Performance Analysis of Hypothesis Testing for Sparse Pairwise Interaction Point Processes

John A. Gubner, *Member, IEEE*, Wei-Bin Chang, and Majeed M. Hayat, *Senior Member, IEEE*

Abstract—The sufficient statistic for performing the likelihood ratio test for pairwise interaction point processes is well-known; however, the evaluation of its performance is a very difficult problem. In this paper it is shown that the distribution of the sufficient statistic can be approximated by the distribution of a Poisson-driven shot-noise random variable, which can be readily computed.

Index Terms—Gibbs point process, pairwise interaction point process, Poisson approximation, shot noise.

I. INTRODUCTION

SPATIAL point processes arise in many applications such as cosmology, ecology, forestry, seismology, and tomography, [4, pp. 578–579], [17, pp. 115–120]. A common task in these areas is that of testing for complete spatial randomness; i.e., whether the observed data is independently and uniformly distributed. When one has a model for the alternative to complete spatial randomness, a standard hypothesis testing problem results. For example, an important class of alternatives is provided by a type of Gibbs point process known as a pairwise interaction point process [4, pp. 669–689], [15, Ch. 4]. These processes are completely characterized by a univariate interaction function. The problem of estimating this function has been addressed in several papers, e.g., [3], [5]–[7], and [11]–[14]. In this paper, we assume that the interaction function is known. One can then show that the sufficient statistic for performing the likelihood ratio test is given by a simple expression. However, since the distribution of this statistic is not readily available, determining the performance of the test is a very challenging problem. In this paper we develop approximations of the desired distribution.

The paper is organized as follows. In Section II, the definition of a pairwise interaction point process is given, and two examples are presented. In Section III, the hypothesis testing problem is formulated, and in Section IV, the distribution of the likelihood ratio statistic is analyzed. To approximate the distribution, numerical methods are presented and illustrated with two examples. The conclusion is in Section V. Appendix A states and proves a limit theorem that allows us to approximate the distribution of the likelihood ratio statistic by the distribution of a Poisson-driven shot-noise random variable. Appendices B and C provide derivations and extensions of some of the results stated in Section IV-B. Finally, Appendix D analyzes certain sparseness hypotheses used in the limit theorem of Appendix A.

II. MATHEMATICAL MODEL

Let D be a bounded subset of the plane[1] $\mathbb{R}^2$ equipped with the Euclidean norm $\| \cdot \|$. In our numerical examples, we take D to be the 10-by-10 square with lower left-hand corner at the origin. A D^n-valued random vector $X = (X_1, \cdots, X_n)$ is called a *pairwise interaction point process* [15] if it has a density of the form

$$f(x) = \frac{1}{Z} \prod_{1 \leq i < j \leq n} \varphi(\|x_i - x_j\|), \qquad x = (x_1, \cdots, x_n) \in D^n$$

where Z is the normalizing constant

$$Z := \int_{D^n} \prod_{1 \leq i < j \leq n} \varphi(\|x_i - x_j\|) \, dx$$

and the function $\varphi \colon [0, \infty) \to [0, \infty)$ is called the *interaction function*. The idea is that if $\varphi(r) < 1$ (resp., > 1), then realizations in which many point pairs have $\|x_i - x_j\| \approx r$ will have low (resp., high) probability. In our applications, $\varphi(r)$ is small for small r and 1 for large r; hence, realizations in which pairs of points are close are discouraged, while pairs of points that are far from each other are neither encouraged nor discouraged. In other words, points at close range tend to repel each other.

Example 1: If $\varphi(r) = 1$ for all r, then the X_i are independent and uniformly distributed. A realization of such a process with $n = 75$ points is shown in Fig. 1. For comparison with Example 2, point pairs that are within distance 0.5 are connected.

Example 2: Let $\varphi(r) = r/r_{\max}$ for $0 \leq r \leq r_{\max}$, and $\varphi(r) = 1$ for $r > r_{\max}$. An $n = 75$-point realization of this process with $r_{\max} = 0.5$ is shown in Fig. 2. Point pairs that are within distance 0.5 are connected. Note that there are fewer such pairs than in Fig. 1 due to the inhibition effects of the interaction function.

[1]The results here and in [3] should extend to any finite-dimensional Euclidean space with the necessary dimensional changes to the sparseness conditions in Appendix D; e.g., areas become volumes, etc. See [16] for specific sparseness formulas.
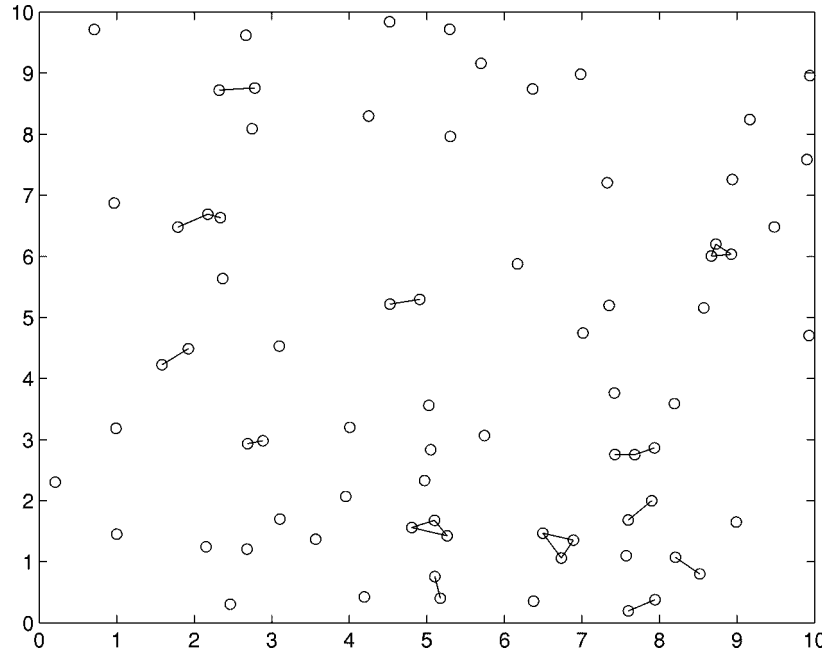
Fig. 1.   A point process with no interaction among points. Pairs closer than 0.5 are indicated by the lines joining them.

