Second order preserving point process permutations

George Mohler[*] Jorge Mateu[†]

Abstract

While random permutations of point processes are useful for generating counterfactuals in bivariate interaction tests, such permutations require that the underlying intensity be separable. In many real-world datasets where clustering or inhibition is present, such an assumption does not hold. Here we introduce a simple combinatorial optimization algorithm that generates second order preserving (SOP) point process permutations, e.g. permutations of the times of events such that the L function of the permuted process matches the L function of the data. We apply the algorithm to synthetic data generated by a self-exciting Hawkes process and a self-avoiding point process, along with data from Los Angeles on earthquakes and arsons and data from Indianapolis on law enforcement drug seizures and overdoses. In all cases we are able to generate a diverse sample of permuted point processes where the distribution of the L functions closely matches that of the data. We then show how SOP point process permutations can be used in two applications: 1) bivariate Knox tests and 2) data augmentation to improve deep learning based space-time forecasts.

Keywords: Bivariate point processes, Data augmentation, Independence, Knox test, Permutations, Second order characteristics, Self-exciting point processes

[*]Department of Computer Science, Boston College
[†]Department of Mathematics, University Jaume I

1 Introduction

In this paper we are motivated by the problem of testing for interactions between two spatial-temporal point processes. For example, in [10] a two sample randomization test was introduced to test for interactions between underground nuclear tests and subsequent earthquakes. Similar methods have been applied to test for interactions between firearm arrests and shootings [20], homicide and other violent or property crimes [5] and law enforcement drug seizures and overdose [13]. In a more general multivariate fashion, the interest is in detecting and modeling cross-interactions amongst any two marginal processes, as shown in [8] and [18]. Indeed, the literature on modeling multivariate spatial point patterns is mainly restricted to the bivariate case (see [4], [7], [6], [14]). In this context, interactions have been modeled by using Gibbs point processes and Cox processes.

In a two sample randomization test, the event times of a point process are permuted, while the spatial coordinates of the process are fixed. This "permuted point process" provides a counterfactual that can be compared to the second point process for which an interaction is being assessed. However, the permuted point process may not have the same statistics as the original data. For example, in Figure 1 we plot a realization of a self-exciting Hawkes process along with a permutation of the Hawkes process. Whereas the Hawkes process exhibits significant space-time clustering, the permuted process only exhibits spatial clustering. Thus two sample randomization tests typically require the assumption that the point process is stationary Poisson or that the intensity of the point process is separable [3, 1, 16].

Our goal here is to generate permutations of a point process that preserve second order statistics of the process. In Section 2, we introduce a simple combinatorial optimization algorithm that, starting from a random permutation, iteratively swaps two random event times to improve the agreement of the second order statistics (as measured by the L function) between the permuted point process and the data. We then apply the method in Section 3 to synthetic data generated by a self-exciting Hawkes process and a self-avoiding point process. We also use the method to assess interactions between a) Los Angeles earthquakes and arson and b) Indianapolis law enforcement drug seizures and overdose. Finally, we use second order preserving point process permutations to augment training data and improve a CNN-LSTM based short-term forecast. We discuss some open questions and directions for future research in Section 4.
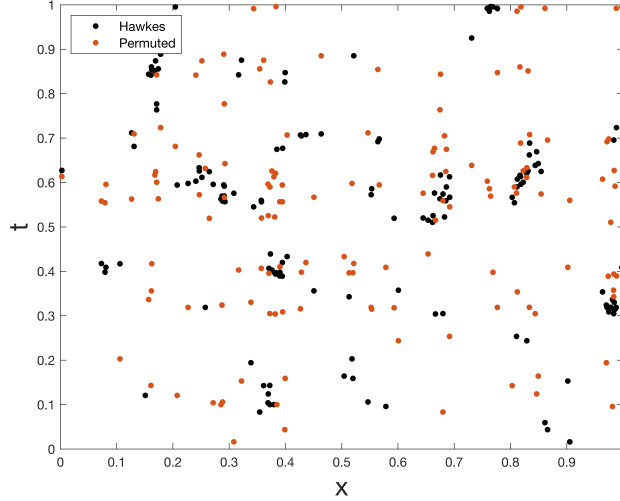
Figure 1: Simulated Hawkes process in black (first spatial coordinate vs. time) along with the same Hawkes process with times randomly permuted in red.

2 Methodology

Let $(\mathbf{x}_i, t_i)_{i=1}^N$ denote the spatial coordinates and event times of a point process and $\widetilde{t}_i$ be a random permutation of the event times. Two common second order statistics of a spatial point process are Ripley's K function [15] and the related L function. We consider space-time extensions of the K and L functions, where we first define $\mathbf{z}_i$ to be a 3-dimensional vector consisting of the two spatial coordinates and one time coordinate of $(\mathbf{x}_i, t_i)$ (and each coordinate is rescaled by min-max scaling to be in $[0, 1]$). The K function is then estimated by,

$$K(r) = \frac{1}{N^2} \sum_{i,j} 1\{\|\mathbf{z}_i - \mathbf{z}_j\| < r\} \tag{1}$$

and the L function is given by $L(r) = [K(r)]^{1/2}$. The K and L functions measure the prevalence of events within a space-time radius r of each event from the process and can be compared to the K and L functions of a Poisson process to determine clustering or inhibition.

