al. (2021). In practice, $D_1(|\boldsymbol{s}|; P_{|S|})$ and $w(|\boldsymbol{s}|; P_{|S|})$ can be easily estimated by samples if the population result is unknown or difficult to obtain. The hyperparameter r is a weighted coefficient to determine the importance on the number of events in a point process. In general, a larger value will rank the data mainly on the number of events, whereas a smaller value will rank the data mainly on the event distribution/pattern. Its proper value may depend on the goal of the ranking process.

The second term $D_c(\boldsymbol{s}; P_{S||S|})$ is the main focus of this paper. Note the conventional multivariate depth may not be directly used because the time events are in a non-Euclidean space (increasing sequence in the finite domain $[T_1, T_2]$). In this paper, we will propose to transform point process data to a Euclidean domain, and then utilize its density function to define the depth. The detailed procedure is given in the next section.

3. Proposed methods

In this section, we will provide details on the proposed methods. At first, we will equivalently represent point process using a vector in simplex, and then introduce the ILR transformation on it.

3.1. ILR transformation on simplex

Given a point process $\boldsymbol{s} = (s_1, s_2, \ldots, s_k)$ with conditioned cardinality k in the time domain $[T_1, T_2]$. Denote $s_0 = T_1$ and $s_{k+1} = T_2$. Then the process can be equivalently represented using a vector of the inter-event times (IET), obtained as $\boldsymbol{u} = (u_1, u_2, \ldots, u_{k+1})^T = (s_1 - s_0, s_2 - s_1, \ldots, s_{k+1} - s_k)^T$. It is easy to see that $\sum_{i=1}^{k+1} u_i = T_2 - T_1$ for $u_i \ge 0$, $i = 1, \ldots, k+1$. That is, all IET vectors form a simplex $\mathcal{S}^{k+1}$ in $\mathbb{R}^{k+1}$, where

$$\mathcal{S}^{k+1} = \{\boldsymbol{u} = (u_1, u_2, \ldots, u_{k+1})^T \mid \sum_{i=1}^{k+1} u_i = T_2 - T_1, u_i \ge 0, i = 1, \cdots, k+1\}.$$

One previous approach assumes the IET observations follow a Dirichlet distribution, which motivates the definition of Dirichlet depth (Qi et al., 2021). However, we point out that a Dirichlet model is only a simplified assumption and in general it is very challenging to characterize data in simplex. In this paper, we propose to adopt the well-known Isometric Log-Ratio (ILR) transformation to map the IET vectors to a conventional vector space. In this manner, we can examine the distribution of the IET vectors in the equivalent, unconstrained Euclidean space.

The ILR transformation is an isometric isomorphism mapping from simplex space $\mathcal{S}^{k+1}$ to Euclidean space $\mathbb{R}^k$. Specifically, the transformation of any $\boldsymbol{u} = (u_1, u_2, \ldots, u_{k+1})^T \in \mathcal{S}^{k+1}$ is given in the following form (Pawlowsky-Glahn et al., 2007):

$$\boldsymbol{u}^* = ilr(\boldsymbol{u}) = \Psi \cdot \Big[\log \frac{u_1}{g(\boldsymbol{u})}, \log \frac{u_2}{g(\boldsymbol{u})}, \ldots, \log \frac{u_{k+1}}{g(\boldsymbol{u})}\Big]^T, \tag{3.1}$$

where $g(\boldsymbol{u})$ is the geometric mean of $\boldsymbol{u}$. Ψ is a matrix in $\mathbb{R}^{k\times(k+1)}$ which satisfies $\Psi\Psi^T = I_k$ and $\Psi^T\Psi = I_{k+1} - \frac{1}{k+1}\mathbf{1}_{k+1}\mathbf{1}_{k+1}^T$, where I_k is the identity matrix in $\mathbb{R}^{k\times k}$, I_{k+1} is the identity matrix in $\mathbb{R}^{(k+1)\times(k+1)}$, and $\mathbf{1}_{k+1}$ is a column vector of ones in $\mathbb{R}^{k+1}$. Based on Eqn. (3.1), the inverse of ILR, i.e. recovering $\boldsymbol{u}$ from $\boldsymbol{u}^*$, takes the following form (Pawlowsky-Glahn et al., 2007):

$$\boldsymbol{u} = ilr^{-1}(\boldsymbol{u}^*) = (T_2 - T_1) \cdot \frac{\exp(\boldsymbol{u}^{*T}\Psi)^T}{\exp(\boldsymbol{u}^{*T}\Psi) \cdot \mathbf{1}_{k+1}}. \tag{3.2}$$

Based on the ILR transformation, we obtain an important property of Ψ, shown in Proposition 3.1, where the proof is given in Appendix A. This property will be used in our newly defined depth.

Proposition 3.1. *All $k+1$ columns of the matrix $\Psi \in \mathbb{R}^{k\times(k+1)}$ form a regular simplex in $\mathbb{R}^k$ centered at origin with edge length $\sqrt{2}$.*

Using the ILR transformation and its inverse, bijective mappings for the three spaces $\mathbb{S}_k$, $\mathcal{S}^{k+1}$ and $\mathbb{R}^k$ can be established. Whenever a point process $\boldsymbol{s} \in \mathbb{S}_k$ within a fixed time domain is given, its IET $\boldsymbol{u} \in \mathcal{S}^{k+1}$ as well as the ILR transformation $\boldsymbol{u}^* \in \mathbb{R}^k$ can be easily obtained. On the other hand, for any $\boldsymbol{u}^* \in \mathbb{R}^k$, the corresponding point process in $\mathbb{S}_k$ can be recovered

by the ILR inverse. Therefore, the distribution of $\boldsymbol{u}^*$ can be derived from the distribution of $\boldsymbol{u}$, which will be illustrated in detail in the next subsection.

We often use various mathematical properties to evaluate the performance of a depth method. A list of four properties for the conditional depth in point process is given below and we will evaluate them in the proposed depth. This list corresponds to the similar four properties in the multivariate case (Zuo and Serfling, 2000), and a qualitative explanation of these properties is given in Appendix B.

1. $D_c(\boldsymbol{s}; P_{S||S|=k})$ is a continuous mapping from $\mathbb{S}_k$ to $\mathbb{R}^+$ and $D_c(\boldsymbol{s}; P_{S||S|=k}) = 0$ if $\boldsymbol{s} \in \mathbb{B}_k$.
2. There exists unique $\boldsymbol{s}_c$ such that $D_c(\boldsymbol{s}_c; P_{S||S|=k}) = \sup_{\boldsymbol{s}\in\mathbb{S}_k} D_c(\boldsymbol{s}; P_{S||S|=k})$, where $\boldsymbol{s}_c$ is the center point given a specific symmetry.
3. For any point process $\boldsymbol{s}$ with IET $\boldsymbol{u}$, denote $d_c(\boldsymbol{u}) = D_c(\boldsymbol{s}; P_{S||S|=k})$. If $\boldsymbol{s}_c$ is the center point and $\boldsymbol{u}_c$ is the corresponding IET, then $d_c(\boldsymbol{u}) \le d_c(\boldsymbol{u}_c \oplus \alpha \odot (\boldsymbol{u} \ominus \boldsymbol{u}_c))$ holds for any $\alpha \in [0, 1]$, where $\oplus, \odot$ and $\ominus$ correspond to specific linear operations (i.e., addition, scaling, and subtraction) in the simplex space (Pawlowsky-Glahn et al., 2007).
4. For any scaling coefficient $a \in \mathbb{R}^+$ and translation coefficient $b \in \mathbb{R}$, $D_c(\boldsymbol{s}; P_{S||S|=k}) = D_c(a\boldsymbol{s} + b; P_{aS+b||S|=k})$.

3.2. Density of the ILR-transformed uniform distribution

For simplicity, we will at first study the homogeneous Poisson process (HPP) (Stoyan et al., 2013). As pointed out in Qi et al. (2021), given the cardinality k of point process, the IET of an HPP is uniformly distributed on simplex space $\mathcal{S}^{k+1}$. In this way, the joint density function of the IET vector $(u_1, u_2, ..., u_{k+1})^T$ is:

$$f_{\boldsymbol{u}}(u_1, u_2, ..., u_{k+1}) = k! \cdot \frac{1}{(T_2 - T_1)^k}, \tag{3.3}$$

Thus, in order to obtain the density function of $\boldsymbol{u}^*$, the Jacobian matrix J of $\boldsymbol{u}$ corresponding to $\boldsymbol{u}^*$ should be derived in closed form by taking derivative of u_j to u_i^* for $i, j = 1, 2, \ldots, k$. Since u_{k+1} is not random given $u_1, \ldots, u_k$, it can be omitted when deriving the density. Then, the density function of $\boldsymbol{u}^*$ is:

$$f_{\boldsymbol{u}^*}(u_1^*, ..., u_k^*) = \frac{k! \cdot |\det(J)|}{(T_2 - T_1)^k}. \tag{3.4}$$

The closed form of Eqn. (3.4) can be summarized in the next theorem (see proof in Appendix C).

Theorem 3.1. *Let $\boldsymbol{s} = (s_1, s_2, \ldots, s_k)$ be a realization of a homogeneous Poisson point process with k time events in $[T_1, T_2]$. Denote $s_0 = T_1$, $s_{k+1} = T_2$ and $\boldsymbol{u}^* = (u_1^*, u_2^*, ..., u_k^*)^T$ as the ILR transformation of the IET $\boldsymbol{u} = (u_1, u_2, \ldots, u_{k+1})^T = (s_1 - s_0, s_2 - s_1, \ldots, s_{k+1} - s_k)^T$. Then the probability density function of $\boldsymbol{u}^*$ conditioned on its cardinality is:*

$$f_{\boldsymbol{u}^*}(\boldsymbol{u}^*) = f_{\boldsymbol{u}^*}(u_1^*, \ldots u_k^*) = \frac{c}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^{k} u_i^* \Psi_{i,p}})^{k+1}}, \tag{3.5}$$

where c is the normalizing constant (to make the integral of the density be 1).

To understand the density in Eqn. (3.5), we show two examples when $k = 1$ and 2: When $k = 1$, the density function can be simplified as $f_{\boldsymbol{u}^*}(u^*) = \frac{\sqrt{2}}{(e^{u^*\Psi_{1,1}} + e^{-u^*\Psi_{1,1}})^2}$. This function is shown in Fig. 1(a). We can see that it has a clear bell-shape, close to a standard normal distribution density. When $k = 2$, the density function is $f_{\boldsymbol{u}^*}(u_1^*, u_2^*) = \frac{6|\Psi_{1,1}\Psi_{2,2} - \Psi_{1,2}\Psi_{2,1}|}{(e^{u_1^*\Psi_{1,1} + u_2^*\Psi_{2,1}} + e^{u_1^*\Psi_{1,2} + u_2^*\Psi_{2,2}} + e^{u_1^*\Psi_{1,3} + u_2^*\Psi_{2,3}})^3}$, and the corresponding graphs of density as well as contours are shown in Fig. 1(b). We also see a 2-dimensional bell shape in density. However, we notice that the contours are not elliptical, but smoothed triangular.

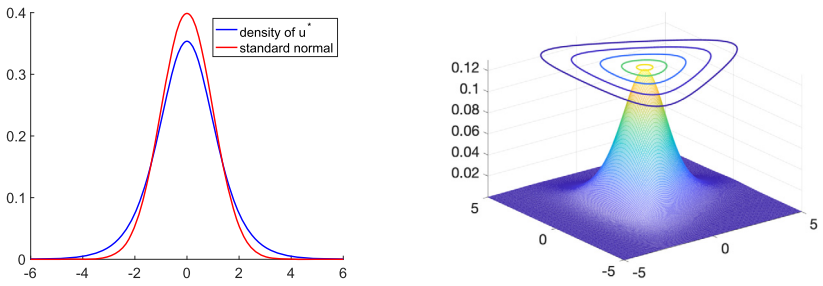
Based on the observations in these special cases, the density in Eqn. (3.5) appears to satisfy two important properties:

1. The density is log-concave and uni-modal.
2. The density has a symmetry with respect to the origin in a simplex. We will use this symmetry in the proposed depth function.

We point out that these properties are indeed true and formally given in Propositions 3.2 and 3.3 as follows.

Proposition 3.2. *The density function given in Eqn. (3.5) is log-concave and uni-modal, and the global maximum point is the origin in $\mathbb{R}^k$.*

The proof of Proposition 3.2 is given in Appendix D.



(a) $k = 1$ (b) $k = 2$

Fig. 1. Two special cases of the density in Eqn. (3.5). (a) The blue line is the density function when $k = 1$, and the red line is the standard normal density function (b) Mesh plot of density when $k = 2$, where the curves at the top are contours of the density function. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Remark 3.1. Based on this result, the density function in Eqn. (3.5) is centered at origin and the value decreases from it in $\mathbb{R}^k$. Thus, if $\boldsymbol{u}^*$ is mapped back to $\mathbb{S}_k$, the density function in Eqn. (3.5) provides an ideal candidate for point process depth conditioned on its cardinality, which will be formally defined in Section 3.3.1.

Next, according to Fig. 1(b), the density contours appear symmetric about the origin for orthogonal transformations on three corners of the smoothed triangular shape. In general, we will show this symmetric property with respect to $(k+1)!$ orthogonal transformations in k-dimensional space. We will at first define those transformations: Assume $r(\cdot)$ is any permutation operation on $1, 2, \ldots, k+1$, and denote Ψ_r as the matrix after permuting the column of Ψ in the order $r(1), r(2), \ldots, r(k+1)$. That is, using column-wise representation, if $\Psi = (\Psi_{:,1}, \cdots, \Psi_{:,k+1})$, then $\Psi_r = (\Psi_{:,r(1)}, \cdots, \Psi_{:,r(k+1)})$. Let

$$A_r = \Psi\Psi_r^T \in \mathbb{R}^{k\times k}. \tag{3.6}$$

It is easy to verify that A_r is orthogonal and $A_r^T\Psi_{:,i} = \Psi_{:,r(i)}$, $i = 1, 2, \ldots, k+1$.