It is now convenient to introduce the *pair-potential function* $\psi(r) := -\ln \varphi(r)$. With this notation

$$\prod_{1 \leq i < j \leq n} \varphi(\|x_i - x_j\|) = \exp\left(-\sum_{1 \leq i < j \leq n} \psi(\|x_i - x_j\|)\right).$$

Next, let $S(r, x)$ denote the number of point pairs in x whose interpoint distance is less than or equal to r. Then $S(r, x)$ is a nondecreasing, piecewise-constant, integer-valued function of r, with jump discontinuities whenever $r = \|x_i - x_j\|$ for some x_i and x_j in x. With this notation, the sum above can be written as a Stieltjes integral

$$\sum_{1 \leq i < j \leq n} \psi(\|x_i - x_j\|) = \int_0^\infty \psi(r) S(dr, x).$$

III. THE HYPOTHESIS-TESTING PROBLEM

Let H_0 denote the hypothesis that X has pair potential $\psi_0 = -\ln \varphi_0$, and let H_1 denote the hypothesis that X has pair potential $\psi_1 = -\ln \varphi_1$. Then the likelihood ratio test for this problem is easily seen to be equivalent to

$$\ell(x) \underset{H_0}{\overset{H_1}{\gtrless}} y$$

where y is an adjustable threshold, and

$$\ell(x) := \int_0^\infty [\psi_0(r) - \psi_1(r)] S(dr, x).$$

The corresponding probabilities of detection and false alarm are

$$\wp_1(\ell(X) > y) \quad \text{and} \quad \wp_0(\ell(X) > y) \tag{1}$$

respectively, where $\wp_1$ indicates that X is an interaction point process with interaction function φ_1, and $\wp_0$ indicates that X is

an interaction point process with interaction function φ_0. Computing the probabilities in (1) is the focus of this paper.

IV. PERFORMANCE ANALYSIS

As noted in Section II, we usually take interaction functions to be one for large r, say $r \geq r_{\max} > 0$. Then $\psi_0(r) = \psi_1(r) = 0$ for $r \geq r_{\max}$. Hence, the integral for $\ell(x)$ above reduces to

$$\ell(x) = \int_0^{r_{\max}} [\psi_0(r) - \psi_1(r)] S(dr, x). \tag{2}$$

Without loss of generality (see the discussion at the end of Appendix A and also at the end of Appendix C), we take for the null hypothesis $\varphi_0(r) \equiv 1$ as in Example 1. Then $\psi_0(r) \equiv 0$, and we have

$$\ell(x) = -\int_0^{r_{\max}} \psi_1(r) S(dr, x). \tag{3}$$

Using this expression for ℓ, our goal is to approximate the probabilities in (1). Since the results below apply to both $\wp_1$ and $\wp_0$, we simplify the notation by writing $\wp(\ell(X) > y)$, where X has interaction φ, and it is understood that φ can be taken as either φ_1 or φ_0 as needed. However, no matter whether X has interaction function φ_1 or φ_0, ℓ in (3) is always defined using ψ_1.

Our first result is a limit theorem, whose precise statement and proof are given in Appendix A. Loosely speaking, the theorem says that if the number of points n is large, and if the region D is large enough that the points are "sparse," then

$$\wp(\ell(X) > y) \approx \wp(Y > y)$$

where the random variable Y is defined as follows. Let $\{S_\infty(r), 0 \leq r \leq r_{\max}\}$ be an inhomogeneous Poisson process with intensity $2\lambda r \varphi(r)$, where φ is the interaction function of X, and
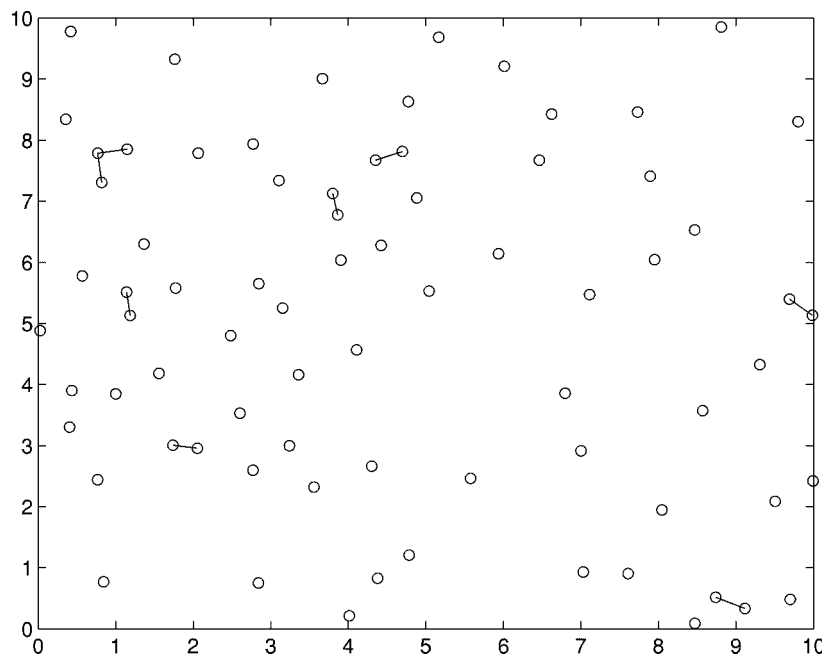
Fig. 2. A point process with inhibition of pairs of close points. Pairs closer than 0.5 are indicated by the lines joining them. Note that there are fewer such pairs than in Fig. 1.

the positive constant λ is determined by the sparseness conditions in Appendix D. Then

$$Y := -\int_0^{r_{\max}} \psi_1(r)\, S_\infty(dr). \tag{4}$$

We have thus reduced the computation of $\wp(\ell(X) > y)$ to that of $\wp(Y > y)$.

Our second result is that for many interaction functions φ_1, it is possible to calculate $\wp(Y > y)$ accurately and quickly. In the next subsection, we address the case in which φ_1 is piecewise-constant. In the following subsection, we address the case in which φ_1 is strictly increasing with continuous derivative on $[0, r_{\max}]$. In Appendix C, this is generalized to the case in which φ_1 is piecewise strictly monotonic with continuous derivative on each segment.

Remark 1: A referee has suggested that an alternative to our methods for computing $\wp(Y > y)$ is to use simulation. The reason being that it is easy to simulate Poisson processes and shot-noise random variables (as opposed to simulating interaction point processes to estimate $\wp(\ell(X) > y)$ empirically as we did for comparison in Figs. 3 and 4).