Here we propose an algorithm to generate permutations such that the L function, $L(r)$, of the data better matches the L function, $\widetilde{L}(r)$, of the

3

permuted point process. We first note that we do not want the L function of each permutation to exactly match that of the data. If one were to simulate multiples realizations of a point process, there would be variation in each sample that would lead to a distribution of L functions. Thus our goal is to generate multiple permuted point processes such that the distribution of the L functions closely matches the distribution of the L function of the data.

The algorithm, which is summarized in Algorithm 1, proceeds in two stages. In the first stage, M independent random permutations $\tilde{\mathbf{z}}^k = (\mathbf{x}_i, \tilde{t}_i^k)$, $k = 1, ..., M$ of the data are generated and the L function $L^k(r)$ is computed for each permutation. Next, the mean over the M L functions is calculated,

$$\mu(r) = \frac{1}{M} \sum_{k=1}^{M} L^k(r), \tag{2}$$

along with the error,

$$\epsilon^k(r) = L^k(r) - \mu(r). \tag{3}$$

In the second stage, we generate M second order preserving (SOP) permutations. For each k, we initialize the permutation with the random permutation $\tilde{\mathbf{z}}^k$ from stage one. We then iteratively swap two random times to generate a proposal permutation $\tilde{\mathbf{q}}^k$. The L function, $L^{prop}(r)$, of this proposal permutation is then computed, along with the proposal error,

$$error_{prop} = \left(\int |L^{prop}(r) - L^{data}(r) - \epsilon^k(r)|^2 dr \right)^{1/2}. \tag{4}$$

Here we want the L function of the permutation to be close to the L function of the data plus the L function error of random permutation k, so that the variation of the SOP permutations around the data L function will be similar to the variation of the random permutation L functions around their mean. We accept proposal permutations if they reduce the error and we terminate the iteration when the error is below a tolerance parameter γ. We note that the integral in Equation 4 can be approximated using numerical integration, for example using the trapezoidal rule (we calculate the L function at 100 equally spaced points and approximate the integral using quadrature at those same points).

2.1 Bivariate Knox test

The motivating application of our second order preserving permutations is a two-sample Knox test [10] for interactions between two point processes.

Algorithm 1 SOP point process permutation pseudo code

Input: Data $(\mathbf{x}_i, t_i)_{i=1}^{N}$, number of permutations M, tolerance γ, L function of data $L^{data}(r)$

Output: SOP permutations $\tilde{\mathbf{z}}^k$, for $k = 1, ...M$

1: **for** k=1:M **do**

      $\tilde{\mathbf{z}}^k = (\mathbf{x}_i, \tilde{t}_i^k)$ \\ permute event times

      $L^k(r) = \left(\frac{1}{N^2} \sum_{i,j} 1\{\|\tilde{\mathbf{z}}_i^k - \tilde{\mathbf{z}}_j^k\| < r\}\right)^{1/2}$ \\ compute L function

2: **end for**

3: $\mu(r) = \frac{1}{M} \sum_{k=1}^{M} L^k(r)$ \\ mean L function of permuted processes

4: **for** k=1:M **do**

      $\epsilon^k(r) = L^k(r) - \mu(r)$ \\ L function error of permuted process k

5: **end for**

6: **for** k=1:M **do**

7:     $(\mathbf{x}_i, s_i) \leftarrow \tilde{\mathbf{z}}^k$ \\ initialize permutation

8:     $error = \gamma + 1$

9:     **while** $error > \gamma$ **do**

10:         $j_1, j_2 = random(1, N)$ \\ pick two random integers

11:         $\widetilde{s} = s$

12:         $\widetilde{s}_{j_1} = s_{j_2}$ \\ swap event times

13:         $\widetilde{s}_{j_2} = s_{j_1}$

14:         $\tilde{\mathbf{q}}^k \leftarrow (\mathbf{x}_i, \widetilde{s}_i)$ \\ new proposal permutation

15:         $L^{prop}(r) = \left(\frac{1}{N^2} \sum_{i,j} 1\{\|\tilde{\mathbf{q}}_i^k - \tilde{\mathbf{q}}_j^k\| < r\}\right)^{1/2}$ \\ compute L function

16:         $error_{prop} = (\int |L^{prop}(r) - L^{data}(r) - \epsilon^k(r)|^2 dr)^{1/2}$ \\ calculate error

17:         **if** $error_{prop} < error$ **then**

18:           $error = error_{prop}$ \\ update best error

19:           $(\mathbf{x}_i, s_i) \leftarrow \tilde{\mathbf{q}}^k$ \\ update optimal permutation

20:         **end if**

21:     **end while**

22:     $\tilde{\mathbf{z}}^k \leftarrow (\mathbf{x}_i, s_i)$ \\ output optimal permutation k

23: **end for**

In particular, given a time cutoff τ and spatial distance cutoff δ, the Knox statistic [11] $\kappa(\tau, \delta)$ is given by,

$$\kappa(\tau, \delta) = \sum_{i,j} 1\{\|\mathbf{x^a}_i - \mathbf{x}_j^b\| \le \delta, |t_i^a - t_j^b| < \tau\}. \tag{5}$$

where the two point processes are $(\mathbf{x}_i^a, t_i^a)$ and $(\mathbf{x}_j^b, t_j^b)$. The Knox statistic counts the number of events of type b within a radius δ and time window τ of events of type a.