Using Eqn. (3.5), we have

$$f_{\boldsymbol{u}^*}(\boldsymbol{u}^*) = \frac{c}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^{k} u_i^*\Psi_{i,p}})^{k+1}} = \frac{c}{(\sum_{p=1}^{k+1} e^{(\boldsymbol{u}^*)^T\Psi_{:,p}})^{k+1}}.$$

Thus,

$$f_{\boldsymbol{u}^*}(A_r\boldsymbol{u}^*) = \frac{c}{(\sum_{p=1}^{k+1} e^{(A_r\boldsymbol{u}^*)^T\Psi_{:,p}})^{k+1}} = \frac{c}{(\sum_{p=1}^{k+1} e^{(\boldsymbol{u}^*)^T\Psi_{:,r(p)}})^{k+1}} = f_{\boldsymbol{u}^*}(\boldsymbol{u}^*).$$

Therefore, the above analysis has shown the following proposition on permutation symmetry:

Proposition 3.3. *For any orthogonal transformation A_r defined via a permutation $r(\cdot)$ in Eqn. (3.6), $f_{\boldsymbol{u}^*}(A_r\boldsymbol{u}^*) = f_{\boldsymbol{u}^*}(\boldsymbol{u}^*)$ for any column vector $\boldsymbol{u}^* \in \mathbb{R}^k$.*

Remark 3.2. In Proposition 3.3, if $k = 2$, then there are $(2+1)! = 6$ different orthogonal matrices in total. Three of them are rotation matrices and the corresponding rotation angles are 0, $\frac{2\pi}{3}$, and $\frac{4\pi}{3}$, respectively. The other three are reflection matrices.

Combining Proposition 3.1 and Proposition 3.3, the key part of density function of $\boldsymbol{u}^*$ can be considered as summing up the exponential of the inner product of $\boldsymbol{u}^*$ with the vertices of a $k+1$ regular simplex centered at origin in $\mathbb{R}^k$. Therefore, the orthogonal transformation of $\boldsymbol{u}^*$ can be viewed as orthogonal transformation of a regular simplex corresponding to the origin. In this way, the permutation symmetry of the density of $\boldsymbol{u}^*$ can be easily interpreted and the center of this density is the same as the center of the regular simplex.

3.3. ILR depth for point process

In this subsection, we will formally define the depth for point process, conditioned on its cardinality. The definition is based on the density of the ILR transformed inter-event times in Eqn. (3.5) and we call the new method the ILR depth.

3.3.1. Definition

The density function of the ILR transformed IET in an HPP is given in Eqn. (3.5). Based on the fact that the ILR transformation is an isometric isomorphism between the simplex $\mathcal{S}^{k+1}$ and Euclidean space $\mathbb{R}^k$ (Pawlowsky-Glahn et al., 2007), this density can capture temporal pattern in the original point process. It is shown in Proposition 3.2 that this density is log-concave and uni-modal, and therefore it provides an ideal form to define the depth.

As the ILR transformation has a closed-form inverse, we can express the depth in terms of the time events. Based on Eqn. (3.5), we have

$$\Big(\sum_{p=1}^{k+1} e^{\sum_{i=1}^{k} u_i^* \Psi_{i,p}}\Big)^{k+1} = \Big(\sum_{p=1}^{k+1} e^{\log \frac{\boldsymbol{u}^T}{g(\boldsymbol{u})} \Psi^T \Psi_{:,p}}\Big)^{k+1}$$
$$= \Big(\sum_{p=1}^{k+1} e^{\log \frac{u_p}{g(\boldsymbol{u})}}\Big)^{k+1} = \Big(\frac{T_2 - T_1}{g(\boldsymbol{u})}\Big)^{k+1} = \frac{(T_2 - T_1)^{k+1}}{\prod_{i=1}^{k+1}(s_i - s_{i-1})}.$$

In this way, the formal depth definition conditioned on cardinality can be derived in terms of a point process $\boldsymbol{s}$, its IET $\boldsymbol{u}$, or the ILR transformation $\boldsymbol{u}^*$ of IET. In other words, the depth can be defined on the three equivalent spaces, $\mathbb{S}_k$, $\mathcal{S}^{k+1}$ and $\mathbb{R}^k$. Similar to the commonly used Mahalanobis depth in Liu and Singh (1993), the density in Eqn. (3.5) will not be directly used to define depth because the value of density will decrease sharply when the data point deviates from the global maximum point. This phenomenon will become more evident when the dimension k is large. Instead, similar to the definition of the Mahalanobis depth, a logarithm-based increasing function $f(\cdot) = \frac{1}{1-\log(\cdot)}$ can be used to the kernel part of density to alleviate the decreasing rate.

For general point process, the well-known *Time Rescaling (TR) method* (Brown et al., 2002) will be adopted to convert the process to an HPP: Denote $T_1 < s_1 < s_2 < \cdots < s_k < T_2$ as a realization from a point process with a conditional intensity function $\lambda(t|H_t) > 0$ for all $t \in [T_1, T_2]$. Then, the sequence $\Lambda_S(s_i) = \int_{T_1}^{s_i} \lambda(t|H_t)dt$, $i = 1, \ldots, k$ is a Poisson process with the unit rate in $(0, \Lambda_S(T_2)]$. In this way, any point process can be transformed to an HPP as long as the conditional intensity function is known (or can be estimated). Therefore, the ILR depth can be formally defined in the following form:

Definition 2. For a point process $\boldsymbol{s} \in \mathbb{S}_k$ in time domain $[T_1, T_2]$ with cardinality k satisfying $T_1 < s_1 < s_2 < \cdots < s_k < T_2$, assume the conditional intensity function $\lambda(t|H_t) > 0$ and denote $\Lambda_S(t) = \int_{T_1}^{t} \lambda(u|H_u)du$. Let $s_0 = T_1$ and $s_{k+1} = T_2$. The ILR depth of $\boldsymbol{s}$ is defined as:

$$D_c(\boldsymbol{s}; P_{S||S|=k}) = \frac{1}{1 - \log\Big(\frac{c}{\Lambda_S(T_2)^{k+1}} \prod_{i=1}^{k+1}(\Lambda_S(s_i) - \Lambda_S(s_{i-1}))\Big)},$$

where c is a positive constant in $(0, e(k+1)^{k+1})$ to make the depth value positive. We can let $c = (k+1)^{k+1}$ to have the maximum depth value being 1. If $\boldsymbol{s} \in \mathbb{B}_k$, we define $D_c(\boldsymbol{s}; P_{S||S|=k}) = 0$.

Remark 3.3. In Theorem 3.1, there is an important assumption that the point process $\boldsymbol{s}$ belongs to the interior of $\mathbb{S}_k$, i.e., $\boldsymbol{s} \notin \mathbb{B}_k$. Otherwise, the ILR transformation cannot be properly conducted. However, in Definition 2, if $\boldsymbol{s} \in \mathbb{B}_k$, the depth value is defined to be 0 since in this case the ILR depth is continuous at boundary set $\mathbb{B}_k$. In addition, if $\boldsymbol{s}$ is an HPP, then the conditional intensity function is a positive constant number λ. In this case, $\Lambda_S(s_i) - \Lambda_S(s_{i-1}) = \lambda(s_i - s_{i-1})$ for $i = 1, 2, \ldots, k+1$ and $\Lambda_S(T_2) = \lambda(T_2 - T_1)$. With some simple algebra, the depth formula can be simplified as $D_c(\boldsymbol{s}; P_{S||S|=k}) = \frac{1}{1-\log\big(\frac{c}{(T_2-T_1)^{k+1}} \prod_{i=1}^{k+1}(s_i - s_{i-1})\big)} = \frac{1}{1-\log\big(\frac{c}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^{k} u_i^* \Psi_{i,p}})^{k+1}}\big)}$.

In the remaining part of this paper, the constant c in Definition 2 is fixed as $(k+1)^{k+1}$ to normalize the maximum value of the ILR depth being 1. Moreover, based on Proposition 3.2, the density function of the ILR transformed data in Eqn. (3.5) can be well approximated by a Gaussian density (by using a Laplacian approximation), which can lead to another type of depth definition. More detail on this simplified version is given in Appendix E.

3.3.2. Mathematical properties

In Section 3.1, we have listed four basic mathematical properties that a conditional depth is expected to satisfy. In order to make the ILR depth a proper conditional depth, a center needs to be defined corresponding to an appropriate symmetry based on Property 2. Previous studies (Zuo and Serfling, 2000) have summarized different types of depth symmetries for multivariate data, e.g. A-symmetry, C-symmetry and H-symmetry. However, none of them can be directly applied to the ILR depth. In the definition of the Dirichlet depth, the center is taken as the conditional mean of IET $\boldsymbol{u} = (\frac{T_2-T_1}{k+1}, \frac{T_2-T_1}{k+1}, \ldots, \frac{T_2-T_1}{k+1})^T$ for HPP case since the depth value reaches the maximum value at this point (Qi et al., 2021). This "center", however, lacks geometric interpretation with respect to a symmetry in the simplex space $\mathcal{S}^{k+1}$.

By Remark 3.3, for HPP case, the ILR depth can be defined in the unconstrained Euclidean space $\mathbb{R}^k$ transformed from the simplicial domain $\mathcal{S}^{k+1}$. Thus, we can examine possible symmetries with respect to probability distribution of the transformed IET. According to Proposition 3.3, the origin is the center of the ILR depth contours in $\mathbb{R}^k$ with respect to $(k+1)!$ orthogonal transformations defined using all vertices in a regular simplex in Eqn. (3.6). Given the bijective mappings between $\mathbb{S}_k$, $\mathcal{S}^{k+1}$, and $\mathbb{R}^k$, we can map the origin back to $\mathcal{S}^{k+1}$ and $\mathbb{S}_k$, and call the two corresponding points as "simplicial centers" with respect to the "permutation symmetry". This result is formally given in the following proposition.

Proposition 3.4. *For any HPP with k events in a time domain $[T_1, T_2]$, the simplicial center with respect to the permutation symmetry in $\mathbb{S}_k$ is*

$$\boldsymbol{s} = (T_1 + \frac{T_2 - T_1}{k+1}, T_1 + \frac{2(T_2 - T_1)}{k+1}, \ldots, T_1 + \frac{k(T_2 - T_1)}{k+1}),$$

and in $\mathcal{S}^{k+1}$ is

$$\boldsymbol{u} = (\frac{T_2 - T_1}{k+1}, \frac{T_2 - T_1}{k+1}, \ldots, \frac{T_2 - T_1}{k+1})^T.$$

Once the notion of simplicial center is given, the mathematical properties of the ILR depth can be easily verified subsequently. This is summarized in the following proposition, where the detailed proof and explanations are given in Appendix F.

Proposition 3.5. *For the ILR conditional depth, given the cardinality, Property 1 and Property 4 hold for any point process; Property 2 holds for any Poisson process; Property 3 holds for any homogeneous Poisson process.*

Based on Proposition 3.5, the ILR depth is an appropriate conditional depth to measure point process. In particular, the method proforms best for homogeneous Poisson process where all properties hold. Compared with the Dirichlet depth introduced in Qi et al. (2021), the ILR depth has a clear theoretical interpretation for the center and the corresponding linear structure; while the concept of center in the Dirichlet depth is vague and the description on its linear structure is inappropriate. In the following sections, we will discuss the computational issues and illustrate the simulation studies in detail.

4. Computational aspects

According to Definition 2, the conditional intensity function will be used when applying the Time Rescaling method to compute depth value of general point processes. However, the true conditional intensity function is often unknown and needs an estimation. In the following subsections, we will examine the estimation procedures according to two aspects. One is the inhomogeneous Poisson process (IPP), and the other is a non-Poisson process with Markovian property on inter-event times.

4.1. ILR depth computation for inhomogeneous Poisson process

For Poisson process, the process events are independent of each other and the intensity function is deterministic. In this subsection, we propose to use a conventional histogram or binning method to estimate the intensity. Such method has been commonly used in density estimation (Silverman, 2018). Based on the definition, the intensity function in an IPP is given as

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\mathbb{E}[N(t + \Delta t) - N(t)]}{\Delta t},$$

where $N(t)$ is the counting measure on $[0, t)$. The intensity can be approximated when the bin size Δt is small. Therefore, the intensity as well as the depth value of an IPP can be estimated by binning method. We have shown the estimation in Algorithm 1. The large sample theory on this estimation is discussed in detail in Appendix I.

4.2. ILR depth computation for non-Poisson process

If the point process is a non-Poisson process in the time domain $[T_1, T_2]$, then there exists history dependence in conditional intensity function and the Histogram method is not applicable. The estimation of the conditional intensity $\lambda(t|H_t)$ is, in general, highly challenging. A tractable simplification assumes the Markovian property in the following form (Kass and Ventura, 2001):

$$\lambda(t|H_t) = \lambda(t, t - s_*(t)) = \lambda_1(t)\lambda_2(t - s_*(t)), \tag{4.1}$$

where $\lambda_1(\cdot), \lambda_2(\cdot)$ are two deterministic intensity functions and $s_*(t)$ is the last time event preceding to t. If there is no time event before time t, then denote $s_*(t) = T_1$. Point process with this simplified conditional intensity function is called an *inhomogeneous Markov interval (IMI) process*.

Algorithm 1 ILR depth estimation for inhomogeneous Poisson process.

Input: n independent realizations of IPP; The time domain $[T_1, T_2]$; Number of bins M.
- Evenly divide $[T_1, T_2]$ into M bins with equal width;
for each $i = 1, 2, \ldots, n$ **do**
- Denote n_i as the number of events of the i-th realization;
- Denote $\boldsymbol{s_i} = (s_{i1}, s_{i2}, \ldots, s_{in_i})$ as the i-th realization;
- Denote $s_{i0} = T_1$, $s_{i(n_i+1)} = T_2$;
end for
for each $j = 1, 2, \ldots, M$ **do**
- Denote the j-th bin as B_j;
- The intensity estimator is: $\hat{\lambda}(t) = \frac{M}{n(T_2-T_1)} \sum_{i=1}^{n} \sum_{r=1}^{n_i} I(s_{ir} \in B_j)$ if $t \in B_j$;
end for
for each $i = 1, 2, \ldots, n$ **do**
for each $j = 0, 1, \ldots, n_i + 1$ **do**
- Compute $\hat{\Lambda}_S^{(n)}(s_{ij}) = \int_{T_1}^{s_{ij}} \hat{\lambda}(t)dt$;
end for
- Depth of the i-th realization is:

$$\hat{D}_i = \frac{1}{1 - \log\left(\frac{(n_i+1)^{n_i+1}}{\hat{\Lambda}_S^{(n)}(T_2)^{n_i+1}} \prod_{j=1}^{n_i+1} (\hat{\Lambda}_S^{(n)}(s_{ij}) - \hat{\Lambda}_S^{(n)}(s_{i(j-1)}))\right)}.$$

end for
Output: Depth values $\hat{D}_1, \ldots, \hat{D}_n$.