For use below, note that the characteristic function of Y is [10, Ch. 3]

$$\mathsf{E}\left[e^{j\omega Y}\right] = \exp\left(\int_0^{r_{\max}} [e^{-j\omega\psi_1(r)} - 1] 2\lambda r\, \varphi(r)\, dr\right). \tag{5}$$

Since Y is analogous to a single time sample of a shot-noise or filtered Poisson process [10, p. 25], we call Y a shot-noise random variable.

A. Piecewise-Constant φ_1

When the interaction function is piecewise-constant, $\wp(Y > y)$ can be approximated as follows. If φ_1 is piecewise-constant, then so is ψ_1. Suppose that ψ_1 takes value q_i on the interval $(r_i, r_{i+1}]$, where $0 = r_1 < \cdots < r_d = r_{\max}$. The random variable Y in (4) can then be written as

$$Y = -\sum_{i=1}^{d-1} q_i S_\infty((r_i, r_{i+1}])$$

where $S_\infty((r_i, r_{i+1}])$ is the number of points of the Poisson process S_∞ that occur in the interval $(r_i, r_{i+1}]$. We now assume that the q_i are rational so that there is a positive integer M such that $\hat{Y} := MY$ is an integer-valued random variable. Since $\wp(Y > y) = \wp(\hat{Y} > My)$, it suffices to compute the complementary cumulative distribution function of $\hat{Y}$. For integer-valued random variables, we have from [8] that

$$\wp(\hat{Y} \geq n) = \lim_{\mathcal{L}\to\infty} \frac{1}{2\mathcal{L}} \sum_{k=0}^{2\mathcal{L}-1} \nu_{\mathcal{L}}(k) C_{\hat{Y}}(k\pi/\mathcal{L}) e^{-jnk\pi/\mathcal{L}} \tag{6}$$

where

$$\nu_{\mathcal{L}}(k) := \begin{cases} \mathcal{L}, & k = 0 \\ 1 - j\cot\left(\frac{k\pi}{2\mathcal{L}}\right), & k \text{ odd} \\ 0, & \text{otherwise} \end{cases}$$

and $C_{\hat{Y}}$ is the characteristic function of $\hat{Y}$

$$C_{\hat{Y}}(\omega) := \mathsf{E}\left[e^{j\omega\hat{Y}}\right] = \mathsf{E}\left[e^{j(\omega M)Y}\right].$$

Combining this with (5) and the fact that ψ_1 is piecewise-constant, we have

$$C_{\hat{Y}}(\omega) = \exp\left(\int_0^{r_{\max}} [e^{-j\omega M\psi_1(r)} - 1] 2\lambda r\, \varphi(r)\, dr\right)$$

$$= \exp\left(\sum_{i=1}^{d-1} 2\lambda[e^{-j\omega M q_i} - 1] \int_{r_i}^{r_{i+1}} r\varphi(r)\, dr\right).$$
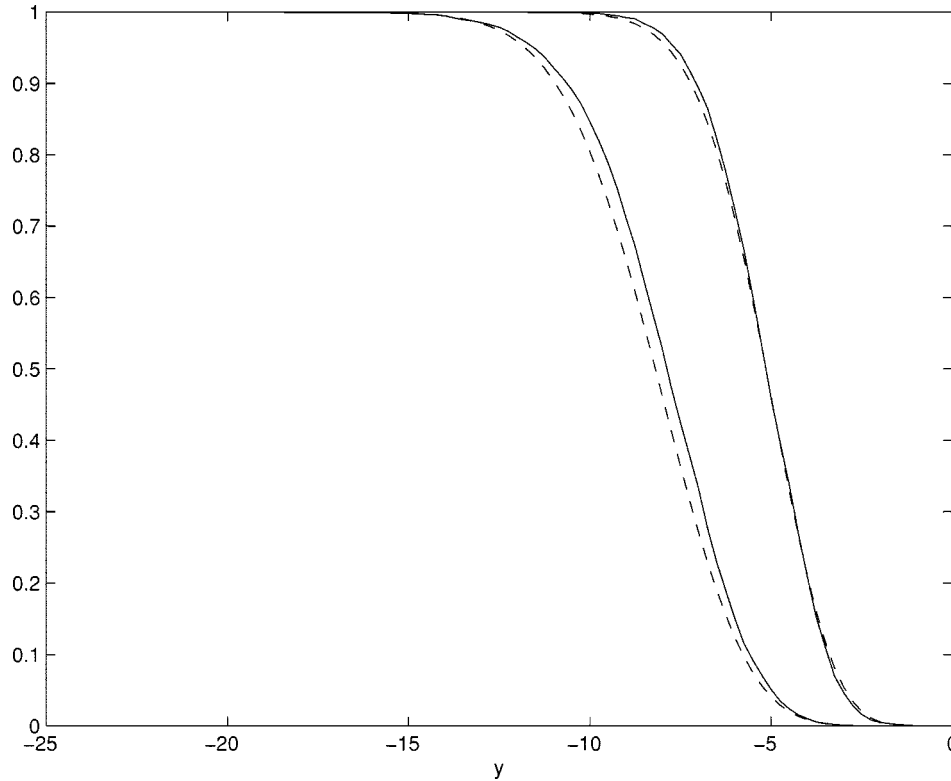
Fig. 3. Empirical estimates (solid lines) and shot-noise estimates (dashed lines) of the probabilities of false alarm (left) and detection (right) for Example 3. The horizontal axis y is the threshold of the test.

Furthermore,

$$\int_{r_i}^{r_{i+1}} r\varphi(r)\,dr = \frac{p_i}{2}\left(r_{i+1}^2 - r_i^2\right)$$

where if $\varphi = \varphi_0 \equiv 1$, $p_i = 1$ for all i, and if $\varphi = \varphi_1$, then $p_i = e^{-q_i}$ (recall that $\psi := -\ln\varphi$).