To determine excess clustering or inhibition, the Knox statistic can be compared to a null distribution where the two processes are independent. The null distribution of the Knox statistic is computed through multiple realizations of,

$$\widetilde{\kappa}(\tau, \delta) = \sum_{i,j} 1\{\|\mathbf{x^a}_i - \mathbf{x}_j^b\| \le \delta, |t_i^a - \widetilde{t}_j^b| < \tau\}, \tag{6}$$

where $\widetilde{t}_j^b$ are random permutations of the event times of process b. In the next section we will explore the benefits of using SOP permutations rather than random permutations to generate $\widetilde{t}_j^b$.

3 Experiments

In this section we illustrate the effectiveness of Algorithm 1 at generating second order preserving permutations using synthetic and real data. We then compare the results of bivariate Knox tests that utilize SOP permutations versus random permutations. Finally, we show how second order preserving permutations can be used to improve a CNN-LSTM based space-time forecast through data augmentation.

3.1 Synthetic data experiment with a Hawkes process

In the first experiment we generate data from a self-exciting Hawkes point process with intensity,

$$\lambda(x, t) = \mu + \sum_{t > t_i} \theta f(t - t_i; \omega) g(x - x_i; \sigma). \tag{7}$$

We simulate the process on the unit cube in space-time with background rate $\mu = 40$, reproduction number $\theta = 0.75$, exponential kernel f in time

with parameter $\omega = 100$ (mean 0.01) and Gaussian kernel g in space (2d) with standard deviation $\sigma = 0.01$. We estimate the L function on $r \in [0, 0.3]$ at 100 discrete (evenly spaced) points and use a tolerance of $\gamma = .01$ for Algorithm 1. The tolerance was selected by visually inspecting the L functions corresponding to different choices of γ (see Figure 2).
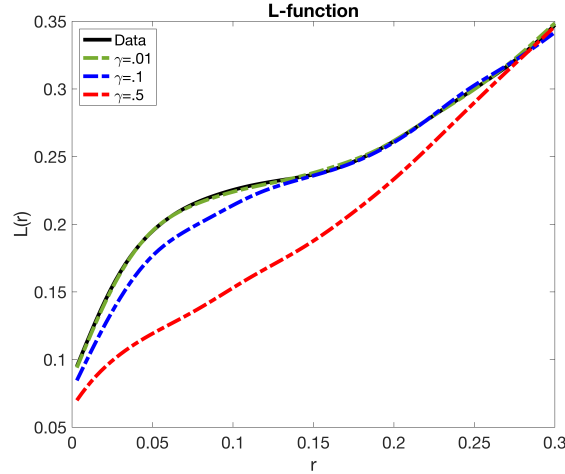

Figure 2: L function of data vs. L functions of SOP permutations for different choices of γ.

In Figure 3, we plot a realization of the Hawkes process (first spatial coordinate versus time) along with an example random permutation. We then plot three example SOP (L function preserving) permutations and the corresponding L function. We see that Algorithm 1 is able to produce diverse point patterns that match the second order statistics of the permutations to the original data.

Next we apply a bivariate Knox test to two independent Hawkes process realizations generated with the above parameters. For the Knox test we use parameters $\delta = 0.1$ and $\tau = 0.1$ for the Knox statistic $\kappa(\tau, \delta)$. In Figure 4, we plot the null distribution of the Knox statistic under both random permutations and SOP permutations for $M = 200$ permutations. We note that the random bivariate Knox test rejects the null hypothesis that the two point patterns are independent, whereas the p-value for the SOP based Knox test is $p = 0.1$. On the bottom of Figure 4 we also plot the L functions for the 200 permutations. Here we see that the random permutations are much
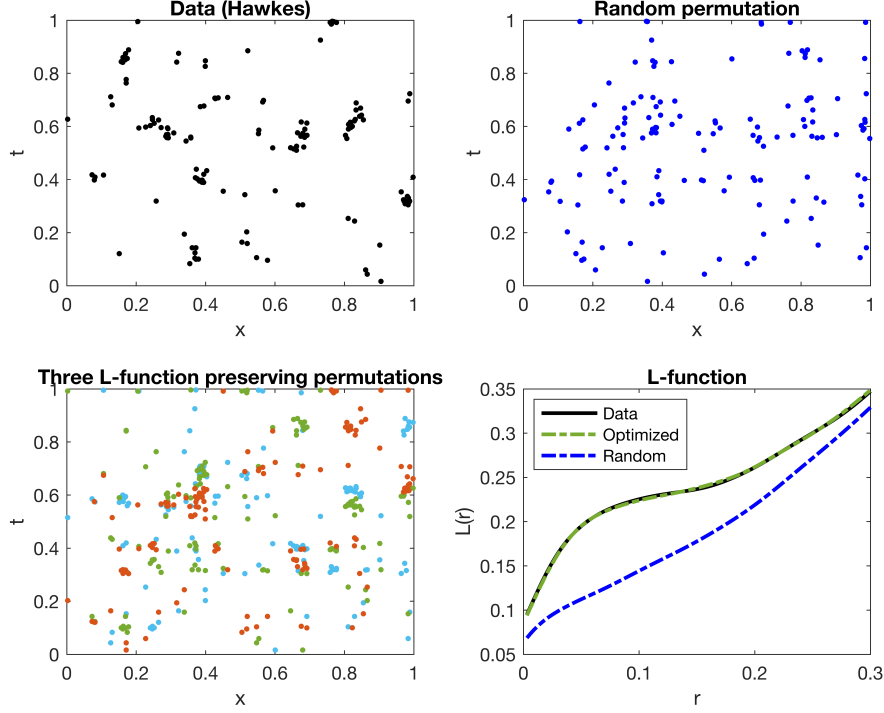
Figure 3: Top Left: Realization of the Hawkes process (first spatial coordinate vs. time). See text for the parameters used. Top right: example random permutation. Lower left: Three example SOP permutations of the Hawkes process. Lower Right: L function of data vs. L functions of SOP and random permutations.

less clustered than the data, whereas the SOP permutations match the L function of the data (with similar variation to the random permutations, by design).