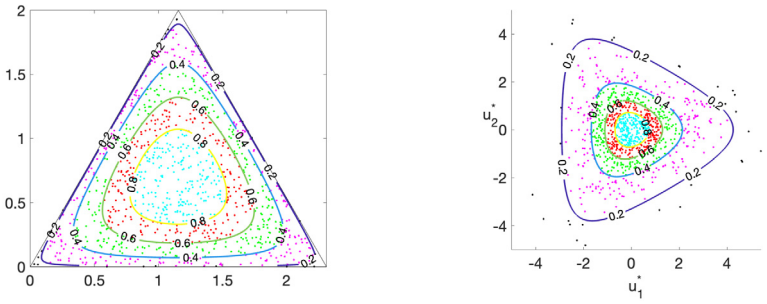
Algorithm 2 Depth estimation for inhomogeneous Markov interval process.

Input: n independent realizations of an IMI process on $[T_1, T_2]$; Number of bins M_1 to estimate λ_1; Number of bins M_2 to estimate λ_2.
- Calculate all inter-event times in the data. Denote the largest one as L;
- Uniformly divide L into M_2 bins and denote the bin width as dt;
for each $i = 1, 2, \ldots, n$ **do**
- Denote n_i as the number of events of the i-th realization; Denote $\boldsymbol{s_i} = (s_{i1}, s_{i2}, \ldots, s_{in_i})$ as the i-th realization; Denote $s_{i0} = T_1$, $s_{i(n_i+1)} = T_2$; Denote $\boldsymbol{u_i} = (u_{i1}, \ldots, u_{in_i}) = (s_{i1} - s_{i0}, \ldots, s_{in_i} - s_{i(n_i-1)})$;
end for
for each $i = 1, 2, \ldots, M_2$ **do**
- Denote the i-th bin as B_i;
- For any $\tau \in B_i$, compute the density of IET $p(\tau) = \frac{\sum_{j=1}^{n} \sum_{k=1}^{n_j} I(u_{jk} \in B_i)}{dt \cdot \sum_{j=1}^{n} n_j}$;
- Compute the conditional intensity function $\lambda_2(\tau) = \frac{p(\tau)}{1 - \int_0^{\tau} p(\tau')d\tau'}$;
end for
- Uniformly divide $[T_1, T_2]$ into M_1 bins and denote the bin width as Δt;
for each $k = 1, 2, \ldots, M_1$ **do**
- Denote the k-th bin as A_k and $t_k = (k - \frac{1}{2})\Delta t$
- Denote $\tau_k^j = t_k - s_*^j(t_k)$, where $s_*^j(t_k)$ is the nearest events before t_k in the jth realization, $j = 1, 2, \ldots, n$
- Probability of an event in the k-th bin $p_k = \frac{\sum_{i=1}^{n} \sum_{r=1}^{n_i} I(s_{ir} \in A_k)}{\sum_{i=1}^{n} n_i}$;
- Compute $\lambda_1(t) = \frac{p_k \cdot n}{\Delta t \cdot \sum_{j=1}^{n} \lambda_2(\tau_k^j)}$ if $t \in A_k$;
end for
for each $i = 1, 2, \ldots, n$ **do**
- Compute $\hat{\Lambda}_S^{(n)}(s_{ij}) = \int_{T_1}^{s_{ij}} \lambda_1(t)\lambda_2(t - s_*(t))dt$, $j = 0, 1, \ldots, n_i + 1$
- $\hat{D}_i = \frac{1}{1 - \log\left(\frac{(n_i+1)^{n_i+1}}{\hat{\Lambda}_S^{(n)}(T_2)^{n_i+1}} \prod_{j=1}^{n_i+1} (\hat{\Lambda}_S^{(n)}(s_{ij}) - \hat{\Lambda}_S^{(n)}(s_{i(j-1)}))\right)}$.
end for
Output: $\hat{D}_1, \ldots, \hat{D}_n$ are the depth values.

To obtain an estimation of Eqn. (4.1) from observed data, we adopt a non-parametric approach proposed by Wójcik et al. (2009). To simplify the notation, denote $\tau = t - s_*(t)$. Thus, Eqn. (4.1) can be rewritten as $\lambda(t, \tau) = \lambda_1(t)\lambda_2(\tau)$. In this case, $\lambda(t, \tau)$ is the product of two intensity functions; one depends on the current time t and the other one only corresponds to the time to the last event τ. The computation of the ILR depth with the IMI model to estimate $\lambda_1(t)$ and $\lambda_2(\tau)$ is shown in Algorithm 2.

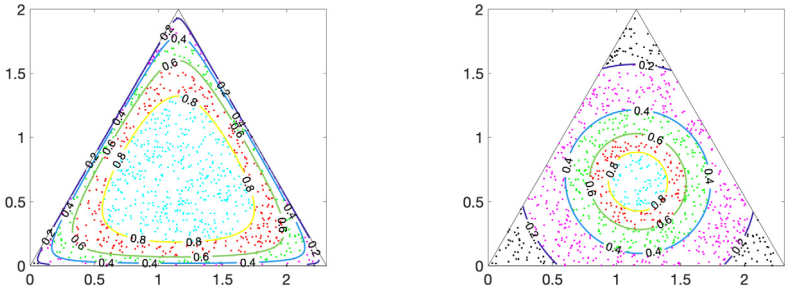
5. Simulation studies

In this section, we will conduct two simulation studies of the ILR depth for the HPP and non-Poisson cases, respectively. The simulation of the IPP case is similar to the one for HPP and therefore given in Appendix K.



(a) data points in Simplex with the ILR depth contour

(b) data points after ILR with the ILR depth contour



(c) data points with the Dirichlet depth contour

(d) data points with the generalized Mahalanobis depth contour

Fig. 2. ILR depth result in HPP. (a) Data points in simplex shown by ternary plot, where the points are colored with respect to ranges of the depth values. The solid lines indicate the depth contours with specific values. (b) Same as (a) except for data points in the ILR-transformed space $\mathbb{R}^2$. (c) & (d) Same as (a) except for the Dirichlet depth and the generalized Mahalanobis depth, respectively.

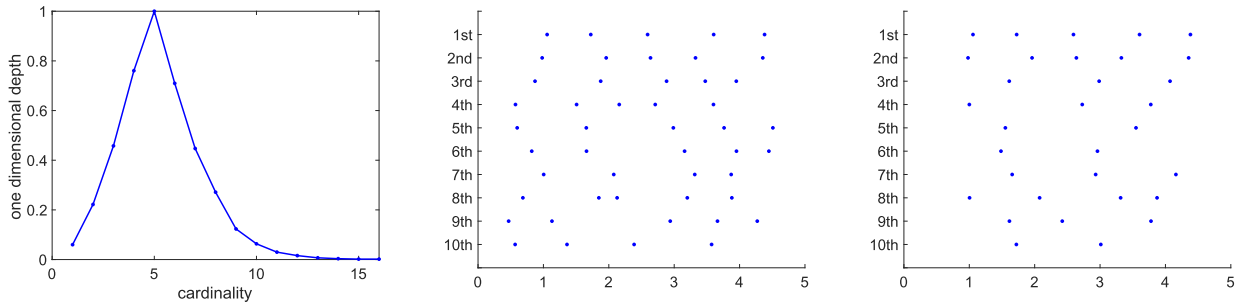
5.1. Simulation of HPP

Assume the time domain is $[0, 2]$ and the intensity function is a constant value 1. Then 1000 HPP realizations with cardinality being 2 are generated. For each realization, the IET is a three dimensional vector in the simplex $\mathcal{S}^3$. Thus, a 2-dimensional ternary plot (Pawlowsky-Glahn et al., 2007) can be conducted together with the ILR depth value as the contour. The result is shown in Fig. 2(a). The depth values and their contours in the transformed Euclidean space are shown in Fig. 2(b). From Fig. 2(a), the contour value decreases from the center to edges, and the depth value approaches 0 if the IET vector approaches the boundary. In addition, based on Fig. 2(b), the shape of the ILR transformed data looks similar to a regular triangle centered at origin. The inner contours of the ILR depth in $\mathbb{R}^2$ are closer to circular shapes, whereas the outside ones are similar to regular, smoothed triangles, which coincide with the shape of the contour of density in Eqn. (3.5) except the scale.

For comparison, we also show result using the Dirichlet depth (Qi et al., 2021) and the generalized Mahalanobis depth (Liu and Wu, 2017) on the same dataset in Fig. 2(c) and 2(d). We can see that the decreasing rate of the Dirichlet depth near the triangular center is much slower than that in the ILR depth. For the Dirichlet depth, many realizations in the middle part have large depth value, whereas the depth values decrease sharply when the data points approach the boundary. On the other hand, the decreasing rate of the ILR depth changes more slowly when the data points vary from center to boundary. Finally, the generalized Mahalanobis depth is not appropriate for HPP since it gives positive value to process at boundary and does not fit the triangular domain. This violates the Property 1 in Section 3.1.

In addition to the illustration of the conditional depth, we will use another simulation to show the rank of HPP realizations using the depth in Definition 1, where we use the ILR depth as the conditional depth given cardinality. Suppose the intensity function is $\lambda = 1$ on the time domain $[0, 5]$. From the definition of HPP, the expected number of time events is 5. This makes the normalized one dimensional depth $w(|\boldsymbol{s}|; P_{|S|})$ in Definition 1 obtains the maximum value when the cardinality of point process is 5. This result is clearly shown in Fig. 3(a).

Based on the overall depth in Definition 1, top-ranked processes are expected to have 5 evenly distributed events. In Fig. 3(b), we show the top 10 processes with largest overall depth value when $r = 1$. In this case, the one-dimensional depth plays an important role and most processes have 5 events and their distributions are nearly uniform on [0, 5]. In contrast, we also show the top 10 processes with largest overall depth value when $r = 0.1$ in Fig. 3(c). With a smaller



(a) one dimensional depth (b) depth rank, $r = 1$ (c) depth rank, $r = 0.1$

Fig. 3. One dimensional depth and point processes with top 10 overall depths. (a) One dimensional depth for different cardinalities in the simulated sample. (b) Each row is a realization of simulated HPP with the rank of depth value shown in vertical axis. The depth is computed by Definition 1, where the ILR depth is the conditional depth and the hyper-parameter $r = 1$. (c) Same as (b) except that $r = 0.1$.

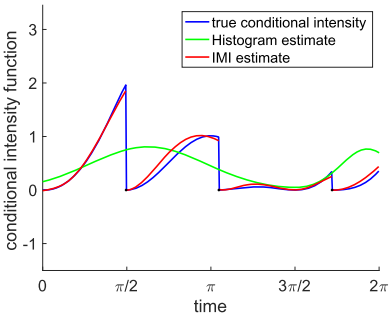


Fig. 4. Estimation of the conditional intensity in a typical realization. The three solid lines denote the true conditional intensity (blue), Histogram estimate (green), and IMI estimate (red), respectively.

weight coefficient, the one-dimensional depth contributes less to the overall depth value and we can see several top-10 processes have 2, 3, 4, or 5 events. Nevertheless, the event distribution in each process is still approximately homogeneous on [0, 5].

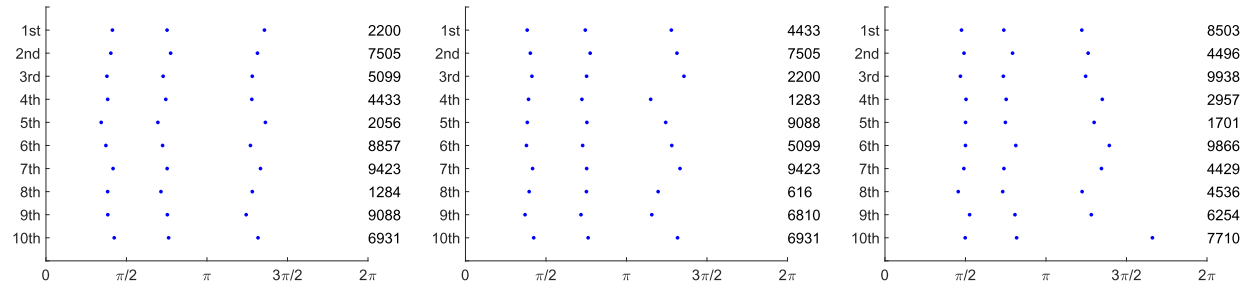
5.2. Simulation of non-Poisson process

We will use one example to illustrate the ILR depth for an IMI process based on Definition 2. Suppose the conditional intensity function is $\lambda(t|H_t) = (\sin(t) + 1) \cdot (\sin(t - s_*(t) - \frac{\pi}{2}) + 1)$, where $s_*(t)$ is the last time event preceding to t. 10000 realizations are generated in the time domain $[0, 2\pi]$. In this case, the conditional intensity function varies with respect to event history. We use Algorithm 2 to estimate the conditional intensity function as well as depth value for each realization. The conditional intensity estimate on one typical realization is shown in Fig. 4. We can see that the IMI model provides a much better estimate than the event-independent Histogram method.

Similar to the IPP study, the ranking performance of the 10000 realizations can be evaluated by selecting realizations with top 10 depth values. The result is shown in Fig. 5, where we display the indexed top 10 depth-valued processes using the true conditional intensity, estimated by the IMI method, and estimated by the Histogram method, respectively. We can see the processes using true intensity (Panel (a)) and IMI method (Panel (b)) have a lot of overlaps. Indeed, 7 indices, 2200, 7505, 5099, 4433, 9423, 9088 and 6931, appear in both Fig. 5(a) and 5(b) out of the top 10 realizations. However, in Fig. 5(c), none of the realizations depth are ranked as top 10 in Fig. 5(a). This result shows that the depth calculated by the IMI model is more accurate for general point process as compared to the Histogram method.