Example 3: Consider the hypothesis-testing problem for realizations of $n = 75$ points in a square region whose sides have length $L = 10$. From the analysis of the sparseness conditions in Appendix D, and (27) in particular, $\lambda = 87.1792$. Let

$$\varphi_1(r) = \begin{cases} e^{-3/4} \approx 0.4724, & 0 \le r \le 1/4 \\ e^{-1/4} \approx 0.7788, & 1/4 < r \le 1/2 \\ 1, & r > 1/2. \end{cases}$$

To approximate $\wp(Y > y)$ using (6), we must use a finite value of $\mathcal{L}$; we found that $\mathcal{L} = 64$ provided sufficient accuracy. A plot of the approximation of $\wp_0(Y > y)$ is given by the dashed line on the left in Fig. 3; the approximation of $\wp_1(Y > y)$ is given by the dashed line on the right. To estimate the true values of $\wp(\ell(X) > y)$, we used empirical estimates obtained from 5000 simulations of X under each hypothesis. These estimates are shown by the solid lines in Fig. 3, with the solid line on the left being the estimate of $\wp_0(\ell(X) > y)$, and the solid line on the right being the estimate of $\wp_1(\ell(X) > y)$.

B. Piecewise-Monotonic φ_1

In this subsection we assume that φ_1 is continuously differentiable on $[0, r_{\max}]$ with $\dot\varphi_1 > 0$ on $(0, r_{\max})$ and $\varphi_1(r) = 1$ for $r \ge r_{\max}$. The discussion of this case easily generalizes

when φ_1 is piecewise-continuously differentiable with derivative strictly of one sign on each piece, though the notation is more cumbersome; see Appendix C.

We begin by analyzing the characteristic function of Y. Let

$$B := \int_0^{r_{\max}} 2\lambda r\,\varphi(r)\,dr \tag{7}$$

and

$$\zeta(\omega) := \int_0^{r_{\max}} e^{-j\omega\psi_1(r)} 2\lambda r\,\varphi(r)\,dr. \tag{8}$$

Note that since φ is either φ_1 or φ_0, $B < \infty$, and the integrand in (8) is absolutely integrable. Using the definitions of B and ζ, the characteristic function of Y in (5) can be written as $\exp[\zeta(\omega) - B]$, which can be factored as

$$e^{-B}e^{\zeta(\omega)} = e^{-B}[1 + \zeta(\omega) + R(\omega)] \tag{9}$$

where

$$R(\omega) := \sum_{k=2}^{\infty} \frac{\zeta(\omega)^k}{k!} = e^{\zeta(\omega)} - \zeta(\omega) - 1. \tag{10}$$

Taking the inverse Fourier transform of (9), we have

$$p_Y(y) = e^{-B}[\delta(y) + \gamma(y) + g(y)] \tag{11}$$

where p_Y is the density of Y, γ is the inverse Fourier transform of ζ, and g is the inverse Fourier transform of R. Now, it is not obvious at this stage that γ and g exist; i.e., ζ and R might be
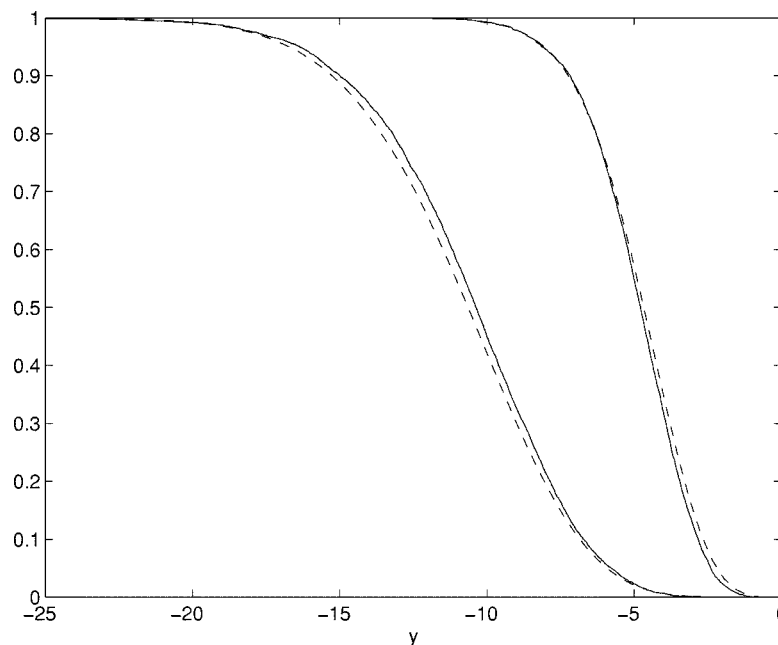
Fig. 4. Empirical estimates (solid lines) and shot-noise estimates (dashed lines) of the probabilities of false alarm (left) and detection (right) for Example 4. The horizontal axis y is the threshold of the test.

Fourier transforms of measures that are not absolutely continuous with respect to Lebesgue measure. Fortunately, these densities do exist, as shown in Appendix B. Next, write[2]

$$\wp(Y > y) = e^{-B}[1 + \Gamma^c(y) + G^c(y)], \qquad y < 0 \qquad (12)$$

where $\Gamma^c(y) := \int_y^\infty \gamma(\theta)\, d\theta$, and $G^c(y) := \int_y^\infty g(\theta)\, d\theta$. It is shown in Appendix B that

$$\Gamma^c(y) = \begin{cases} 0, & y \geq 0 \\ \int_{\varphi_1^{-1}(e^y)}^{r_{\max}} 2\lambda r\, \varphi(r)\, dr, & \ln \varphi_1(0) < y < 0 \\ B, & y \leq \ln \varphi_1(0) \end{cases} \qquad (13)$$

and that

$$G^c(y) = \lim_{\mathcal{L} \to \infty} \sum_{k=-\infty}^{\infty} b_k R(k\pi/\mathcal{L}) e^{-jk\pi y/\mathcal{L}} \qquad (14)$$

where $b_0 = 1/2$, $b_k = -j/(k\pi)$ for k odd, and $b_k = 0$ otherwise.

Remark 2: For numerical calculation, we use the approximation

$$G^c(y) \approx \sum_{k=-\mathcal{N}}^{\mathcal{N}} b_k R(k\pi/\mathcal{L}) e^{-jk\pi y/\mathcal{L}} \qquad (15)$$

where $\mathcal{L}$ and $\mathcal{N}$ are finite. Note that for plotting equally spaced samples, the approximation is efficiently computed using a fast Fourier transform routine.

---

[2]The assumptions on φ_1 imply that it is upper-bounded by one; hence ψ_1 is nonnegative, and Y in (4) is nonpositive. Since Y is nonpositive, $\wp(Y > y) = 0$ for $y \geq 0$.