3.2 Synthetic data experiment with a regular self-avoiding point process

In the next experiment we generate data from a self-avoiding "regular" point process. We iteratively simulate N points in the unit cube using rejection. The fist point is generated at random, then the next point is generated at
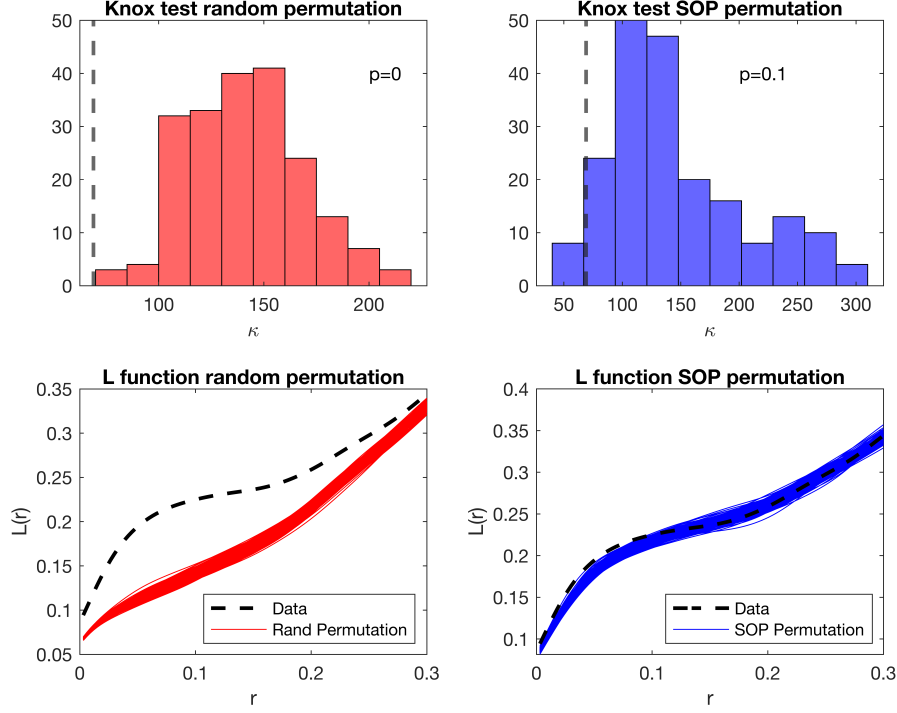
Figure 4: Two independent Hawkes processes experiment. Top: Null distribution of the Knox statistic under both random permutations (left) and SOP permutations (right) for $M = 200$ permutations. Data Knox statistic represented by vertical dashed line. Bottom: L functions for the 200 random (right) and SOP (left) permutations.

random subject to being at least a distance c (space-time Euclidean distance) from all previous points (otherwise the point is rejected and a new point is added). We let $N = 100$, $c = 0.2$, and estimate the L function on $r \in [0, 0.3]$ at 100 discrete (evenly spaced) points and use a tolerance of $\gamma = .01$ for Algorithm 1.

In Figure 5, we plot a realization of the regular process (first spatial coordinate vs. time) along with an example random permutation. We then plot three example SOP (L function preserving) permutations and the corresponding L function. We see that Algorithm 1 is able to produce diverse regular point patterns that match the second order statistics of the permu-
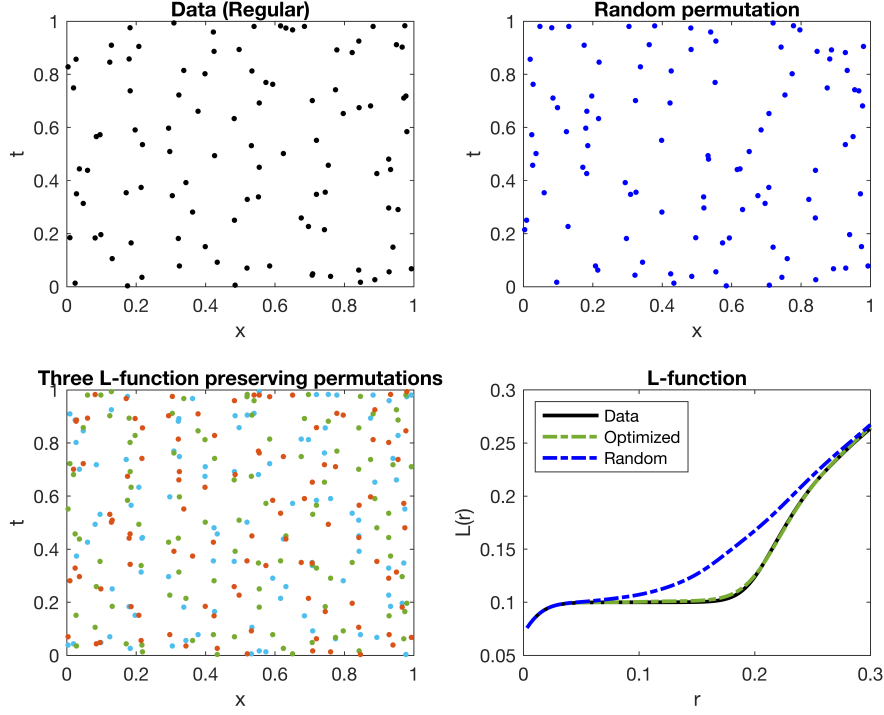
tations to the original data.