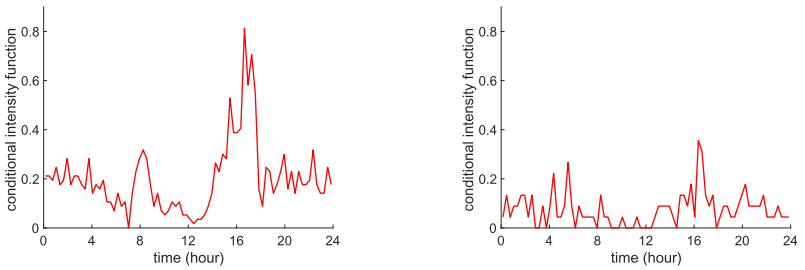
6. Real data application

In this section, we will illustrate the proposed depth method in a dataset from the real world. We consider the occurrence times of car accidents from 2016 to 2020 in Tallahassee, Florida, USA. The data can be retrieved at the link https://www.kaggle.com/sobhanmoosavi/us-accidents and was previously used in Moosavi et al. (2019a,b). Tallahassee is the capital city of Florida and there is a highway *I-10* located at the northern region of the city. Majority of people live in Tallahassee commute via local roads and the highway is mainly used by long-distance travelers. Our dataset includes accident occurrence times in two types of roads: highway (i.e., I-10) and local roads. For each type, the occurrence times are recorded in the time domain $[0, 24]$ in the units of hours.



(a) true intensity (b) IMI model (c) Histogram method

Fig. 5. Top 10 depth-valued processes for different conditional intensity function estimations. (a) Each row is a realization of simulated point process with the rank of depth value shown on left vertical axis and the index shown on right vertical axis. The depth is computed by Definition 1, where the conditional ILR depth is estimated using true conditional intensity in the Time Rescaling transformation and the hyper-parameter $r = 1$. (b) Same as (a) except that the IMI model is used in the Time Rescaling. (c) Same as (a) except that the Histogram method is used in the Time Rescaling.



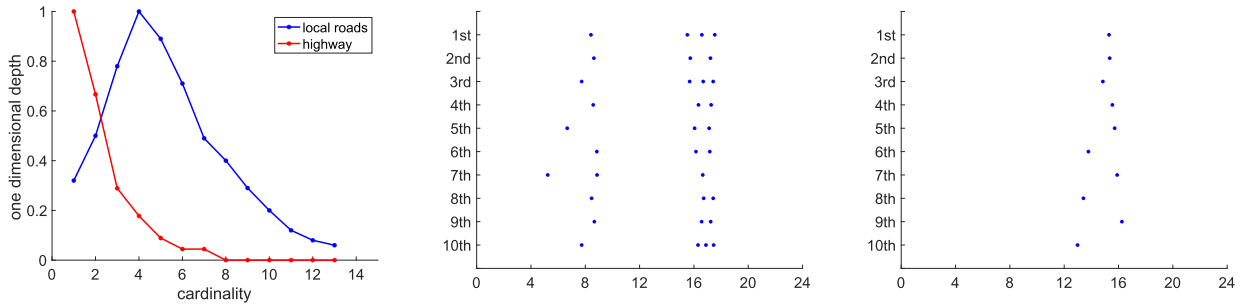
(a) local roads (b) highway

Fig. 6. Estimation result. (a) Estimated intensity function of local roads. (b) Estimated intensity function of highway.

If at least one car accident was recorded for a specific day, the accident occurrence times in that day are treated as a realization of a point process. Since each accident can be assumed to be independent of each other, we will consider the point process as an inhomogeneous Poisson process and focus on the accident occurrence pattern for both local roads and the highway I-10. Hence, the Histogram method introduced in Section 4.1 can be used to both data groups to estimate the intensity functions, and the result is shown in Fig. 6. We also conduct hypothesis test to verify the Poisson process assumptions and the sample independence for both local roads and highway. The detail is given in Appendix L.

From Fig. 6(a), we can see there exist two large peak regions for the estimated conditional intensity function for local roads: One is a global maximum region at around 5pm, which is the rush hour in the afternoon. The other is a local maximum region at around 8am, which is the rush hour in the morning. It is also interesting to notice that the value of intensity function is even higher at night than at noon in local roads, which indicates noon may be the safest time to drive in local roads of Tallahassee. In contrast, the peak of the intensity function in the highway, shown in Fig. 6(b), does not have any obvious pattern. There is a global maximum region around 4pm (starting time of rush hour in the afternoon), but there is no apparent peak in morning rush hours. This result clearly shows that accidents in the local roads are mainly affected by the rush hour traffic, whereas those on the highway are not very related to it.

Similar to the previous simulation studies, the realizations with top 10 depth values can be collected for each group (with hyper-parameter $r = 1$), and the result is shown in Fig. 7. Based on the estimated intensity function, the expected numbers of time events in a single realization are 4.25 and 1.38 for local road and highway, respectively, which means the frequency of realizations with 4 and 1 time events may be relatively larger in the local roads sample and highway sample, respectively. From Fig. 7(a), one can find the one dimensional depth in Definition 1 obtains larger values when the cardinality is around 3, 4 and 5 for local roads. This result can also be seen in Fig. 7(b) where the depths of realizations with 3 or 4 time events are ranked in the top 10. Similarly, from Fig. 7(a), one can find the one dimensional depth obtains a significant larger value when the cardinality is 1 for highway, which can also be seen in 7(c) where the depths of realizations with 1 time event are ranked in the top 10. Moreover, we can see that the patterns of the top 10 depth-valued realizations in local roads and highway perfectly match the estimated conditional intensity functions, respectively. That is, for local roads, the high rank processes have two or three events during the afternoon rush hour (4-6pm), and one event during the morning rush hour (8-9am). For highway, there is only one event in top processes at around 4pm. In summary, the ILR depth method provides a reasonable center-outward rank to summarize the given point process data.



(a) one dimensional depths (b) local roads depth rank (c) highway depth rank

Fig. 7. One dimensional depth and processes with top depth ranks. (a) One dimensional depth in Definition 1 for different number of time events for local roads (blue line) and highway (red line) samples. (b) Each row is a real case of one day car accidents occurred in local roads of Tallahassee with the rank of depth value shown on vertical axis. The depth is computed by Definition 1, where the ILR depth is used as the conditional depth and the hyper-parameter $r = 1$. (c) Same as (b) except for accidents occurred in the highway.

7. Summary and future work

In this paper, we have introduced a new framework to define depth for point process. The definition can be divided into two parts: (1) One dimensional depth that characterizes the depth for point process cardinality. (2) The ILR depth that measures the center-outward ranks for location of time events conditioned on cardinality. In the formal definition, we adopt the approach proposed in Qi et al. (2021) to define the one dimensional depth and we focus on the derivation of the ILR depth. The approach is based on the ILR transformation of HPP's inter-event time from a simplex space to a Euclidean space, and the depth is defined using the probability distribution of the ILR transformed data combined with the Time Rescaling method. We also conduct a thorough study on the mathematical properties of the ILR depth, in which we introduce a new notion called *simplicial center* to satisfy the desirable properties. We introduce two algorithms to compute the ILR depth and conduct the asymptotic study when the process is a Poisson process. We also compare the ILR depth with two previous methods in literature, the generalized Mahalanobis depth (Liu and Wu, 2017) and the Dirichlet depth (Qi et al., 2021), and demonstrate the superiority of the new method. Finally, we use a real world dataset on car accidents to show the effectiveness of our new framework.

The ILR depth has various potential applications. Firstly, it can define a "median point process" which is a robust first-order summary statistic used to measure the central tendency of point processes. Secondly, the depth-based outlier detection method can be applied to time sequence data. Previous studies (Zhu et al., 2020; Liu and Hauskrecht, 2021) have mainly focused on identifying unexpected occurrences or absence of individual time events, while research on evaluating if a whole point process is an outlier is still under-explored. Our depth provides a reasonable approach to address this problem. Additionally, the ILR depth will contribute to point process classification, which is a common pattern recognition problem in neuroscience studies.

The ILR depth is a novel and rigorous approach to define depth of point process conditioned on cardinality. The apparent advantage is that the ILR depth is a distribution-based depth that can capture the real data pattern and provide geometric interpretation in Euclidean space. There are clear topics for us to further investigate in the future. At first, the density function in Eqn. (3.5) is under-explored, and more work can be conducted to examine its shape in Euclidean space when the norm of $\boldsymbol{u}^*$ gets larger. In addition, the difference between the density in Eqn. (3.5) and Gaussian density can be studied via the Kullback-Leibler divergence or Fisher-Rao distance (Srivastava et al., 2011) to obtain a more comprehensive understanding of the density function of $\boldsymbol{u}^*$. Next, other than the Time Rescaling, we can explore more generalization methods of the ILR depth from HPP to general point process. Finally, in this paper, the ILR depth is only applied to one-dimensional temporal point process. Further investigations will be explored to extend the proposed depth framework to multivariate point process.

Appendix A. Proof of Proposition 3.1

Since $\Psi\Psi^T = I_k$ and $\Psi^T\Psi = I_{k+1} - \frac{1}{k+1}\mathbf{1}_{k+1}^T\mathbf{1}_{k+1}$, it is easy to verify that $\Psi\mathbf{1}_{k+1} = \mathbf{0}$, thus, origin is the center of this polyhedron. Moreover, for each column $\Psi_{:,p}$, $p = 1, 2, \ldots, k+1$, $\|\Psi_{:,p}\|^2 = \frac{k}{k+1}$, where $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^k$, $\langle\Psi_{:,p}, \Psi_{:,q}\rangle = -\frac{1}{k+1}$ for distinct $p, q = 1, 2, \ldots, k+1$, then for any $i, j = 1, 2, \ldots, k+1$ and $i \neq j$, $\|\Psi_{:,i} - \Psi_{:,j}\|^2 = \|\Psi_{:,i}\|^2 + \|\Psi_{:,j}\|^2 - 2\langle\Psi_{:,i}, \Psi_{:,j}\rangle = \frac{2k}{k+1} + \frac{2}{k+1} = 2$.

Appendix B. Qualitative explanation of the depth properties

1. The conditional depth is a continuous function in point process space $\mathbb{S}_k$. Also, the depth value approaches 0 iff $\boldsymbol{s}$ approaches the boundary.

2. There exists a unique center in $\mathbb{S}_k$ such that the depth obtains the maximal value at the center. In addition, there should be a well-defined symmetry to establish the center.
3. Once the unique center with maximal depth value is fixed, the depth value should decrease along any linear trajectory within the space. However, there is no well-established linear operation in point process space $\mathbb{S}_k$. Thus, this property can only be introduced and verified in simplex space $\mathcal{S}^{k+1}$.
4. If the point process as well as the time domain is shifted or zoomed, then the depth value should remain unchanged.

Appendix C. Derivation about the pdf of ILR transformation of IET

In Eqn. (3.4), if the cardinality of a given homogeneous Poisson process is k, to figure out the kernel of the density function, $\det(J)$ need to be solved in terms of $\boldsymbol{u}^*$. For any $j, s = 1, \ldots, k$, the element of J is:

$$\begin{aligned}\frac{du_j}{du_s^*} = \; & (T_2 - T_1) \cdot \frac{\Psi_{s,j} \exp(\sum_{i=1}^k u_i^* \Psi_{i,j})(\sum_{p=1}^{k+1} \exp(\sum_{i=1}^k u_i^* \Psi_{i,p}))}{(\sum_{p=1}^{k+1} \exp(\sum_{i=1}^k u_i^* \Psi_{i,p}))^2} \\ & -(T_2 - T_1) \cdot \frac{\exp(\sum_{i=1}^k u_i^* \Psi_{i,j})(\sum_{p=1}^{k+1} \Psi_{s,p} \exp(\sum_{i=1}^k u_i^* \Psi_{i,p}))}{(\sum_{p=1}^{k+1} \exp(\sum_{i=1}^k u_i^* \Psi_{i,p}))^2}.\end{aligned}$$

Therefore, $\det(J)$ can be expressed as follows. To save space, denote $A_{i,j,p} = \Psi_{i,j} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}}$.