Remark 3: If a closed-form antiderivative of $r\,\varphi(r)$ is known, then B in (7) and $\Gamma^c(y)$ in (13) are trivial to evaluate. In particular, when $\varphi = \varphi_0 \equiv 1$, $B = \lambda r_{\max}^2$, and $\Gamma^c(y) = \lambda[r_{\max}^2 - \varphi_1^{-1}(e^y)^2]$.

Example 4: Consider the hypothesis-testing problem for realizations of $n = 75$ points in a square region whose sides have length $L = 10$. From the analysis of the sparseness conditions in Appendix D, and (27) in particular, $\lambda = 87.1792$. Let φ_1 be the interaction function of Example 2. Then the quantities B, Γ^c, $\zeta(\omega)$, and $R(\omega)$ are all easily computed in closed form under each hypothesis. Taking $\mathcal{L} = 25$ and $\mathcal{N} = 128$ provided sufficient accuracy in (15). We then substituted (15) into (12). A plot of the approximation of $\wp_0(Y > y)$ is given by the dashed line on the left in Fig. 4; the approximation of $\wp_1(Y > y)$ is given by the dashed line on the right. To estimate the true values of $\wp_0(\ell(X) > y)$ and $\wp_1(\ell(X) > y)$, we used empirical estimates obtained from 5000 simulations of X under each hypothesis. These estimates are shown by the solid lines in Fig. 4, with the solid line on the left being the estimate of $\wp_0(\ell(X) > y)$, and the solid line on the right being the estimate of $\wp_1(\ell(X) > y)$.

V. Conclusion

Although the sufficient statistic $\ell(X)$ for performing the likelihood ratio test for pairwise interaction point processes is easy to compute, the evaluation of its performance is a challenging problem. The limit theorem of Appendix A shows that in the case of sparseness, the distribution of $\ell(X)$ could be approximated by the distribution of the shot-noise random variable Y, and it is then shown that the distribution of Y could be computed using (6) or (12), depending on the form of the interaction function. While the analysis of (6) in Section IV-A is straightforward, the derivation of (12) in Section IV-B is more complicated, and the details are found in Appendix B. Extensions

are treated in Appendix C. Finally, we note that our analysis of the sparseness conditions in Appendix D in the case of square regions leads to a simply computable value for the constant λ which appears in the characteristic function of Y in (5).

APPENDIX A
LIMIT THEOREM

This result establishes the approximation

$$\wp(\ell(X) > y) \approx \wp(Y > y),$$

where Y is defined in (4).

Theorem: Let $\{D_n\}$ be an increasing sequence of bounded Borel subsets of the plane for which the sparseness conditions in Appendix D hold. For example, if D_n is a square whose sides have length proportional to n, then the sparseness conditions hold (see Appendix D). Suppose that $\varphi_1(r)$ is piecewise-continuous, positive almost everywhere with respect to Lebesgue measure, and upper-bounded by 1. Also assume $\varphi_1(r) = 1$ for $r \geq r_{\max}$. For each n, let $X^{(n)} = (X_1^{(n)}, \cdots, X_n^{(n)})$ be a $(D_n)^n$-valued interaction point process with common interaction function φ, where φ is either φ_1 or $\varphi_0 \equiv 1$. Then $\ell(X^{(n)})$ converges to Y in distribution; i.e., for all y at which $\wp(Y > y)$ is continuous

$$\lim_{n \to \infty} \wp(\ell(X^{(n)}) > y) = \wp(Y > y).$$

Proof: Let $\mathcal{S}$ denote the space of right-continuous, piecewise-constant, integer-valued functions on $[0, r_{\max}]$. For $s \in \mathcal{S}$, put

$$\hat{\ell}(s) := -\int_0^{r_{\max}} \psi_1(r)\, s(dr).$$

With this notation, we can write (recall (3) and (4)) $\ell(X^{(n)}) = \hat{\ell}(S(\cdot, X^{(n)}))$ and $Y = \hat{\ell}(S_\infty)$.

Now, the hypotheses of our theorem are sufficient for us to apply [3, Theorem 2.3], which says that the sequence of processes $\{S(r, X^{(n)}), 0 \leq r \leq r_{\max}\}$ converges in distribution to the Poisson process $\{S_\infty(r), 0 \leq r \leq r_{\max}\}$. It remains to show that $\hat{\ell}(S(\cdot, X^{(n)}))$ converges in distribution to $\hat{\ell}(S_\infty) = Y$ as well.

Following Billingsley [1, pp. 30–31, Theorem 5.1], it suffices to find a set $\mathcal{S}_1 \subset \mathcal{S}$ such that not only is $\wp(S_\infty \in \mathcal{S}_1) = 0$, but also $\hat{\ell}$ is continuous on $\mathcal{S}_1^c$. To this end, let $\mathcal{D}_d$ denote the set of discontinuities of φ_1 in $[0, r_{\max}]$. Let

$$\mathcal{D}_z := \{r \in [0, r_{\max}]: \varphi_1(r) = 0\}.$$

Put $\mathcal{D}_1 := \mathcal{D}_d \cup \mathcal{D}_z$. Then $\mathcal{D}_1$ is the union of a finite set and a set of measure zero. Thus $\mathcal{D}_1$ has measure zero. Next, let $\mathcal{S}_1$ denote the set of $s \in \mathcal{S}$ such that s has at least one jump discontinuity in $\mathcal{D}_1$. Then

$$\wp(S_\infty \in \mathcal{S}_1) = \wp(S_\infty(\mathcal{D}_1) \geq 1) = 1 - \wp(S_\infty(\mathcal{D}_1) = 0)$$

where $S_\infty(\mathcal{D}_1)$ is the number of points of the Poisson process that occur in $\mathcal{D}_1$. Since $\mathcal{D}_1$ has measure zero, and since S_∞ is a Poisson process with an integrable, nonimpulsive intensity, $\wp(S_\infty(\mathcal{D}_1) = 0) = 1$. Hence, $\wp(S_\infty \in \mathcal{S}_1) = 0$.