Figure 5: Top Left: Realization of the regular process (first spatial coordinate vs. time). See text for the parameters. Top right: example random permutation. Lower left: Three example SOP permutations of the regular process. Lower Right: L function of data vs. L functions of SOP and random permutations.

Next we apply a bivariate Knox test to two independent regular process realizations generated with the above parameters. For the Knox test we use parameters $\delta = 0.1$ and $\tau = 0.1$ for the Knox statistic $\kappa(\tau, \delta)$. In Figure 6, we plot the null distribution of the Knox statistic under both random permutations and SOP permutations for $M = 200$ permutations. We note that in this case both the random and SOP bivariate Knox tests fail to reject the null hypothesis that the two point patterns are independent (which is the correct result). However, in the bottom of Figure 6 we see that the random permutations are much more clustered than the data, whereas the

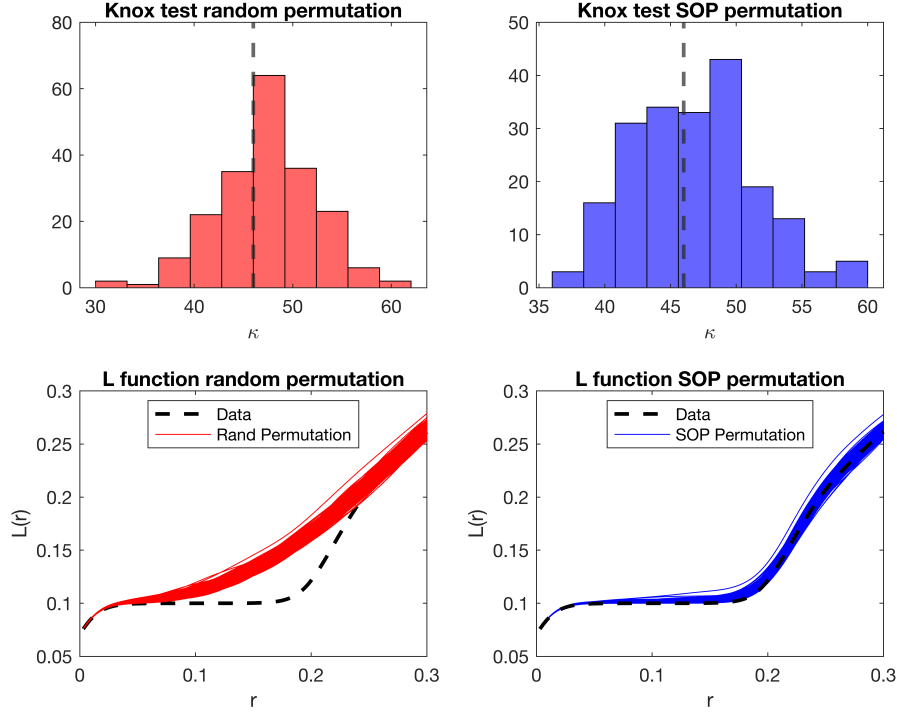SOP permutations better match the L function of the data.

Figure 6: Two independent regular point processes experiment. Top: Null distribution of the Knox statistic under both random permutations (left) and SOP permutations (right) for $M = 200$ permutations. Data Knox statistic represented by vertical dashed line. Bottom: L functions for the 200 random (right) and SOP (left) permutations.

3.3 Association between earthquakes and arson in Los Angeles

Next we apply a bivariate Knox test to earthquake and arson data from Los Angeles during 2020-2021. The earthquake magnitudes range from 1.5 (query cutoff) to 4.28 being the largest. The arson events are based on Los Angeles crime reports from the same time period. Given the small magnitude of the earthquakes, we do not expect there to be an association between these

two types of events. We plot the distribution of events in space and time in Figure 7.
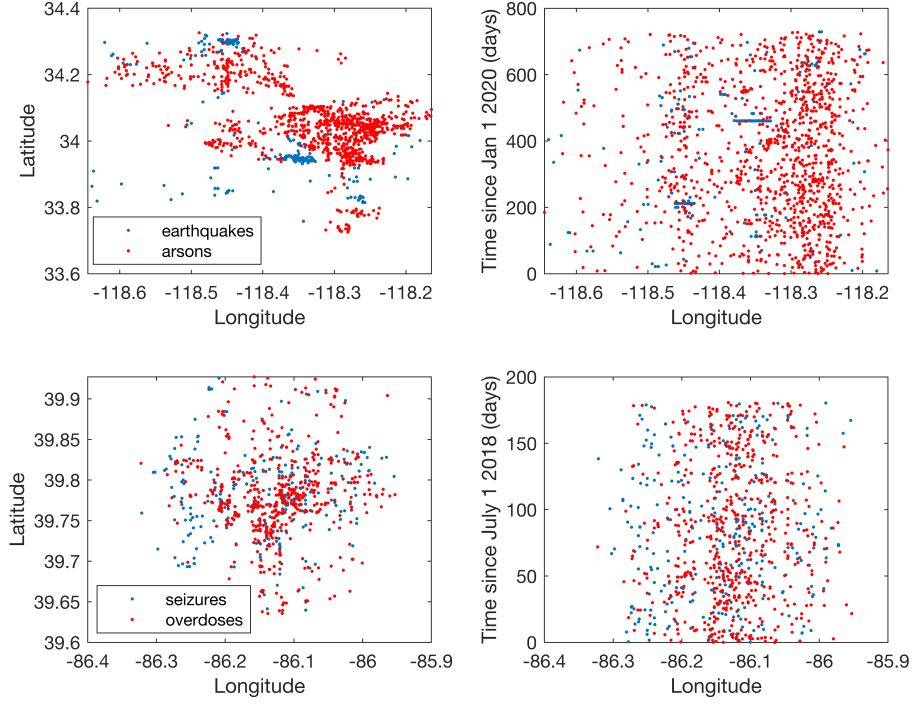