$$\begin{aligned}\det(J) = & \frac{(T_2 - T_1)^k}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}})^{2k}} \times \\ & \begin{vmatrix} e^{\sum_{i=1}^k u_i^* \Psi_{i,1}} \left(\sum_{p=1}^{k+1} (A_{1,1,p} - A_{1,p,p}) \right) & \cdots & e^{\sum_{i=1}^k u_i^* \Psi_{i,1}} \left(\sum_{p=1}^{k+1} (A_{k,1,p} - A_{k,p,p}) \right) \\ e^{\sum_{i=1}^k u_i^* \Psi_{i,2}} \left(\sum_{p=1}^{k+1} (A_{1,2,p} - A_{1,p,p}) \right) & \cdots & e^{\sum_{i=1}^k u_i^* \Psi_{i,2}} \left(\sum_{p=1}^{k+1} (A_{k,2,p} - A_{k,p,p}) \right) \\ \vdots & \ddots & \vdots \\ e^{\sum_{i=1}^k u_i^* \Psi_{i,k}} \left(\sum_{p=1}^{k+1} (A_{1,k,p} - A_{1,p,p}) \right) & \cdots & e^{\sum_{i=1}^k u_i^* \Psi_{i,k}} \left(\sum_{p=1}^{k+1} (A_{k,k,p} - A_{k,p,p}) \right) \end{vmatrix} \\ = & \frac{(T_2 - T_1)^k \prod_{p=1}^k e^{\sum_{i=1}^k u_i^* \Psi_{i,p}}}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}})^{2k}} \begin{vmatrix} \sum_{p=1}^{k+1} (A_{1,1,p} - A_{1,p,p}) & \cdots & \sum_{p=1}^{k+1} (A_{k,1,p} - A_{k,p,p}) \\ \sum_{p=1}^{k+1} (A_{1,2,p} - A_{1,p,p}) & \cdots & \sum_{p=1}^{k+1} (A_{k,2,p} - A_{k,p,p}) \\ \vdots & \ddots & \vdots \\ \sum_{p=1}^{k+1} (A_{1,k,p} - A_{1,p,p}) & \cdots & \sum_{p=1}^{k+1} (A_{k,k,p} - A_{k,p,p}) \end{vmatrix} \\ = & \frac{(T_2 - T_1)^k \prod_{p=1}^k e^{\sum_{i=1}^k u_i^* \Psi_{i,p}}}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}})^{2k}} \begin{vmatrix} \sum_{p=1}^{k+1} (A_{1,1,p} - A_{1,p,p}) & \cdots & \sum_{p=1}^{k+1} (A_{k,1,p} - A_{k,p,p}) \\ \sum_{p=1}^{k+1} (A_{1,2,p} - A_{1,1,p}) & \cdots & \sum_{p=1}^{k+1} (A_{k,2,p} - A_{k,1,p}) \\ \vdots & \ddots & \vdots \\ \sum_{p=1}^{k+1} (A_{1,k,p} - A_{1,1,p}) & \cdots & \sum_{p=1}^{k+1} (A_{k,k,p} - A_{k,1,p}) \end{vmatrix} \\ = & \frac{(T_2 - T_1)^k \prod_{p=1}^k e^{\sum_{i=1}^k u_i^* \Psi_{i,p}}}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}})^{k+1}} \begin{vmatrix} \sum_{p=1}^{k+1} (A_{1,1,p} - A_{1,p,p}) & \cdots & \sum_{p=1}^{k+1} (A_{k,1,p} - A_{k,p,p}) \\ \Psi_{1,2} - \Psi_{1,1} & \cdots & \Psi_{k,2} - \Psi_{k,1} \\ \vdots & \ddots & \vdots \\ \Psi_{1,k} - \Psi_{1,1} & \cdots & \Psi_{k,k} - \Psi_{k,1} \end{vmatrix} \\ = & \frac{(T_2 - T_1)^k \prod_{p=1}^k e^{\sum_{i=1}^k u_i^* \Psi_{i,p}}}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}})^{k+1}} \left(\sum_{p=1}^{k+1} \begin{vmatrix} A_{1,1,p} - A_{1,p,p} & \cdots & A_{k,1,p} - A_{k,p,p} \\ \Psi_{1,2} - \Psi_{1,1} & \cdots & \Psi_{k,2} - \Psi_{k,1} \\ \vdots & \ddots & \vdots \\ \Psi_{1,k} - \Psi_{1,1} & \cdots & \Psi_{k,k} - \Psi_{k,1} \end{vmatrix} \right) \\ = & \frac{(T_2 - T_1)^k \prod_{p=1}^k e^{\sum_{i=1}^k u_i^* \Psi_{i,p}}}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}})^{k+1}} \left(\sum_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}} \begin{vmatrix} \Psi_{1,1} - \Psi_{1,p} & \cdots & \Psi_{k,1} - \Psi_{k,p} \\ \Psi_{1,2} - \Psi_{1,1} & \cdots & \Psi_{k,2} - \Psi_{k,1} \\ \vdots & \ddots & \vdots \\ \Psi_{1,k} - \Psi_{1,1} & \cdots & \Psi_{k,k} - \Psi_{k,1} \end{vmatrix} \right). \end{aligned}$$

Next, denote $D_p = \begin{vmatrix} \Psi_{1,1}-\Psi_{1,p} & \dots & \Psi_{k,1}-\Psi_{k,p} \\ \Psi_{1,2}-\Psi_{1,1} & \dots & \Psi_{k,2}-\Psi_{k,1} \\ \vdots & \ddots & \vdots \\ \Psi_{1,k}-\Psi_{1,1} & \dots & \Psi_{k,k}-\Psi_{k,1} \end{vmatrix}$ for $p = 1, \dots, k$. If $p = 1$,

$$D_1 = \begin{vmatrix} 0 & \dots & 0 \\ \Psi_{1,2}-\Psi_{1,1} & \dots & \Psi_{k,2}-\Psi_{k,1} \\ \vdots & \ddots & \vdots \\ \Psi_{1,k}-\Psi_{1,1} & \dots & \Psi_{k,k}-\Psi_{k,1} \end{vmatrix} = 0.$$

If $p = 2, \dots, k$, when calculating D_p, the first row can be added with the p-th row, then the first row will become 0 and the determinant is unchanged. Thus, if $p = 2, \dots, k$,

$$D_p = \begin{vmatrix} 0 & \dots & 0 \\ \Psi_{1,2}-\Psi_{1,1} & \dots & \Psi_{k,2}-\Psi_{k,1} \\ \vdots & \ddots & \vdots \\ \Psi_{1,k}-\Psi_{1,1} & \dots & \Psi_{k,k}-\Psi_{k,1} \end{vmatrix} = 0.$$

Therefore, apply this result to the formula of $\det(J)$,

$$\begin{aligned} \det(J) &= \frac{(T_2-T_1)^k \prod_{p=1}^k e^{\sum_{i=1}^k u_i^* \Psi_{i,p}}}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}})^{k+1}} \left(e^{\sum_{i=1}^k u_i^* \Psi_{i,k+1}} \begin{vmatrix} \Psi_{1,1}-\Psi_{1,k+1} & \dots & \Psi_{k,1}-\Psi_{k,k+1} \\ \Psi_{1,2}-\Psi_{1,1} & \dots & \Psi_{k,2}-\Psi_{k,1} \\ \vdots & \ddots & \vdots \\ \Psi_{1,k}-\Psi_{1,1} & \dots & \Psi_{k,k}-\Psi_{k,1} \end{vmatrix} \right) \\ &= \frac{(T_2-T_1)^k \prod_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}}}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}})^{k+1}} \begin{vmatrix} \Psi_{1,1}-\Psi_{1,k+1} & \dots & \Psi_{k,1}-\Psi_{k,k+1} \\ \Psi_{1,2}-\Psi_{1,1} & \dots & \Psi_{k,2}-\Psi_{k,1} \\ \vdots & \ddots & \vdots \\ \Psi_{1,k}-\Psi_{1,1} & \dots & \Psi_{k,k}-\Psi_{k,1} \end{vmatrix} \\ &= \frac{(T_2-T_1)^k e^{\sum_{i=1}^k u_i^* (\sum_{p=1}^{k+1} \Psi_{i,p})}}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}})^{k+1}} \begin{vmatrix} \Psi_{1,1}-\Psi_{1,k+1} & \dots & \Psi_{k,1}-\Psi_{k,k+1} \\ \Psi_{1,2}-\Psi_{1,1} & \dots & \Psi_{k,2}-\Psi_{k,1} \\ \vdots & \ddots & \vdots \\ \Psi_{1,k}-\Psi_{1,1} & \dots & \Psi_{k,k}-\Psi_{k,1} \end{vmatrix} \\ &= \frac{(T_2-T_1)^k}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}})^{k+1}} \begin{vmatrix} \Psi_{1,1}-\Psi_{1,k+1} & \dots & \Psi_{k,1}-\Psi_{k,k+1} \\ \Psi_{1,2}-\Psi_{1,1} & \dots & \Psi_{k,2}-\Psi_{k,1} \\ \vdots & \ddots & \vdots \\ \Psi_{1,k}-\Psi_{1,1} & \dots & \Psi_{k,k}-\Psi_{k,1} \end{vmatrix}. \end{aligned}$$

Therefore, the pdf of the ILR transformation for k-event HPP is:

$$\begin{aligned} f_{\boldsymbol{u}^*}(u_1^*, \dots u_k^*) &= \frac{k!}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}})^{k+1}} \begin{vmatrix} \Psi_{1,1}-\Psi_{1,k+1} & \dots & \Psi_{k,1}-\Psi_{k,k+1} \\ \Psi_{1,2}-\Psi_{1,1} & \dots & \Psi_{k,2}-\Psi_{k,1} \\ \vdots & \ddots & \vdots \\ \Psi_{1,k}-\Psi_{1,1} & \dots & \Psi_{k,k}-\Psi_{k,1} \end{vmatrix} \\ &= \frac{c}{(\sum_{p=1}^{k+1} e^{\sum_{i=1}^k u_i^* \Psi_{i,p}})^{k+1}}, \end{aligned}$$

where c is the positive constant that guarantees the integral of the density as 1.

Appendix D. Proof of log-concavity of the density in Eqn. (3.5)

By the property of probability density function, if the density is log-concave, then it is a uni-modal shape curve. Therefore, the remaining task is to prove the density function in Eqn. (3.5) is log-concave for any positive integer k.

To prove $\log(f_{\boldsymbol{u}^*}(u_1^*, \dots u_k^*))$ is a concave function, the Hessian matrix need to be found in closed form. In this way, the first and second order partial derivative of $\log(f_{\boldsymbol{u}^*}(u_1^*, \dots u_k^*))$ can be computed as follows:

$$\frac{\partial}{\partial u_s^*}\log(f_{\boldsymbol{u}^*}) = -(k+1)\cdot\frac{\sum_{p=1}^{k+1}\Psi_{s,p}e^{\sum_{i=1}^{k}u_i^*\Psi_{i,p}}}{\sum_{p=1}^{k+1}e^{\sum_{i=1}^{k}u_i^*\Psi_{i,p}}},$$

$$\frac{\partial^2}{\partial u_s^{*2}}\log(f_{\boldsymbol{u}^*}) = -(k+1)\cdot\frac{\sum_{p=1}^{k}\sum_{q>p}(\Psi_{s,p}-\Psi_{s,q})^2 e^{\sum_{i=1}^{k}u_i^*\Psi_{i,p}}e^{\sum_{i=1}^{k}u_i^*\Psi_{i,q}}}{\left(\sum_{p=1}^{k+1}e^{\sum_{i=1}^{k}u_i^*\Psi_{i,p}}\right)^2},$$

$$\frac{\partial^2}{\partial u_s^* du_t^*}\log(f_{\boldsymbol{u}^*}) = -(k+1)\cdot\frac{\sum_{p=1}^{k}\sum_{q>p}(\Psi_{s,p}-\Psi_{s,q})(\Psi_{t,p}-\Psi_{t,q})e^{\sum_{i=1}^{k}u_i^*\Psi_{i,p}}e^{\sum_{i=1}^{k}u_i^*\Psi_{i,q}}}{\left(\sum_{p=1}^{k+1}e^{\sum_{i=1}^{k}u_i^*\Psi_{i,p}}\right)^2}.$$

Next, denote the Hessian matrix H as:

$$H = \begin{pmatrix} \frac{\partial^2}{\partial u_1^{*2}}\log(f_{\boldsymbol{u}^*}) & \frac{\partial^2}{\partial u_1^*\partial u_2^*}\log(f_{\boldsymbol{u}^*}) & \dots & \frac{\partial^2}{\partial u_1^*\partial u_k^*}\log(f_{\boldsymbol{u}^*}) \\ \frac{\partial^2}{\partial u_2^*\partial u_1^*}\log(f_{\boldsymbol{u}^*}) & \frac{\partial^2}{\partial u_2^{*2}}\log(f_{\boldsymbol{u}^*}) & \dots & \frac{\partial^2}{\partial u_2^*\partial u_k^*}\log(f_{\boldsymbol{u}^*}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial u_k^*\partial u_1^*}\log(f_{\boldsymbol{u}^*}) & \frac{\partial^2}{\partial u_k^*\partial u_2^*}\log(f_{\boldsymbol{u}^*}) & \dots & \frac{\partial^2}{\partial u_k^{*2}}\log(f_{\boldsymbol{u}^*}) \end{pmatrix}.$$

Due to the property of concavity, a multivariate function is concave if and only if its Hessian matrix is negative definite. Denote $A_{s,t} = \sqrt{e^{\sum_{i=1}^{k}u_i^*\Psi_{i,s}}e^{\sum_{i=1}^{k}u_i^*\Psi_{i,t}}}$ and a $(k\times\frac{k(k+1)}{2})$ matrix B as:

$$B = \begin{pmatrix} (\Psi_{1,1}-\Psi_{1,2})A_{1,2} & (\Psi_{1,1}-\Psi_{1,3})A_{1,3} & \dots & (\Psi_{1,k}-\Psi_{1,k+1})A_{k,k+1} \\ (\Psi_{2,1}-\Psi_{2,2})A_{1,2} & (\Psi_{2,1}-\Psi_{2,3})A_{1,3} & \dots & (\Psi_{2,k}-\Psi_{2,k+1})A_{k,k+1} \\ \vdots & \vdots & \ddots & \vdots \\ (\Psi_{k,1}-\Psi_{k,2})A_{1,2} & (\Psi_{k,1}-\Psi_{k,3})A_{1,3} & \dots & (\Psi_{k,k}-\Psi_{k,k+1})A_{k,k+1} \end{pmatrix}.$$

After some algebra, H can be expressed as:

$$H = -\frac{k+1}{\left(\sum_{p=1}^{k+1}e^{\sum_{i=1}^{k}u_i^*\Psi_{i,p}}\right)^2}\cdot BB^T.$$

Since BB^T is a positive definite matrix, H is negative definite. Thus, the pdf in Eqn. (3.5) is log-concave and uni-modal. Finally, when $\boldsymbol{u}^* = (0,0,\dots,0)^T$, take into account that the sum of each row of Ψ is 0, then, the first partial derivative $\frac{\partial}{\partial u_s^*}\log(f_{\boldsymbol{u}^*})$ equals to 0 for each $s = 1,2,\dots,k$. Therefore, the origin in Euclidean space $\mathbb{R}^k$ is the global maximum point of the density in Eqn. (3.5).

Appendix E. Simplified version of the ILR depth

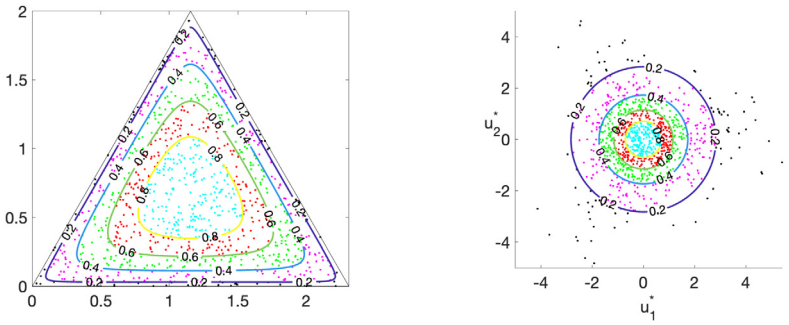
In general, because of log-concavity, one may propose to use a normal distribution to approximate the density function in Eqn. (3.5) for the purposes of simplification or efficiency. This normal approximation adopts the idea of the Laplacian approximation, and it will actually lead to an alternative approach of depth definition later in this paper. In our framework, it is interesting to find that the approximated normal distribution has a standard form (proof is given in Appendix G):

Proposition Appendix E.1. *The normal approximation of the density function in Eqn.* (3.5) *is the k-dimensional standard multivariate normal distribution* $N(0, I_k)$.