We now turn to the continuity of $\hat{\ell}$ on $\mathcal{S}_1^c$. Let s_k be a sequence in $\mathcal{S}_1^c$ that converges to $s \in \mathcal{S}_1^c$. We must show that $\hat{\ell}(s_k) \to \hat{\ell}(s)$. Now, in [3, Theorem 2.3], which establishes the convergence in distribution of $S(\cdot, X^{(n)})$ to S_∞, the space $\mathcal{S}$ is equipped with the Skorohod topology [1, pp. 111–112]. To say that $s_k \to s$ in the Skorohod topology means the following. Let $\mathcal{M}$ be the class of strictly increasing, continuous mappings of $[0, r_{\max}]$ onto itself. Then there exist $\mu_k \in \mathcal{M}$ such that

$$\lim_{k \to \infty} s_k(\mu_k^{-1}(r)) = s(r), \qquad \text{uniformly in } r \qquad (16)$$

and

$$\lim_{k \to \infty} \mu_k^{-1}(r) = r, \qquad \text{uniformly in } r. \qquad (17)$$

Since s_k and s take only integer values, (16) implies that for sufficiently large k, $s_k = s \circ \mu_k$. For such k

$$|\hat{\ell}(s_k) - \hat{\ell}(s)| = |\hat{\ell}(s \circ \mu_k) - \hat{\ell}(s)|.$$

Suppose s has J jumps in $[0, r_{\max}]$ located at $r_1 < \cdots < r_J$. Then

$$|\hat{\ell}(s \circ \mu_k) - \hat{\ell}(s)| = \left| \sum_{j=1}^{J} \psi_1(\mu_k^{-1}(r_j)) - \psi_1(r_j) \right|$$

$$\leq \sum_{j=1}^{J} |\psi_1(\mu_k^{-1}(r_j)) - \psi_1(r_j)|.$$

By (17), $\mu_k^{-1}(r_j) \to r_j$ as $k \to \infty$. Since $s \in \mathcal{S}_1^c$, ψ_1 is finite and continuous at all the r_j. Hence, $\psi_1(\mu_k^{-1}(r_j)) \to \psi_1(r_j)$ as well. $\qquad \square$

Discussion

The preceding proof appeals to [3, Theorem 2.3]. The hypotheses of that theorem are identical to those of our limit theorem, except for our extra assumption that φ_1 be positive almost everywhere. To understand this extra assumption, it is helpful to consider a simple modification of our limit theorem to handle the case in which φ_0 is not identically one. In this case, ℓ is given by (2), and the modified hypotheses for our limit theorem are that both φ_0 and φ_1 should be piecewise-continuous and upper-bounded by 1 *and* that the sets

$$\mathcal{D}_{01} := \{r \in [0, r_{\max}]: \varphi_0(r) = 0 \text{ and } \varphi_1(r) > 0\}$$

and

$$\mathcal{D}_{10} := \{r \in [0, r_{\max}]: \varphi_0(r) > 0 \text{ and } \varphi_1(r) = 0\}$$

should have measure zero. The proof would be as before, except that $\mathcal{D}_d$ is replaced by the union of the discontinuities of φ_0 and φ_1, again a finite set, and $\mathcal{D}_z$ is replaced by $\mathcal{D}_{01} \cup \mathcal{D}_{10} \cup \mathcal{D}_{00}$, where

$$\mathcal{D}_{00} := \{r \in [0, r_{\max}]: \varphi_0(r) = \varphi_1(r) = 0\}. \qquad (18)$$

Note that since S_∞ has intensity $2\lambda r \varphi(r)$, $\wp(S_\infty(\mathcal{D}_{00}) = 0) = 1$ for both $\varphi = \varphi_0$ and $\varphi = \varphi_1$.

APPENDIX B
DERIVATION OF (11), (13), AND (14)

Let φ_1 be continuously differentiable on $[0, r_{\max}]$, with $\dot{\varphi}_1 > 0$ on $(0, r_{\max})$, and $\varphi_1(r) = 1$ for $r \geq r_{\max}$.

The key to the derivations is to show that ζ in (8) is the Fourier transform of a nonnegative, integrable function γ. The assumptions on φ_1 imply that it has an inverse

$$\varphi_1^{-1}\colon [\varphi_1(0), 1] \to [0, r_{\max}].$$

Furthermore, since $\dot{\varphi}_1(r)$ is continuous and positive for $r \in (0, r_{\max})$

$$\frac{d}{dt}\, \varphi_1^{-1}(t) = \frac{1}{\dot{\varphi}_1(\varphi_1^{-1}(t))}, \qquad t \in (\varphi_1(0), 1)$$

is also continuous. Making the change of variable $\theta = -\psi_1(r)$ in (8) yields

$$\zeta(\omega) = \int_{\ln \varphi_1(0)}^{0} e^{j\omega\theta} \gamma(\theta)\, d\theta \qquad (19)$$

where

$$\gamma(\theta) := \frac{2\lambda \varphi_1^{-1}(e^{\theta})\, \varphi(\varphi_1^{-1}(e^{\theta})) e^{\theta}}{\dot{\varphi}_1(\varphi_1^{-1}(e^{\theta}))} \qquad (20)$$

for $\ln \varphi_1(0) < \theta < 0$. For θ outside this range, set $\gamma(\theta) := 0$. Since γ is nonnegative, we can show it is integrable by setting $\omega = 0$ in (19); then from (8) and (7), $\zeta(0) = B < \infty$.

Now recall that $\gamma(\theta)$ is defined to be zero for $\theta \leq \ln \varphi_1(0)$ and for $\theta \geq 0$. This implies that for $y < 0$

$$\Gamma^c(y) = \int_y^{\infty} \gamma(\theta)\, d\theta = \int_y^0 \gamma(\theta)\, d\theta. \qquad (21)$$

For $\ln \varphi_1(0) < y < 0$, we make the change of variable $\theta = -\ln \psi_1(r)$, and obtain the integral in (13). For $y \leq \ln \varphi_1(0)$, the lower limit in (21) can be set to $\ln \varphi_1(0)$; but this integral is exactly $\zeta(0)$ in (19), which equals B.

We next show the existence of g in (11). Since ζ is the Fourier transform of γ, we see from the definition of R in (10) that it is the Fourier transform of

$$g(y) := \sum_{k=2}^{\infty} \frac{\gamma^{*k}(y)}{k!} \qquad (22)$$

where γ^{*k} is the k-fold convolution of γ with itself. We remark that since γ is zero on $[0, \infty)$, the same holds for the convolutions γ^{*k}, and thus for g as well.