Figure 7: Top: distribution of earthquake and arson events in Los Angeles during 2020 and 2021. Bottom: distribution of drug seizure and overdose events in Indianapolis during the second half of 2018.

For the Knox test we use parameters $\delta = 1000$m and $\tau = 30$ days for the Knox statistic $\kappa(\tau, \delta)$. In Figure 8, we see that the random permutations are less clustered than the data, whereas the L functions of the SOP permutations better match the L function of the data. In Figure 8, we also plot the null distribution of the Knox statistic under both random permutations and SOP permutations for $M = 400$ permutations. Whereas the random permutation test would reject the null hypothesis of independence at the $p = 0.01$ (two-sided) level, the SOP based Knox test fails to reject the null hypothesis at the $p = 0.05$ level.

If we look again at Figure 7, we see that earthquakes are clustered along

12

fault lines, which are in different areas from housing (where arsons are clustered). In time, earthquake aftershocks are highly clustered in short windows (for example a few hours). Thus it appears that arsons "avoid" earthquakes, hence the rejection of the null hypothesis by the standard two-sample Knox test. However, the standard test assumes that the processes themselves are not clustered, and clearly earthquakes are. The SOP based Knox test accounts for randomly occurring avoidance among two independent cluster processes, and the test produces a more conservative p-value.
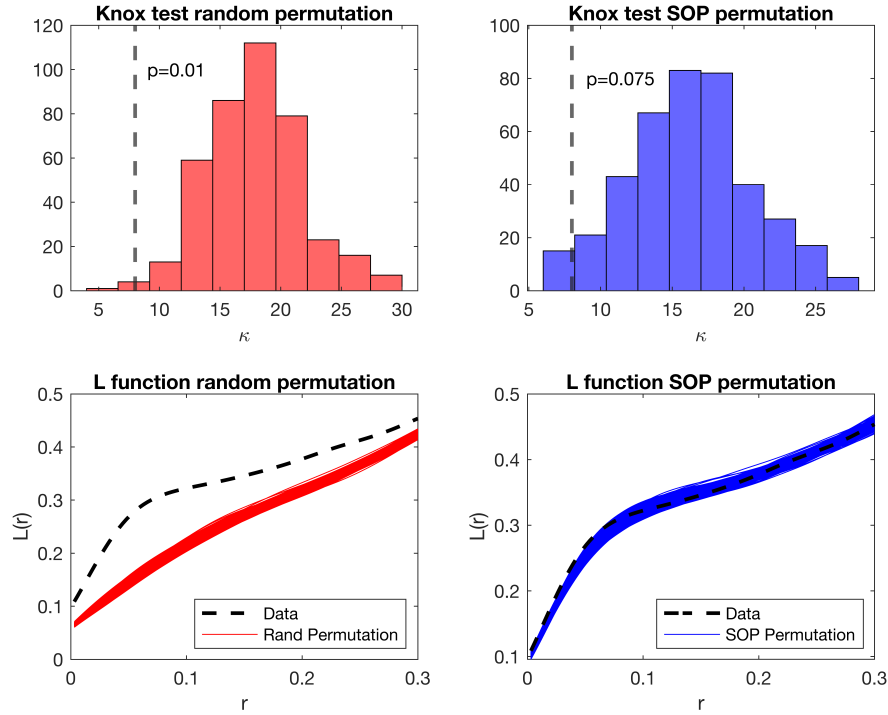

Figure 8: Arson versus Earthquakes. Top: Null distribution of the Knox statistic under both random permutations (left) and SOP permutations (right) for $M = 400$ permutations. Data Knox statistic represented by vertical dashed line. Bottom: L functions for the 400 random (right) and SOP (left) permutations.

3.4 Association between drug seizures and overdose in Indianapolis

In our last experiment, we assess interactions between law enforcement drug seizure events and overdoses where naloxone was administered by Indianapolis Emergency Medical Services. Data consist of latitude, longitude, date and time of the incident and come from Indianapolis, Indiana from July 1, 2018 to December 31, 2018. A space-time positive association between law enforcement drug seizures and overdose was observed in [13]. One hypothesis is that, when law enforcement officers make an arrest for drug dealing and seize drugs, users may need to seek out alternative sources and may be at higher risk of overdose in the near future.

For the Knox test we use parameters $\delta = 250$m and $\tau = 21$ days for the Knox statistic $\kappa(\tau, \delta)$ (similar to the values used in [13]). In Figure 9 we see similar agreement between the L functions of random permutations, SOP permutations and the L function of the data. One possible explanation is that the intensity of law enforcement seizures are separable, and thus random permutations generate similar second order statistics to SOP permutations. We note that the L functions are tightly clustered around the mean, due to larger number of events in the dataset compared to the previous three examples. In Figure 9, we plot the null distribution of the Knox statistic under both random permutations and SOP permutations for $M = 400$ permutations. Both tests reject the null independence hypothesis, as none of the $M = 400$ permutations yield Knox statistics as extreme as the data.

