Statistics and Probability Letters 242 (2027) 111008

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

A counterexample to a conjecture on the bandwidth selection criterion of Cronie and Van Lieshout

O. Cronie [a,*], M.N.M. van Lieshout [b]

[a] Department of Mathematical Sciences, Chalmers University of Technology & University of Gothenburg, Gothenburg, Sweden
[b] Centrum Wiskunde & Informatica & University of Twente, Amsterdam & Enschede, The Netherlands

ARTICLE INFO

MSC:
60G55
62G07
62M30

Keywords:
Bandwidth selection
Gaussian kernel
Kernel estimation
Spatial point pattern
Stoyan–Grabarnik statistic

ABSTRACT

Cronie and Van Lieshout (2018) introduced a nonparametric bandwidth selector for kernel estimation of spatial point process intensity functions. Using the Gaussian kernel and no edge correction term in their criterion, they conjectured that the associated statistic $T_\kappa(h;\psi,W)$ is a monotone function of the bandwidth h. We show that there exist point patterns ψ for which the derivative of this statistic is strictly negative at a positive bandwidth h, which, because the statistic tends to zero for $h \to 0$ and to infinity for $h \to \infty$, disproves the conjecture. Such point patterns consist of a sufficiently large cluster of tightly packed points together with a single well-separated point.

1. Preliminaries

Consider a simple point process X in $\mathbb{R}^d$, $d \geq 1$. Formally, given an underlying probability space, X can be defined as a random element in the measurable space $\mathcal{N}_{\mathrm{lf}}$ of countable configurations $\psi = \{x_i\}_{i \in I_\psi}$ with the local finiteness property that $\#(\psi \cap A)$ is finite for any bounded Borel set $A \subseteq \mathbb{R}^d$ (Chiu et al., 2013). In particular, the index set I_ψ is either finite or countable. An observed realization $\psi \in \mathcal{N}_{\mathrm{lf}}$ of X, observed within some bounded window W, is referred to as a point pattern and an $x_i \in \psi$ is called an event.

1.1. Kernel intensity estimation

The intensity function $\lambda : \mathbb{R}^d \to \mathbb{R}^+$ of X has the heuristic interpretation that the probability of X having a point in an infinitesimal neighbourhood $\mathrm{d}x$ of $x \in \mathbb{R}^d$ is given by $\lambda(x)\mathrm{d}x$. Formally, it can be defined through the Campbell formula (Chiu et al., 2013, Section 4.3), which states that

$$\mathrm{E}\left\{\sum_{\xi \in X} f(\xi)\right\} = \int_{\mathbb{R}^d} f(x)\lambda(x)\mathrm{d}x$$

for any non-negative and measurable function f on $\mathbb{R}^d$.

When analysing point pattern data, the first step is typically to generate a non-parametric estimate of the underlying intensity function λ, and here the most prominent approach is kernel estimation (Chiu et al., 2013, Section 4.7). Given a point pattern ψ

* Corresponding author.
E-mail address: ottmar@chalmers.se (O. Cronie).

https://doi.org/10.1016/j.spl.2026.111008
Received 1 July 2026; Received in revised form 21 September 2026; Accepted 29 October 2026
Available online 5 November 2026
0167-7152/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

within a non-empty, open and bounded Borel set $W \subseteq \mathbb{R}^d$, with $\psi \cap W \neq \emptyset$, the kernel estimate of the intensity function, when no edge correction has been applied, is given by

$$\widehat{\lambda}(x; h, \psi, W) = h^{-d} \sum_{y \in \psi \cap W} \kappa\{(x-y)/h\}, \qquad x \in W. \tag{1}$$

Here, $h > 0$ is the smoothing bandwidth and the kernel κ is a symmetric density function on $\mathbb{R}^d$. Clearly, a variety of kernels can be applied here, but a natural and common choice is the Gaussian kernel

$$\kappa(x) = (2\pi)^{-d/2} \exp(-x^T x/2), \qquad x \in \mathbb{R}^d. \tag{2}$$

Other kernels include, for example, the beta kernel (Wand and Jones, 1995, Section 2.7).