To conclude, we turn to (14). This equation is exactly [9, p. 753, Theorem 1], the hypothesis of which is

$$\lim_{W \to \infty} \frac{1}{2W} \int_{-W}^{W} |R(\omega)|^2\, d\omega = 0.$$

However, this hypothesis is equivalent to the requirement that the distribution corresponding to R have no point masses [2, p. 306]. In other words, the density g must not contain impulses, which is clearly true in (22) since γ contains no impulses.

APPENDIX C
MORE GENERAL INTERACTION FUNCTIONS

We now generalize the results of Section IV-B and of Appendix B. Let φ_1 be upper-bounded by one, and suppose that for some $d \geq 3$, there exist points $0 = r_1 < \cdots < r_d = r_{\max}$ such that φ_1 is continuously differentiable with derivative strictly of one sign on (r_i, r_{i+1}). Furthermore, assume that φ_1 and $\dot{\varphi}_1$ have finite left and/or right limits at each r_i. We continue to assume $\varphi_1(r) = 1$ for $r \geq r_{\max}$.

Under the above assumptions, $\psi_1(r) \neq 0$ for $r \neq r_i$. From (8) we can write

$$\zeta(\omega) = \sum_{i=1}^{d-1} \zeta_i(\omega)$$

where

$$\zeta_i(\omega) := \int_{r_i}^{r_{i+1}} e^{-j\omega\psi_1(r)} 2\lambda r\, \varphi(r)\, dr. \qquad (23)$$

The restriction of φ_1 to $[r_i, r_{i+1}]$, taking φ_1 and $\dot{\varphi}_1$ to have the their right and left limit values at r_i and r_{i+1}, respectively, is continuously differentiable with derivative strictly of one sign on (r_i, r_{i+1}). Hence, φ_1 has a local inverse

$$\varphi_{1,i}^{-1}\colon [t_i, t_{i+1}] \to [r_i, r_{i+1}]$$

where, if φ_1 is increasing on $[r_i, r_{i+1}]$, then $t_i = \varphi_1(r_i)$ and $t_{i+1} = \varphi_1(r_{i+1})$; if φ_1 is decreasing on $[r_i, r_{i+1}]$, then $t_i = \varphi_1(r_{i+1})$ and $t_{i+1} = \varphi_1(r_i)$. Next, making the change of variable $\theta = -\psi_1(r)$ in the definition of ζ_i yields

$$\zeta_i(\omega) = \int_{\ln t_i}^{\ln t_{i+1}} e^{j\omega\theta} \gamma_i(\theta)\, d\theta$$

where

$$\gamma_i(\theta) := \frac{2\lambda \varphi_{1,i}^{-1}(e^{\theta})\, \varphi(\varphi_{1,i}^{-1}(e^{\theta})) e^{\theta}}{|\dot{\varphi}_1(\varphi_{1,i}^{-1}(e^{\theta}))|}$$

for $\ln t_i < \theta < \ln t_{i+1}$, and $\gamma_i(\theta) := 0$ otherwise. Since we have defined γ_i to be zero outside $\ln t_i < \theta < \ln t_{i+1}$, the range of integration above may be replaced by $(-\infty, \infty)$. Hence

$$\begin{aligned} \zeta(\omega) &= \sum_{i=1}^{d-1} \zeta_i(\omega) \\ &= \sum_{i=1}^{d-1} \int_{-\infty}^{\infty} e^{j\omega\theta} \gamma_i(\theta)\, d\theta \\ &= \int_{-\infty}^{\infty} e^{j\omega\theta} \gamma(\theta)\, d\theta \end{aligned}$$

where

$$\gamma(\theta) := \sum_{i=1}^{d-1} \gamma_i(\theta).$$

The analog of (13) is then

$$\Gamma^c(y) = \sum_{i=1}^{d-1} \Gamma_i^c(y)$$

where for $\ln t_i < \theta < \ln t_{i+1}$

$$\Gamma_i^c(y) = \int_{\varphi_{1,i}^{-1}(e^y)}^{r_{i+1}} 2\lambda r\,\varphi(r)\,dr, \quad \varphi_1 \text{ increasing on } [r_i, r_{i+1}]$$

and

$$\Gamma_i^c(y) = \int_{r_i}^{\varphi_{1,i}^{-1}(e^y)} 2\lambda r\,\varphi(r)\,dr, \quad \varphi_1 \text{ decreasing on } [r_i, r_{i+1}].$$

Discussion

To handle the case in which φ_0 is not identically one, we note that ℓ is then given by (2) instead of (3). Also, in (5), (8), and (23), $-\psi_1(r)$ is replaced by $\psi_0(r) - \psi_1(r)$. The above analysis can be carried out as before if we use the change of variable $\theta = \psi_0(r) - \psi_1(r)$, and if we assume that on (r_i, r_{i+1}), $\varphi_1(r)/\varphi_0(r)$ is continuously differentiable with derivative strictly of one sign. We can even handle the case in which for some intervals $(r_i, r_{i+1}]$, both $\varphi_0(r)$ and $\varphi_1(r)$ are zero (cf. the set $\mathcal{D}_{00}$ in (18) in Appendix A), since for such i, $\zeta_i(\omega)$ in (23) is zero.

APPENDIX D
THE SPARSENESS CONDITIONS

The sparseness conditions referred to in the paper are given for a sequence of regions $\{D_n\}$ whose areas grow as $n \to \infty$ [3], [16]. The sparseness conditions are purely geometric constraints on how the region D_n grows as the number of points n increases. We use the following notation. Let $\mathcal{B}_r(\xi)$ denote the closed ball (disk in $\mathbb{R}^2$) centered at ξ with radius r

$$\mathcal{B}_r(\xi) := \{\xi' \in \mathbb{R}^2 : \|\xi' - \xi\| \le r\}$$

and let $\mathcal{I}_n(r)$ denote the "r-interior"

$$\mathcal{I}_n(r) := \{\xi \in D_n : \mathcal{B}_r(\xi) \subseteq D_n\}.$$

For example, if D_n is a square whose sides have length L_n, then $\mathcal{I}_n(r)$ is the square with the same center but whose sides have length $L_n - 2r$.