Thus, similar with the definition of the ILR depth, another depth formula can be formally defined via the standard Gaussian density. For simplicity, we will only show the definition for HPP case. It can be extended to more general case by Time Rescaling method. The definition is given in the form of a Mahalanobis depth (Liu and Singh, 1993) as follows:

Definition 3. Let $\boldsymbol{s} = (s_1, s_2, \dots, s_k) \in \mathbb{S}_k$ be a realization of an HPP in the time domain $[T_1, T_2]$ with $T_1 < s_1 < s_2 < \cdots < s_k < T_2$. Denote $s_0 = T_1$, $s_{k+1} = T_2$, $g_s = (\prod_{i=1}^{k+1}(s_i - s_{i-1}))^{\frac{1}{k+1}}$, $\boldsymbol{u} = (u_1, u_2, \dots, u_{k+1})^T = (s_1 - s_0, s_2 - s_1, \dots, s_{k+1} - s_k)^T$ as the IET and $\boldsymbol{u}^* = (u_1^*, u_2^*, \dots, u_k^*)^T$ as the ILR transformation of $\boldsymbol{u}$. Then, a simplified version of the ILR depth of $\boldsymbol{s}$ conditioned on $|\boldsymbol{s}| = k$ is defined as:

$$D_c(\boldsymbol{s}; P_{S||S|=k}) = \frac{1}{1+\frac{1}{2}\|\boldsymbol{u}^*\|^2} = \frac{1}{1+\frac{1}{2}\sum_{i=1}^{k+1}(\log\frac{s_i-s_{i-1}}{g_s})^2}.$$



(a) data points in simplex (b) data points after ILR

Fig. E.8. Simplified ILR depth result in HPP. (a) Data points in simplex shown by ternary plot, where the points are colored with respect to ranges of the simplified ILR depth values. The solid lines indicate the depth contours with specific values. (b) Same as (a) except for data points in the ILR-transformed space $\mathbb{R}^2$.

If $\boldsymbol{s} \in \mathbb{B}_k$, i.e., at least two of $s_0, s_1, \ldots, s_{k+1}$ are identical, then $D_c(\boldsymbol{s}; P_{S||S|=k}) = 0$.

Definition 3 is a simplified conditional depth definition for homogeneous Poisson process. It cannot capture the real data pattern as accurately as the ILR depth because of the approximation. However, the contours of this new depth are all hyperspheres centered at origin. All classical symmetries can be satisfied in this simplified depth and there is no need to introduce a simplicial center as in the ILR depth. This simplified depth represents the transformed IET using conventional Mahalanobis depth (Liu and Singh, 1993) and satisfies all necessary important properties. This result is given in the following proposition (see proof in Appendix H).

Proposition Appendix E.2. *All of the four mathematical properties in Section 3.1 hold for the simplified ILR depth in Definition 3.*

We also illustrate the simplified ILR depth using simulations. Using the same simulated realizations of HPP in Section 5.1, the result is shown in Fig. E.8. Comparing Fig. 2(a) and Fig. E.8(a), there is no obvious difference between contours in ternary plot for two depths. However, when comparing corresponding depths in Euclidean space in Fig. 2(b) and Fig. E.8(b), the contours of ILR depth match the data pattern more accurately. This is reasonable as the model-based depth should rely on the accuracy of the model – better models are expected to result in better ranks.

Appendix F. Proof of mathematical properties of the ILR depth

We will first verify the properties for HPP case, then extend it to more general case. For HPP case, the four properties are verified as follows:

1. Based on Definition 2, this part is trivial.
2. Based on Proposition 3.4, this part is trivial.
3. Denote the center of ILR depth as $\boldsymbol{u}_c$ in $\mathcal{S}^{k+1}$. Based on Proposition 3.4, $\boldsymbol{u}_c = (\frac{T_2-T_1}{k+1}, \frac{T_2-T_1}{k+1}, \ldots, \frac{T_2-T_1}{k+1})^T$. Thus, according to Pawlowsky-Glahn et al. (2007), for any $\boldsymbol{u} = (u_1, u_2, \ldots, u_{k+1})^T \in \mathcal{S}^{k+1}$ and $\alpha \in [0, 1]$, denote $\mathcal{C}(\cdot)$ as the closure operation in simplex space, then:

$$
\begin{aligned}
&\boldsymbol{u}_c \oplus \alpha \odot (\boldsymbol{u} \ominus \boldsymbol{u}_c) \\
&= \left(\frac{T_2 - T_1}{k+1}, \ldots, \frac{T_2 - T_1}{k+1}\right)^T + \mathcal{C}\left(\left(\left(\frac{u_1(k+1)}{T_2 - T_1}\right)^\alpha, \ldots, \left(\frac{u_{k+1}(k+1)}{T_2 - T_1}\right)^\alpha\right)^T\right) \\
&= \mathcal{C}\left(\left(u_1^\alpha \left(\frac{k+1}{T_2 - T_1}\right)^{\alpha-1}, \ldots, u_{k+1}^\alpha \left(\frac{k+1}{T_2 - T_1}\right)^{\alpha-1}\right)^T\right) \\
&= \left(\frac{(T_2 - T_1)u_1^\alpha}{\sum_{i=1}^{k+1} u_i^\alpha}, \ldots, \frac{(T_2 - T_1)u_{k+1}^\alpha}{\sum_{i=1}^{k+1} u_i^\alpha}\right)^T.
\end{aligned}
$$

Thus, denote $\left(\boldsymbol{u}_c \oplus \alpha \odot (\boldsymbol{u} \ominus \boldsymbol{u}_c)\right)^*$ and $\boldsymbol{u}^*$ as the ILR transformation of $\boldsymbol{u}_c \oplus \alpha \odot (\boldsymbol{u} \ominus \boldsymbol{u}_c)$ and $\boldsymbol{u}$ respectively. Based on the ILR transformation formula, we have:

$$
\begin{aligned}
&\left(\boldsymbol{u}_c \oplus \alpha \odot (\boldsymbol{u} \ominus \boldsymbol{u}_c)\right)^* \\
&= \Psi \cdot \left(\log \frac{u_1^\alpha}{(\prod_{i=1}^{k+1} u_i)^{\frac{\alpha}{k+1}}}, \ldots, \log \frac{u_{k+1}^\alpha}{(\prod_{i=1}^{k+1} u_i)^{\frac{\alpha}{k+1}}} \right)^T \\
&= \alpha \Psi \cdot \left(\log \frac{u_1}{(\prod_{i=1}^{k+1} u_i)^{\frac{1}{k+1}}}, \ldots, \log \frac{u_{k+1}}{(\prod_{i=1}^{k+1} u_i)^{\frac{1}{k+1}}} \right)^T \\
&= \alpha \boldsymbol{u}^*.
\end{aligned}
$$

Since $\alpha \in [0, 1]$ and the ILR depth contour takes the same shape as the log-concave density of $\boldsymbol{u}^*$, it is easy to verify that $d_c(\boldsymbol{u}) \le d_c(\boldsymbol{u}_c \oplus \alpha \odot (\boldsymbol{u} \ominus \boldsymbol{u}_c))$.

4. According to Definition 2, ILR depth remains invariant to scaling and translation. Thus, this part is verified.

Next, for more general case, the verification can be extended as follows:

1. Based on Qi et al. (2021), $\Lambda_S(\cdot)$ is a continuous function for any general point process. Thus, the continuity will hold automatically. What is more, if there exists $i = 1, 2, \ldots, k+1$ such that $u_i \to 0$, then, no matter whether the conditional intensity function $\lambda(\cdot)$ is deterministic or not, $u_i' = \Lambda_S(s_i) - \Lambda_S(s_{i-1}) = \int_{T_1}^{s_i} \lambda(t|H_t)dt - \int_{T_1}^{s_{i-1}} \lambda(t|H_t)dt = \int_{s_{i-1}}^{s_i} \lambda(t|H_t)dt \to 0$. Therefore, the depth value will vanish at boundary.
2. If the process is inhomogeneous Poisson process, the conditional intensity function can be considered as a positive deterministic function $\lambda(\cdot)$. Based on the definition of $\Lambda_S(\cdot)$, $\Lambda_S(\cdot)$ is a strictly increasing function, and therefore is bijective function. Then, this property can be easily verified. If the point process is non-Poisson process, then there exist counter-examples such that the point process with maximum depth value is not unique (Qi et al., 2021). Thus, this property does not hold for non-Poisson process in general.
3. Similar as Property 2, this property does not hold for non-Poisson case. For IPP case, this property does not hold for the original IET $\boldsymbol{u}$. Since the depth is computed after Time Rescaling and $\Lambda_S(\cdot)$ is not linear function, the log-concave structure for $\boldsymbol{u}$ will not hold. However, if the process after Time Rescaling transformation is considered, it is trivial to find that the transformed IET $\boldsymbol{u}' = (\Lambda_S(s_1) - \Lambda_S(s_0), \ldots, \Lambda_S(s_{k+1}) - \Lambda_S(s_k))^T$ satisfies this property corresponding to the transformed center.
4. This part is trivial.

Appendix G. Proof of normal approximation result

According to Proposition 3.2, the global maximum point of the density in Eqn. (3.5) is the origin. Thus, with Taylor series expansion, $\log(f_{\boldsymbol{u}^*})$ can be rewritten as:

$$
\log(f_{\boldsymbol{u}^*}(u_1^*, \ldots, u_k^*)) \approx \log(f_{\boldsymbol{u}^*}(0, 0, \ldots, 0)) + D^T \boldsymbol{u}^* - \frac{1}{2} \boldsymbol{u}^{*T} H \boldsymbol{u}^*, \tag{G.1}
$$

where D is the first derivative of $\log(f_{\boldsymbol{u}^*})$ evaluated at origin and H is the negative of Hessian matrix of $\log(f_{\boldsymbol{u}^*})$ evaluated at origin. Considering that origin is the global maximum point, D is a column vector with all entries 0 and the second term at the right hand side of Eqn. (G.1) can be omitted. The remaining task is to figure out the closed form of H. According to Appendix D, given the properties of the matrix Ψ in ILR transformation such that $\sum_{p=1}^{k} \Psi_{s,p} = 0$, $\sum_{p=1}^{k} \Psi_{s,p}^2 = 1$ and $\sum_{p=1}^{k} \Psi_{s,p} \Psi_{t,p} = 0$ for each $s, t = 1, \ldots, k$, the following result can be easily obtained for each $s, t = 1, \ldots, k$.

$$
\begin{aligned}
-\frac{\partial^2}{\partial u_s^{*2}} \log(f_{\boldsymbol{u}^*})(0, 0, \ldots, 0) &= (k+1) \frac{(k+1) \sum_{p=1}^{k+1} \Psi_{s,p}^2 - (\sum_{p=1}^{k+1} \Psi_{s,p})^2}{(k+1)^2} \\
&= 1, \\
-\frac{\partial^2}{\partial u_s^* du_t^*} \log(f_{\boldsymbol{u}^*})(0, 0, \ldots, 0) &= (k+1) \frac{(k+1) \sum_{p=1}^{k+1} \Psi_{s,p} \Psi_{t,p} - (\sum_{p=1}^{k+1} \Psi_{s,p})(\sum_{p=1}^{k+1} \Psi_{t,p})}{(k+1)^2} \\
&= 0.
\end{aligned}
$$

Therefore, H is a $k \times k$ identity matrix. Finally, take exponential on both sides of Eqn. (G.1), the result is:

$$
f_{\boldsymbol{u}^*}(u_1^*, \ldots, u_k^*) \approx C \cdot e^{-\frac{1}{2} \boldsymbol{u}^{*T} H \boldsymbol{u}^*},
$$

where C is a positive constant and $e^{-\frac{1}{2} \boldsymbol{u}^{*T} H \boldsymbol{u}^*}$ is the kernel of a multivariate normal distribution with mean as the origin and covariance matrix as the inverse of H, which is the $k \times k$ identity matrix.

Appendix H. Proof of mathematical properties of the simplified version of ILR depth in Definition 3

1. Based on the definition, the depth value is continuous if $\boldsymbol{s} \notin \mathbb{B}_k$. Thus, the remaining task is to prove the depth function is continuous at boundary set $\mathbb{B}_k$, which is equivalent to prove the depth value approaches 0 if the point process approaches $\mathbb{B}_k$. For a given $\boldsymbol{s} = (s_1, s_2, \ldots, s_k) \in \mathbb{S}_k$, if there exists at least one $t = 1, \ldots, k+1$ such that $s_t - s_{t-1} \to 0$, one can find at least one $p = 1, \ldots, k+1$ such that $s_p - s_{p-1} \neq 0$ and is a finite positive number. In this case, the depth can be rewritten as:
$$D_c(\boldsymbol{s}; P_{S||S|=k}) = \frac{1}{1 + \frac{1}{2}\sum_{i=1}^{k+1}\left(\log \frac{s_i - s_{i-1}}{(\prod_{j=1}^{k+1}(s_j - s_{j-1}))^{\frac{1}{k+1}}}\right)^2} = \frac{1}{1 + \frac{1}{2}\sum_{i=1}^{k+1}\left(\log\left((s_i - s_{i-1})^{\frac{k}{k+1}} \cdot (\prod_{j\neq i}(s_j - s_{j-1}))^{-\frac{1}{k+1}}\right)\right)^2}.$$
According to the notation above, $(s_p - s_{p-1})^{\frac{k}{k+1}} \cdot (\prod_{j\neq p}(s_j - s_{j-1}))^{-\frac{1}{k+1}} \to \infty$ since $\prod_{j\neq p}(s_j - s_{j-1}) \to 0$. Consider the fact that the denominator part is the sum of 1 and $k+1$ positive terms, if one term approaches infinity, the denominator will approach infinity and this part is verified.
2. Consider the contour of standard multivariate Gaussian density function, origin is the center based on all classical symmetries (Zuo and Serfling, 2000). According to Proposition 3.4, $\boldsymbol{s} = (T_1 + \frac{T_2 - T_1}{k+1}, T_1 + \frac{2(T_2-T_1)}{k+1}, \ldots, T_1 + \frac{k(T_2-T_1)}{k+1})$ is the center.
3. This part is trivial due to the shape of contour of normal distribution and the proof of the ILR depth property.
4. This part is trivial.