3.5 Data augmentation for CNN-LSTM 1-day ahead point process forecast

A number of neural network based models for space-time point processes have been introduced recently [9] [12] [2] [19]. These models can improve accuracy of forecasts over simpler parametric models, but typically require larger datasets due to the increased number of parameters. Here we show how SOP permutations can be used to improve model performance on held out data through augmentation of training data.

We first generate data from a space-time Hawkes process defined on $[0, 1] \times [0, 1] \times [0, 730]$ with background rate $\mu = 40/730$, reproduction number $\theta = 0.75$, exponential kernel f in time with parameter $\omega = 100/730$ and Gaussian kernel g in space (2d) with standard deviation $\sigma = 0.01$. Next we
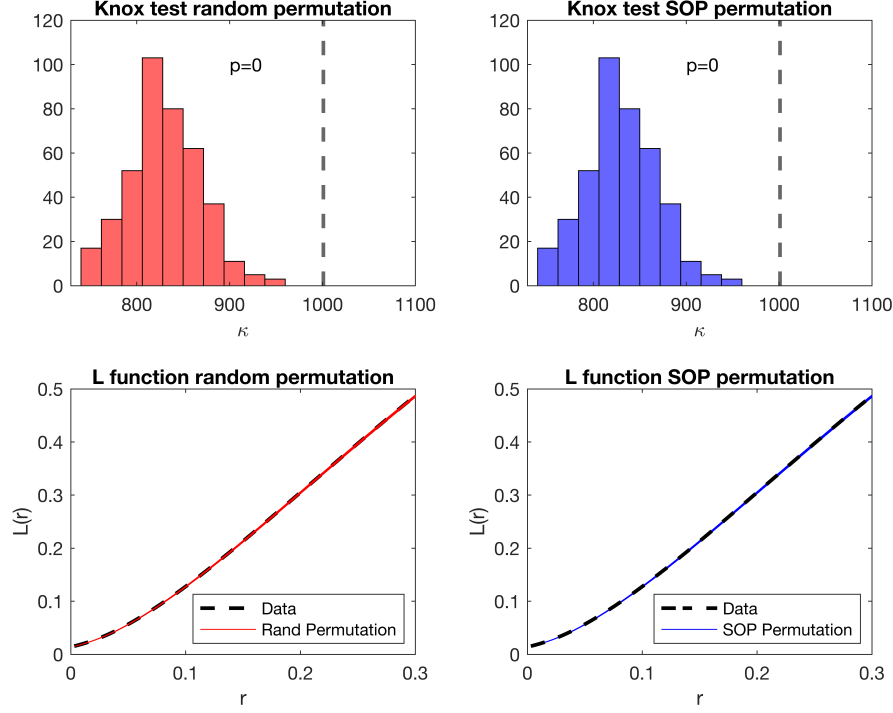
Figure 9: Law enforcement drug seizures versus overdose. Top: Null distribution of the Knox statistic under both random permutations (left) and SOP permutations (right) for $M = 400$ permutations. Data Knox statistic represented by vertical dashed line. Bottom: L functions for the 400 random (right) and SOP (left) permutations.

discretize space into 25×25 grid cells and time into 1 "day" intervals from 0 to 730. We then use a sliding window and create $14 \times 25 \times 25$ features consisting of a binary indicator in each space-time cell for whether at least one event occurred ($y = 1$) or no events occurred ($y = 0$). We use this feature as input to a CNN-LSTM [17] to predict whether an event will occur or not in each grid cell in the following day. The CNN-LSTM consists of three ConvL-STM2D layers (implemented in Keras) followed by batch normalization. The last layer is a Conv3D layer with a sigmoid activation. The model is trained using Adam optimization with a binary crossentropy loss. The data is split into the first 486 days used for training (after waiting 14 days to create the

first set of features) and the last 216 days used for testing.

First we train the model with no data augmentation. In Figure 10, we plot the AUC of the model on held out data over 20 epochs. Recall that AUC stands for the Area Under the ROC Curve, and provides an aggregate measure of performance across all possible classification thresholds. Next we use SOP permutations to increase the data set size by a factor of 10. We again use 20 epochs, so in this case on the 0th and 10th epochs the actual data is used to train the model, whereas on the other epochs SOP permuted data is used. We see in Figure 8 that the sop-augmented training yields an improved AUC value on the held out data.
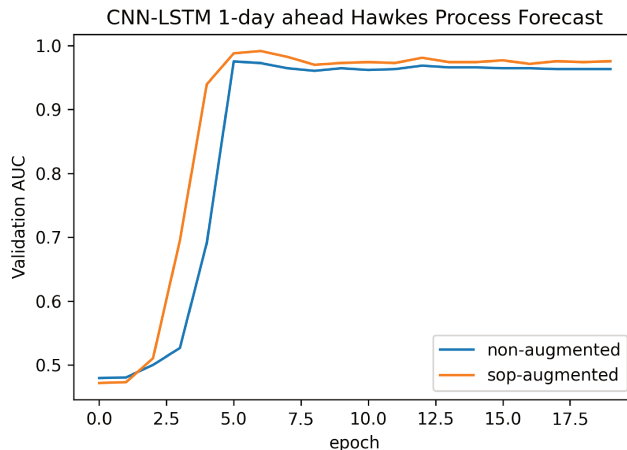

Figure 10: AUC on held-out data versus training epochs for a CNN-LSTM forecast using sop-augmented training and non-augmented training.