Next, let

$$\lambda_n(r) := \frac{\pi n(n-1)}{2} \cdot \frac{\text{area}(\mathcal{I}_n(r))}{\text{area}(D_n)^2},$$

and let

$$v_n(r) := \frac{n(n-1)}{2} \int_{D_n \backslash \mathcal{I}_n(r)} \frac{\text{area}(\mathcal{B}_r(\xi) \cap D_n)}{\text{area}(D_n)^2}\,d\xi.$$

The first sparseness condition is $\lim_{n \to \infty} \lambda_n(r) = \lambda$, where the limit λ is finite, positive, and does not depend on r. The second sparseness condition is $\lim_{n \to \infty} v_n(r) = 0$.

For the square whose sides have length L_n

$$\frac{\text{area}(\mathcal{I}_n(r))}{\text{area}(D_n)^2} = \frac{L_n^2 - 4L_n r + 4r^2}{L_n^4}$$

$$= \frac{1}{L_n^2} - \frac{4r}{L_n^3} + \frac{4r^2}{L_n^4}.$$

Then

$$\lambda_n(r) = \frac{\pi n(n-1)}{2L_n^2}\left(1 - \frac{4r}{L_n} + \frac{4r^2}{L_n^2}\right)$$

and we see that the first sparseness condition will hold if and only if L_n^2 is asymptotically proportional to $n(n-1)$. If in fact

$L_n^2 = Kn(n-1)$ for some constant K, which can always be arranged in practice as shown below, then

$$\lambda_n(r) = \frac{\pi}{2K}\left(1 - \frac{4r}{L_n} + \frac{4r^2}{L_n^2}\right)$$

and $\lambda_n(r) \to \lambda = \pi/(2K)$. However, the extent to which the first sparseness condition holds when n is finite is determined by absolute error

$$\lambda - \lambda_n(r) = \lambda \cdot 4\frac{r}{L_n}\left(1 - \frac{r}{L_n}\right) \qquad (24)$$

and the relative error

$$\frac{\lambda - \lambda_n(r)}{\lambda} = 4\frac{r}{L_n}\left(1 - \frac{r}{L_n}\right). \qquad (25)$$

We now turn to the second sparseness condition when D_n is a square whose sides have length L_n. In this case

$$\frac{\text{area}(\mathcal{B}_r(\xi) \cap D_n)}{\text{area}(D_n)^2} \le \frac{\text{area}(\mathcal{B}_r(\xi))}{\text{area}(D_n)^2} = \frac{\pi r^2}{L_n^4}.$$

Since $\text{area}(D_n \backslash \mathcal{I}_n(r)) = L_n^2 - (L_n - 2r)^2 = 4rL_n - 4r^2$

$$v_n(r) \le \frac{n(n-1)}{2} \cdot \frac{\pi r^2 (4rL_n - 4r^2)}{L_n^4}.$$

This goes to zero if L_n^3 grows faster than $n(n-1)$, as is the case if $L_n^2 = Kn(n-1)$. To see the extent to which sparseness holds when n is finite, substitute $n(n-1) = L_n^2/K$ and $K = \pi/(2\lambda)$. Then

$$v_n(r) \le \lambda \cdot 4\frac{r^3}{L_n}\left(1 - \frac{r}{L_n}\right). \qquad (26)$$

Now, in practice, we are given a fixed region D, say a square with sides of length L, and a fixed number of points, say N. Using $L_n^2 = Kn(n-1)$, we require that $L_N^2 = L^2$. This leads to

$$K = \frac{L^2}{N(N-1)}$$

and hence

$$\lambda = \frac{\pi}{2K} = \frac{\pi N(N-1)}{2L^2}. \qquad (27)$$

REFERENCES

[1] P. Billingsley, *Convergence of Probability Measures*. New York: Wiley, 1968.
[2] ——, *Probability and Measure*. New York: Wiley, 1979.
[3] W.-B. Chang and J. A. Gubner, "Poisson limits and nonparametric estimation for pairwise interaction point processes," *J. Appl. Probab.*, vol. 37, no. 1, Mar. 2000, to be published.
[4] N. A. C. Cressie, *Statistics for Spatial Data*, revised ed. New York: Wiley, 1993.
[5] P. J. Diggle, D. J. Gates, and A. Stibbard, "A nonparametric estimator for pairwise interaction point processes," *Biometrika*, vol. 74, pp. 763–770, 1987.
[6] P. J. Diggle, T. Fiksel, P. Grabarnik, Y. Ogata, D. Stoyan, and M. Tanemura, "On parameter estimation for pairwise interaction point processes," *Int. Statist. Rev.*, vol. 62, pp. 99–117, 1994.
[7] C. J. Geyer and J. Møller, "Simulation procedures and likelihood inference for spatial point processes," *Scand. J. Statist.*, vol. 21, pp. 359–373, 1994.
[8] J. A. Gubner and M. M. Hayat, "A method to recover counting distributions from their characteristic functions," *IEEE Signal Processing Lett.*, vol. 3, pp. 184–186, June 1996.
[9] J. A. Gubner, "Computation of shot-noise probability distributions and densities," *SIAM J. Sci. Comput.*, vol. 17, no. 3, pp. 750–761, May 1996.
[10] J. F. C. Kingman, *Poisson Processes*. Oxford, U.K.: Clarendon , 1993.

[11] R. A. Moyeed and A. J. Baddeley, "Stochastic approximation of the MLE for a spatial point pattern," *Scand. J. Statist.*, vol. 18, pp. 39–50, 1991.

[12] Y. Ogata and M. Tanemura, "Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure," *Ann. Inst. Statist. Math.*, pt. B, vol. 33, pp. 315–338, 1981.

[13] ——, "Likelihood analysis of spatial point patterns," *J. Roy. Statist. Soc. Ser. B*, vol. 46, pp. 496–518, 1984.

[14] A. Penttinen, "Modelling interaction in spatial point patterns: Parameter estimation by the maximum likelihood method," *Jyväsklä Stud. Comput. Sci. Econ. Statist.*, vol. 7, pp. 1–107, 1984.

[15] B. D. Ripley, *Statistical Inference for Spatial Processes*.   Cambridge, U.K.: Cambridge Univ. Press, 1988.

[16] R. Saunders and G. M. Funk, "Poisson limits for a clustering model of Strauss," *J. Appl. Probab.*, vol. 14, pp. 776–784, 1977.

[17] D. L. Snyder and M. I. Miller, *Random Point Processes in Time and Space*, 2nd ed.   New York: Springer-Verlag, 1991.