Appendix I. Asymptotic theory for Algorithm 1

In this part, we will study the asymptotics of the estimated depth values using Algorithm 1. We will prove that the sample depth value computed with estimated $\hat{\lambda}(t)$ converges to the population depth value computed by using the true intensity function $\lambda(t)$. Notice that Algorithm 1 adopts the Histogram method, which is only applicable for Poisson process. We at first have a uniform convergence result about the estimated integrated intensity function in the following lemma, where the proof is given in Appendix J.

Lemma Appendix I.1. *Suppose $\lambda(t)$ is the true intensity function of a sample of IPP in $\mathbb{S}$ with sample size n. Assume $\sup_t \lambda(t) \leq R$ (R is a positive finite number) and $\lambda(t)$ is L-Lipschitz continuous on $[T_1, T_2]$, i.e., for any $x, y \in [T_1, T_2]$, $|\lambda(x) - \lambda(y)| \leq L|x - y|$ for a finite L. Denote $\hat{\lambda}(t)$ as the estimated intensity function based on Algorithm 1. Let M be the number of bins in Algorithm 1, and $\Lambda_S(x) = \int_{T_1}^{x} \lambda(t)dt$, $\hat{\Lambda}_S^{(n)}(x) = \int_{T_1}^{x} \hat{\lambda}(t)dt$, $T_1 \leq x \leq T_2$. Then the following uniform convergence rate holds:*
$$\sup_x |\hat{\Lambda}_S^{(n)}(x) - \Lambda_S(x)| = O\left(\frac{1}{M}\right) + O_P\left(\sqrt{\frac{M^2}{n}}\right).$$

Based on Lemma Appendix I.1, with some simple algebra, it is straightforward to conclude $M_{opt} = O(n^{\frac{1}{4}})$ is the optimal choice of M. Therefore, if $n \to \infty$ and $M \propto n^{\frac{1}{4}}$, $\hat{\Lambda}_S^{(n)}(x)$ uniformly converges to $\Lambda_S(x)$ in probability.

Using this result, we can obtain the main conclusion on the convergence of the sample ILR depth in the following theorem.

Theorem Appendix I.1. *Under the same assumptions as given in Lemma Appendix I.1, denote $D_{c-TR}(s; P_{S||S|=k}, \hat{\Lambda}_S^{(n)})$ as the sample ILR conditional depth computed by Algorithm 1 with estimated intensity function. Also denote $D_{c-TR}(s; P_{S||S|=k}, \Lambda_S)$ as the population ILR conditional depth computed by using the true conditional intensity function. Then we have,*
$$\sup_{\boldsymbol{s}\in\mathbb{S}_k} |D_{c-TR}(\boldsymbol{s}; P_{S||S|=k}, \hat{\Lambda}_S^{(n)}) - D_{c-TR}(\boldsymbol{s}; P_{S||S|=k}, \Lambda_S)| \to 0$$
in probability as $n \to \infty$.

Proof. Since $\boldsymbol{s} \in \mathbb{S}_k$ and $\mathbb{S}_k = \{\boldsymbol{s} = (s_1, s_2, \ldots, s_k) | T_1 \leq s_1 \leq s_2 \leq \cdots \leq s_k \leq T_2\}$, the Time Rescaling result $\Lambda_S(\boldsymbol{s})$ belongs to a bounded and closed set $\Lambda_S(\mathbb{S}_k) = \{(s_1, s_2, \ldots, s_k) | \Lambda_S(T_1) \leq s_1 \leq s_2 \leq \cdots \leq s_k \leq \Lambda_S(T_2)\}$. Thus, $\Lambda_S(\mathbb{S}_k)$ is a bounded and closed subset of Euclidean space $\mathbb{R}^k$. From *Heine–Borel theorem*, $\Lambda_S(\mathbb{S}_k)$ is a compact set and a continuous function defined on $\Lambda_S(\mathbb{S}_k)$ is a uniform continuous function. Consequently, $D_{c-TR}(s; P_{S||S|=k}, \Lambda_S)$ is a uniform continuous function on $\Lambda_S(\mathbb{S}_k)$. Therefore, from the *continuous mapping theorem*, $\sup_{s\in\mathbb{S}_k} |D_{c-TR}(s; P_{S||S|=k}, \hat{\Lambda}_S^{(n)}) - D_{c-TR}(s; P_{S||S|=k}, \Lambda_S)| \to 0$ in probability as $n \to \infty$. □

Appendix J. Proof of uniform convergence rate of Lemma Appendix I.1

First, rewrite $\hat{\lambda}(t)-\lambda(t)=\big(\hat{\lambda}(t)-\mathbb{E}[\hat{\lambda}(t)]\big)+\big(\mathbb{E}[\hat{\lambda}(t)]-\lambda(t)\big)$. We will consider the second part first, denote $N_i(t)$ as the number of events occurred until time t in realization i, then $\mathbb{E}[\hat{\lambda}(t)]-\lambda(t)$ can be rewritten as follows:

$$\begin{aligned}\mathbb{E}[\hat{\lambda}(t)]-\lambda(t)&=\frac{M}{n(T_2-T_1)}\sum_{i=1}^{n}\mathbb{E}\Big[\sum_{r=1}^{n_i}I(s_{ir}\in B_j)\Big]-\lambda(t)\\&=\frac{M}{n(T_2-T_1)}\sum_{i=1}^{n}\mathbb{E}\left[N_i\left(\frac{j(T_2-T_1)}{M}\right)-N_i\left(\frac{(j-1)(T_2-T_1)}{M}\right)\right]-\lambda(t)\\&=\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{E}[N_i(\frac{j(T_2-T_1)}{M})-N_i(\frac{(j-1)(T_2-T_1)}{M})]}{\frac{j(T_2-T_1)}{M}-\frac{(j-1)(T_2-T_1)}{M}}-\lambda(t).\end{aligned}$$

From Mean value theorem and the definition of $\lambda(t)$, since all of the n realizations have the same intensity function, thus, there exists $t^*\in\big[\frac{(j-1)(T_2-T_1)}{M},\frac{j(T_2-T_1)}{M}\big]$ such that for each $i=1,2,\ldots,n$, $\lambda(t^*)=\frac{\mathbb{E}[N_i(\frac{j(T_2-T_1)}{M})-N_i(\frac{(j-1)(T_2-T_1)}{M})]}{\frac{j(T_2-T_1)}{M}-\frac{(j-1)(T_2-T_1)}{M}}$. Then, $\mathbb{E}[\hat{\lambda}(t)]-\lambda(t)=\frac{1}{n}\sum_{i=1}^{n}\lambda(t^*)-\lambda(t)=\lambda(t^*)-\lambda(t)$. Since $\lambda(t)$ is L-Lipschitz continuous, $\mathbb{E}[\hat{\lambda}(t)]-\lambda(t)=\lambda(t^*)-\lambda(t)\le|\lambda(t^*)-\lambda(t)|\le L|t^*-t|\le\frac{L(T_2-T_1)}{M}$. This result can be generalized to every point t, then $\sup_t|\mathbb{E}[\hat{\lambda}(t)]-\lambda(t)|=O\left(\frac{1}{M}\right)$.

Next, consider the variance of $\hat{\lambda}(t)$, which will be used in later proof. For any $t\in B_j$, $j=1,2,\ldots,M$, $Var[\hat{\lambda}(t)]=\frac{M^2}{n^2(T_2-T_1)^2}\sum_{i=1}^{n}Var[\sum_{r=1}^{n_i}I(s_{ir}\in B_j)]$. Since $\sum_{r=1}^{n_i}I(s_{ir}\in B_j)$ denotes the total number of events in B_j for the i-th realization, then $\sum_{r=1}^{n_i}I(s_{ir}\in B_j)\sim Poisson(\int_{\frac{(j-1)(T_2-T_1)}{M}}^{\frac{j(T_2-T_1)}{M}}\lambda(t)dt)$, Thus,

$$\begin{aligned}Var[\hat{\lambda}(t)]&=\frac{M^2}{n^2(T_2-T_1)^2}\sum_{i=1}^{n}\int_{\frac{(j-1)(T_2-T_1)}{M}}^{\frac{j(T_2-T_1)}{M}}\lambda(t)dt\\&=\frac{M^2}{n^2(T_2-T_1)^2}\sum_{i=1}^{n}\frac{\lambda(t^{**})}{M}\\&\le R\cdot\frac{M}{n(T_2-T_1)^2},\end{aligned}$$

where t^{**} is a point within $[\frac{(j-1)(T_2-T_1)}{M},\frac{j(T_2-T_1)}{M}]$.

The remaining part is to focus on $\hat{\lambda}(t)-\mathbb{E}[\hat{\lambda}(t)]$. Denote $A_j=\frac{1}{n(T_2-T_1)}\sum_{i=1}^{n}\sum_{r=1}^{n_i}I(s_{ir}\in B_j)-\frac{1}{n(T_2-T_1)}\sum_{i=1}^{n}\mathbb{E}[N_i(\frac{j(T_2-T_1)}{M})-N_i(\frac{(j-1)(T_2-T_1)}{M})]$. Since,

$$\begin{aligned}\sup_t|\hat{\lambda}(t)-\mathbb{E}[\hat{\lambda}(t)]|&=\max_{j=1,2,\ldots,M}|M\cdot A_j|\\&=M\cdot\max_{j=1,2,\ldots,M}|A_j|.\end{aligned}$$

Thus, for any $\epsilon>0$,

$$\begin{aligned}\mathbb{P}\Big[\sup_t|\hat{\lambda}(t)-\mathbb{E}[\hat{\lambda}(t)]|>\epsilon\Big]&=\mathbb{P}\Big[M\cdot\max_{j=1,2,\ldots,M}|A_j|>\epsilon\Big]\\&=\mathbb{P}\Big[\max_{j=1,2,\ldots,M}|A_j|>\frac{\epsilon}{M}\Big]\\&=\mathbb{P}\Big[\bigcup_{j=1}^{M}\Big(|A_j|>\frac{\epsilon}{M}\Big)\Big]\\&\le\sum_{j=1}^{M}\mathbb{P}\Big[|A_j|>\frac{\epsilon}{M}\Big]\\&=\sum_{j=1}^{M}\mathbb{P}\Big[A_j^2>\frac{\epsilon^2}{M^2}\Big]\end{aligned}$$

$$\begin{aligned}
(\text{Chebyshev's inequality}) &\leq \sum_{j=1}^{M} \frac{Var(\frac{1}{n(T_2-T_1)} \sum_{i=1}^{n} \sum_{r=1}^{n_i} I(s_{ir} \in B_j))}{\epsilon^2/M^2} \\
(\text{Previous result about variance}) &\leq \sum_{j=1}^{M} \frac{R/(nM(T_2-T_1)^2)}{\epsilon^2/M^2} \\
&= R \cdot \frac{M^2}{n(T_2-T_1)^2\epsilon^2}.
\end{aligned}$$

Therefore,

$$\sup_t |\hat{\lambda}(t) - \mathbb{E}[\hat{\lambda}(t)]| = O_P\Big(\sqrt{\frac{M^2}{n}}\Big).$$

Combine with the previous result about $\mathbb{E}[\hat{\lambda}(t)] - \lambda(t)$, the uniform convergence rate of $\hat{\lambda}(t)$ is:

$$\sup_t |\hat{\lambda}(t) - \lambda(t)| = O\Big(\frac{1}{M}\Big) + O_P\Big(\sqrt{\frac{M^2}{n}}\Big).$$

Finally, the uniform convergence rate about $\hat{\Lambda}_S^{(n)}(x)$ can be derived as follows:

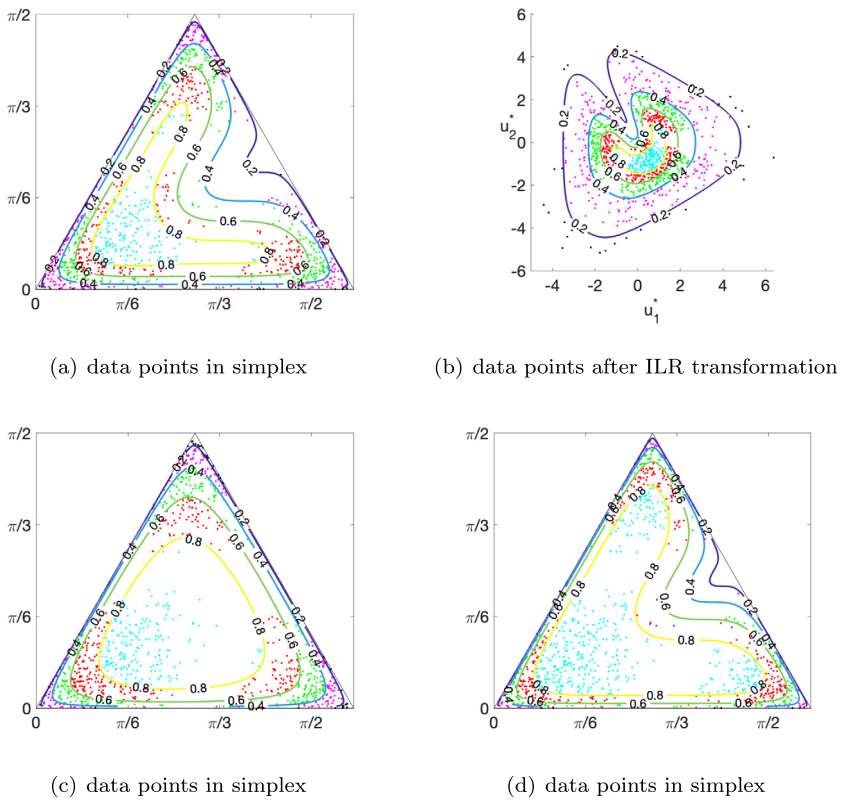
$$\begin{aligned}
\sup_x |\hat{\Lambda}_S^{(n)}(x) - \Lambda_S(x)| &= \sup_x \Big| \int_{T_1}^{x} \hat{\lambda}(t)dt - \int_{T_1}^{x} \lambda(t)dt \Big| \\
&= \sup_x \Big| \int_{T_1}^{x} (\hat{\lambda}(t) - \lambda(t))dt \Big| \\
&\leq \sup_x \int_{T_1}^{x} |\hat{\lambda}(t) - \lambda(t)|dt \\
&\leq \sup_x \int_{T_1}^{x} \sup_t |\hat{\lambda}(t) - \lambda(t)|ds \\
(\text{The integrand is non-negative}) &\leq \int_{T_1}^{T_2} \sup_t |\hat{\lambda}(t) - \lambda(t)|ds \\
&= O\Big(\frac{1}{M}\Big) + O_P\Big(\sqrt{\frac{M^2}{n}}\Big).
\end{aligned}$$

Appendix K. Simulation of IPP case

We will use two examples to illustrate the conditional ILR depth using ternary plots in IPP based on Definition 2.