4   Discussion

In this paper we introduced an algorithm for generating point process permutations that preserve second order statistics of the process. We showed that the SOP permutations may be useful for bivariate Knox tests when the underlying process intensities are not separable. In two of our experiments, standard bivariate Knox tests suffered from false discovery and rejected the independence null hypothesis, whereas the more conservative SOP based

Knox test did not. In the other two cases we found that random and SOP based Knox tests produced similar results.

Another potential application for SOP point process permutations is data augmentation in deep learning models. One challenge in point process research is that a dataset typically consists of a single realization. The algorithm we have introduced here could be used to simulate other realizations, and our results indicate that such realizations may be able to improve accuracy of deep learning based space-time models.

There are several limitations of the present article. First, we have no guarantees that Algorithm 1 will converge for a given tolerance. In our experiments the algorithm was able to find permutations (other than the identity) that satisfied the error objective, but for certain point processes there may not be a solution. Second, we have not analyzed the theoretical properties of SOP based bivariate Knox tests. In our four examples we observed improved type 1 errors without loss of power, but more research is needed to better understand the statistical properties of SOP permutations and associated interaction tests. At present, our recommendation would be to use the SOP based Knox test in tandem with the standard random Knox test. If the results fail to agree, then one should question whether the assumptions of the standard Knox test hold. Future research should also explore data augmentation applications of SOP permutations, as well as more sophisticated combinatorial optimization techniques to speed up convergence.

5   Acknowledgements

References

[1] Rajinder S Bhopal, Peter Diggle, and Barry Rowlingson. Pinpointing clusters of apparently sporadic cases of legionnaires' disease. *British Medical Journal*, 304(6833):1022–1027, 1992.

[2] Ricky TQ Chen, Brandon Amos, and Maximilian Nickel. Neural spatio-

temporal point processes. In *International Conference on Learning Representations*, 2020.

[3] Peter J Diggle. *Statistical analysis of spatial and spatio-temporal point patterns*. CRC press, 2013.

[4] Peter J Diggle and Robin K Milne. Bivariate cox processes: some models for bivariate spatial point patterns. *Journal of the Royal Statistical Society: Series B (Methodological)*, 45(1):11–21, 1983.

[5] Seth R Flaxman, Daniel B Neill, and Alex Smola. Correlates of homicide: new space/time interaction tests for spatiotemporal point processes, 2013.

[6] Pavel Grabarnik and Aila Särkkä. Modelling the spatial structure of forest stands by multivariate point processes with hierarchical interactions. *Ecological Modelling*, 220(9-10):1232–1240, 2009.

[7] RD Harkness and Valerie Isham. A bivariate spatial point pattern of ants' nests. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 32(3):293–303, 1983.

[8] Abdollah Jalilian, Yongtao Guan, Jorge Mateu, and Rasmus Waagepetersen. Multivariate product-shot-noise cox point process models. *Biometrics*, 71(4):1022–1033, 2015.

[9] Abdollah Jalilian and Jorge Mateu. Assessing similarities between spatial point patterns with a siamese neural network discriminant model. *Advances in Data Analysis and Classification*, pages 1–22, 2022.

[10] Melville R Klauber. Two-sample randomization tests for space-time clustering. *Biometrics*, pages 129–142, 1971.

[11] Ernest G Knox and Maurice S Bartlett. The detection of space-time interactions. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 13(1):25–30, 1964.

[12] Jorge Mateu and Abdollah Jalilian. Spatial point processes and neural networks: A convenient couple. *Spatial Statistics*, page 100644, 2022.

[13] G Mohler, S Mishra, B Ray, L Magee, P Huynh, M Canada, D O'Donnell, and S Flaxman. A modified two-process knox test for investigating the relationship between law enforcement opioid seizures and overdoses. *Proceedings of the Royal Society A*, 477(2250):20210195, 2021.

[14] Nicolas Picard, Avner Bar-Hen, Frédéric Mortier, and Joel Chadoeuf. The multi-scale marked area-interaction point process: a model for the spatial pattern of trees. *Scandinavian journal of statistics*, 36(1):23–41, 2009.

[15] Brian D Ripley. The second-order analysis of stationary point processes. *Journal of applied probability*, 13(2):255–266, 1976.

[16] Carl P Schmertmann, Renato M Assunção, and Joseph E Potter. Knox meets cox: Adapting epidemiological space-time statistics to demographic studies. *Demography*, 47(3):629–650, 2010.

[17] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, 28, 2015.

[18] Rasmus Waagepetersen, Yongtao Guan, Abdollah Jalilian, and Jorge Mateu. Analysis of multispecies point patterns by using multivariate log-gaussian cox processes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 65(1):77–96, 2016.

[19] Bao Wang, Duo Zhang, Duanhao Zhang, P Brantingham, and Andrea L Bertozzi. Deep learning for real time crime forecasting. *IEICE Proceeding Series*, 29(A3L-E-2-4):330–333, 2017.

[20] Brian R Wyant, Ralph B Taylor, Jerry H Ratcliffe, and Jennifer Wood. Deterrence, firearm arrests, and subsequent shootings: A micro-level spatio-temporal analysis. *Justice Quarterly*, 29(4):524–545, 2012.