Example 1: Let the intensity function $\lambda(t) = \cos(4t) + 1$ in the time domain $[0, \frac{\pi}{2}]$. This function decreases from 0 to $\frac{\pi}{4}$ and increases from $\frac{\pi}{4}$ to $\frac{\pi}{2}$ with $t = \frac{\pi}{4}$ as the global minimum point. One can generate 1000 independent realizations with cardinality 2 and make the ternary plot to show the contours of the ILR depth. The ILR depth values in the IET domain and Euclidean space are shown in Fig. K.9(a) and (b), respectively. We can clearly see the non-triangularly shaped contours in both plots, which demonstrates the inhomogeneity of the Poisson processes.

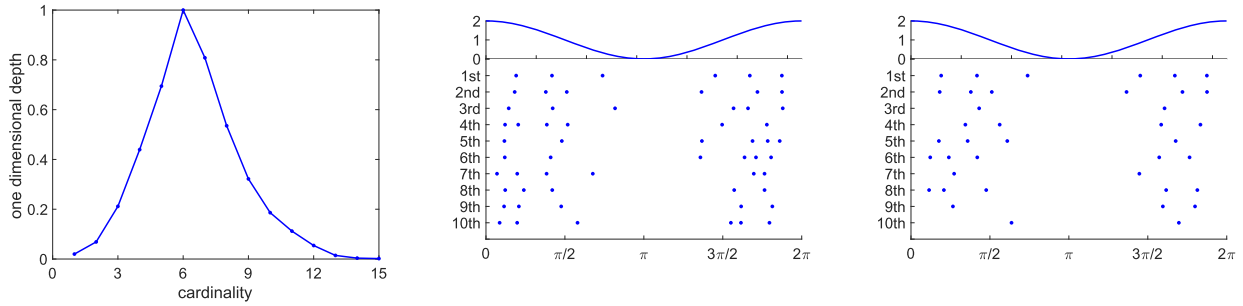
Next, the sample Dirichlet depth and the Dirichlet depth with Time Rescaling method in Qi et al. (2021) can be applied to the same IPP sample to compute conditional depth with cardinality 2. The results are shown in Fig. K.9(c) and (d). Given the fact that the value of intensity function is close to 0 around the middle range, it is not likely to have a process realization such that the second IET is small, meanwhile the first and third IET are almost the same. However, in Fig. K.9(c), this characteristic cannot be reflected. Thus, the sample Dirichlet depth (Qi et al., 2021) is not a good depth choice for IPP. Next, comparing Fig. K.9(a) with K.9(d), the shape of the two contours are identical. Nevertheless, there are fewer realizations located around the middle range of the right side in the ternary plot, e.g. there are fewer realizations with the two time events occurred around the middle part within the time domain. A reasonable depth should capture this feature and give



(a) data points in simplex (b) data points after ILR transformation

(c) data points in simplex (d) data points in simplex

Fig. K.9. Depth results of IPP with $\lambda(t) = \cos(4t) + 1$. (a) Data points in simplex shown by ternary plot, where the points are colored with respect to ranges of the ILR depth values given in Definition 2. The solid lines indicate the depth contours with specific values. (b) Same as (a) except for data points in the ILR transformed space $\mathbb{R}^k$. (c) Same as (a) except by using the sample Dirichlet depth in Qi et al. (2021). (d) Same as (a) except by using the Dirichlet depth after Time Rescaling in Qi et al. (2021).



(a) one dimensional depth (b) depth ranking if $r = 1$ (c) depth ranking if $r = 0.1$

Fig. K.10. One dimensional depth and top 10 depth-valued inhomogeneous point processes with $\lambda(t) = \cos(t) + 1$. (a) One dimensional depth for different cardinalities in the simulated sample. (b) $\lambda(t)$ is displayed in the top panel. In the bottom panel, each row is a realization of simulated IPP with the ranking of depth value shown in vertical axis. The depth is computed by Definition 1, where the ILR depth combined with Time Rescaling in Definition 2 is the conditional depth and the hyper-parameter $r = 1$. (c) Same as (b) except for $r = 0.1$.

much lower depth value to these realizations. Hence, the ILR depth is more reasonable since it satisfies this property better than the Dirichlet depth with Time Rescaling.

Example 2: We have shown the ILR depth given that the cardinality is 2. Now we will check the depth ranking performance for the overall depth in Definition 1. 1000 independent realizations are simulated with intensity function $\lambda(t) = \cos(t) + 1$ in the time domain $[0, 2\pi]$. Then the expected value of the number of time events of a realization is $\mathbb{E}[N(2\pi) - N(0)] = \int_0^{2\pi} (\cos(t) + 1)dt = 2\pi \approx 6$. Thus, a process with 6 events is expected to have a large one dimensional depth in Definition 1. In Fig. K.10, we display the one-dimensional depth with respect to different cardinalities. We also show the top 10 processes with the largest overall depths when the weight coefficient r = 1 and 0.1, respectively.

From Fig. K.10(a), one can conclude that the one dimensional depth obtains the maximum value when cardinality is 6. When $r = 1$, the one dimensional depth in Definition 1 is dominant, which makes realizations with cardinality 6 rank at the top places. This can be seen in Fig. K.10(b), where most realizations in top 10 places have 6 time events. When $r = 0.1$, the number of events is less dominant, and the distribution of event times becomes more important in the ranking process. Fig. K.10(c) shows that the realizations with depth in top 10 places have cardinality different from 6. However, the locations of time events of top 10 depth realizations are close to the density pattern of the intensity function (i.e., there are more events when intensity is large, and fewer events when intensity is small).

Appendix L. Support of the Poisson model and independence assumption of real data

To verify the Poisson model and independence assumption, consider the following hypothesis test for both the local roads/highway groups:

- H_0: All realizations are i.i.d. observations with a Poisson process model.
- H_1: All realizations are not i.i.d. observations with a Poisson process model.

Since there is no well-defined method to check "identically distributed" and "independent" simultaneously, we split the test into two parts to check the Poisson model and independence assumption.

1. Check if all point processes realizations follow one Poisson process model.
2. Check if all point processes are independent.

For the first part, we adopt the Sum-of-squared-spacings (3S) statistic introduced in Shchur et al. (2021) to evaluate if the realizations of two groups (local road and highway) can be modeled by inhomogeneous Poisson processes. The 3S statistic is able to test if a given realization follows a particular intensity function. Using the notation of our paper, the 3S statistic is defined as $\phi(\boldsymbol{s}) = \frac{1}{T_2 - T_1} \sum_{i=1}^{k+1} u_i^2$ for homogeneous Poisson process, where $\boldsymbol{s}$ is a given realization, k is the point process cardinality and u_i is each inter-event time. The test procedure for each group is given below:

- Estimate the intensity function based on Algorithm 1.
- Transform each realization based on Time Rescaling method.
- Compute 3S statistics for each realization.
- Compare $\phi(\boldsymbol{s})$ to the distribution of ϕ of a homogeneous Poisson process in $[T_1, T_2]$ and obtain its corresponding two-tailed p-value.
- After obtaining the p-values of all realizations, compute their FDR (false discovery rate)-adjusted p-values by Benjamini-Hochberg correction. Take the minimum one as the p-value for the first test part. The false discovery rate is 0.05.

After the computation, we find the minimum of the adjusted p-value for local road and highway group are 0.38 and 0.9, respectively. Both are greater than the commonly used significant level 0.05.

Next, for the second part, we move forward to check if the realizations in local roads/highway groups are independent with each other. To the best of our knowledge, there is no available hypothesis test to evaluate the joint independence among all temporal point process realizations. Cebrián et al. (2020) introduced a Poisson test to check if two inhomogeneous Poisson processes are independent. When conducting this test, the intensity function of at least one process must be known. Thus, the estimated intensity function from the first part is applied here. As our data are in chronological order, we choose to examine the pairwise independence between all temporally adjacent realizations to make the computation more efficient and robust. For each adjacent pair realizations, the null hypothesis is that they are independent. Then, we obtain a p-value for this pair. After the p-value of each adjacent pair is computed, the FDR correction will be adopted to obtain a single p-value. The minimum of the adjusted p-value for local roads and highway are 0.07 and 0.11, respectively. Both are greater than 0.05.

Finally, combining the results of the two parts, the FDR correction is applied one more time to obtain a single p-value for each group. Since both p-values are greater than 0.05, we can conclude that all realizations are approximately i.i.d. with a Poisson process model for both local roads and highway groups.

References

Aitchison, J., Barceló-Vidal, C., Martín-Fernández, J.A., Pawlowsky-Glahn, V., 2000. Logratio analysis and compositional distance. Math. Geol. 32, 271–275.
Brown, E.N., Barbieri, R., Ventura, V., Kass, R.E., Frank, L.M., 2002. The time-rescaling theorem and its application to neural spike train data analysis. Neural Comput. 14, 325–346.
Cebrián, A.C., Abaurrea, J., Asín, J., 2020. Testing independence between two nonhomogeneous point processes in time. J. Stat. Comput. Simul. 90, 2878–2901.
Dai, X., Lopez-Pintado, S., Initiative, A.D.N., 2021. Tukey's depth for object data. J. Am. Stat. Assoc., 1–37.
Dyckerhoff, R., Mosler, K., Koshevoy, G., 1996. Zonoid data depth: theory and computation. In: COMPSTAT. Springer, pp. 235–240.
Egozcue, J.J., Pawlowsky-Glahn, V., Mateu-Figueras, G., Barcelo-Vidal, C., 2003. Isometric logratio transformations for compositional data analysis. Math. Geol. 35, 279–300.

Fraiman, R., Meloche, J., García-Escudero, L.A., Gordaliza, A., He, X., Maronna, R., Yohai, V.J., Sheather, S.J., McKean, J.W., Small, C.G., et al., 1999. Multivariate l-estimation. Test 8, 255–317.
Geenens, G., Nieto-Reyes, A., Francisci, G., 2021. Statistical depth in abstract metric spaces. arXiv preprint. arXiv:2107.13779.
Kass, R.E., Ventura, V., 2001. A spike-train probability model. Neural Comput. 13, 1713–1720.
Liu, R.Y., 1990. On a notion of data depth based on random simplices. Ann. Stat., 405–414.
Liu, R.Y., Singh, K., 1993. A quality index based on data depth and multivariate rank tests. J. Am. Stat. Assoc. 88, 252–260.
Liu, S., Hauskrecht, M., 2021. Event outlier detection in continuous time. In: International Conference on Machine Learning. PMLR, pp. 6793–6803.
Liu, S., Wu, W., 2017. Generalized Mahalanobis depth in point process and its application in neural coding. Ann. Appl. Stat., 992–1010.
López-Pintado, S., Romo, J., 2009. On the concept of depth for functional data. J. Am. Stat. Assoc. 104, 718–734.
Makinde, O.S., 2019. Classification rules based on distribution functions of functional depth. Stat. Pap. 60, 629–640.
Moosavi, S., Samavatian, M.H., Parthasarathy, S., Ramnath, R., 2019a. A countrywide traffic accident dataset. arXiv preprint. arXiv:1906.05409.
Moosavi, S., Samavatian, M.H., Parthasarathy, S., Teodorescu, R., Ramnath, R., 2019b. Accident risk prediction based on heterogeneous sparse data: new dataset and insights. In: Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 33–42.
Mosler, K., Polyakova, Y., 2012. General notions of depth for functional data. arXiv preprint. arXiv:1208.1981.
Narisetty, N.N., Nair, V.N., 2016. Extremal depth for functional data and applications. J. Am. Stat. Assoc. 111, 1705–1714.
Nieto-Reyes, A., 2011. On the properties of functional depth. In: Recent Advances in Functional Data Analysis and Related Topics. Springer, pp. 239–244.
Oja, H., 1983. Descriptive statistics for multivariate distributions. Stat. Probab. Lett. 1, 327–332.
Pawlowsky-Glahn, V., Egozcue, J.J., Tolosana Delgado, R., 2007. Lecture Notes on Compositional Data Analysis.
Qi, K., Chen, Y., Wu, W., 2021. Dirichlet depths for point process. Electron. J. Stat. 15, 3574–3610.
Shchur, O., Turkmen, A.C., Januschowski, T., Gasthaus, J., Günnemann, S., 2021. Detecting anomalous event sequences with temporal point processes. Adv. Neural Inf. Process. Syst. 34, 13419–13431.
Silverman, B.W., 2018. Density Estimation for Statistics and Data Analysis. Routledge.
Srivastava, A., Wu, W., Kurtek, S., Klassen, E., Marron, J.S., 2011. Registration of functional data using Fisher-Rao metric. arXiv preprint. arXiv:1103.3817.
Stoyan, D., Kendall, W.S., Chiu, S.N., Mecke, J., 2013. Stochastic Geometry and Its Applications. John Wiley & Sons.
Tukey, J.W., 1975. Mathematics and the picturing of data. In: Proceedings of the International Congress of Mathematicians, vol. 2. Vancouver, 1975, pp. 523–531.
Wójcik, D.K., Mochol, G., Jakuczun, W., Wypych, M., Waleszczyk, W.J., 2009. Direct estimation of inhomogeneous Markov interval models of spike trains. Neural Comput. 21, 2105–2113.
Wu, W., Srivastava, A., 2011. An information-geometric framework for statistical inferences in the neural spike train space. J. Comput. Neurosci. 31, 725–748.
Zhu, S., Yuchi, H.S., Xie, Y., 2020. Adversarial anomaly detection for marked spatio-temporal streaming data. In: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 8921–8925.
Zuo, Y., 2000. A note on finite sample breakdown points of projection based multivariate location and scatter statistics. Metrika 51, 259–265.
Zuo, Y., Serfling, R., 2000. General notions of statistical depth function. Ann. Stat., 461–482.