1.2. The bandwidth selection criterion

Assuming that $\lambda(x) > 0$ for Lebesgue-almost every $x \in W$, the Campbell formula applied to $f(x) = 1/\lambda(x)$ whenever $\lambda(x) > 0$, with $f(x) = 0$ otherwise, yields $\mathrm{E}\{\sum_{\xi \in X \cap W} 1/\lambda(\xi)\} = |W|$, where $|W|$ denotes the Lebesgue measure of W. Motivated by this identity, Cronie and Van Lieshout (2018) proposed to choose the bandwidth h by minimizing

$$h \mapsto \{T_\kappa(h; \psi, W) - |W|\}^2, \qquad h > 0, \tag{3}$$

where

$$T_\kappa(h; \psi, W) = \sum_{x \in \psi \cap W} \frac{1}{\widehat{\lambda}(x; h, \psi, W)} = h^d \sum_{x \in \psi \cap W} \frac{1}{\sum_{y \in \psi \cap W} \kappa\{(x-y)/h\}}. \tag{4}$$

This is an instance of the statistic of Stoyan and Grabarnik (1991) , which can also be used for parametric inference (Kresin, 2023). In the case of the Gaussian kernel,

$$T_\kappa(h; \psi, W) = (2\pi)^{d/2} \widetilde{T}_\kappa(h; \psi, W),$$
$$\widetilde{T}_\kappa(h; \psi, W) = h^d \sum_{x \in \psi \cap W} \frac{1}{\sum_{y \in \psi \cap W} \exp\{-\|x-y\|^2/(2h^2)\}}.$$

Theorem 1 of Cronie and Van Lieshout (2018) implies, among other things, that for the Gaussian kernel and for beta kernels other than the box kernel $T_\kappa(h; \psi, W)$ is continuous in h, tends to zero as $h \downarrow 0$ and tends to infinity as $h \to \infty$. Consequently, there exists at least one minimizer of (3) and plugging such a minimizer into (3) results in the value zero.

A minimizer of (3) is not necessarily unique, unless $T_\kappa(h; \psi, W)$ is a monotonically increasing function of h. Cronie and Van Lieshout (2018) noted that monotonicity may fail for general kernels, but provided the following conjecture.

Conjecture 1. *Monotonicity of $T_\kappa(h; \psi, W)$, $h > 0$, holds for any point pattern $\psi \subseteq W$ and any non-empty open and bounded $W \subseteq \mathbb{R}^d$, $d \geq 1$, if κ is the Gaussian kernel.*

The purpose of this note is to disprove this conjecture. Consequently, the Gaussian kernel displays behaviour similar to the continuous beta kernels. In practice, in case the graph of $h \mapsto T_\kappa(h; \psi, W)$ hits the value $|W|$ multiple times, we pick the first crossing.

2. A counterexample to the conjecture

For the Gaussian kernel, note that we can focus on $\widetilde{T}_\kappa(h; \psi, W)$ since the scaling $(2\pi)^{d/2}$ neither depends on the bandwidth nor the point pattern in question.

2.1. Motivation

Considering $\psi \cap W = \{x_1, \ldots, x_n\}$, we have

$$\widetilde{T}_\kappa(h; \psi, W) = h^d \sum_{i=1}^{n} \frac{1}{S_{x_i;\psi}(h)}, \quad S_{x_i;\psi}(h) = 1 + \sum_{j \neq i} \exp\left\{-\frac{\|x_i - x_j\|^2}{2h^2}\right\}, \tag{5}$$

since the term corresponding to $j = i$ is equal to $\exp(0) = 1$. If $\|x_i - x_j\| \geq Lh$ for all $j \neq i$, then

$$1 \leq S_{x_i;\psi}(h) \leq 1 + (n-1)\exp(-L^2/2),$$

which tends to one as $L \to \infty$. Thus, if all points are separated by a large multiple of h, $\tilde{T}_\kappa(h; \psi, W)$ is approximately equal to nh^d, which is a strictly increasing function of h. Similarly, if all points are close together, $\tilde{T}_\kappa(h; \psi, W)$ is approximately equal to h^d , which is also strictly increasing in h. This suggests that in order to obtain a counterexample to Conjecture 1, one needs to pack many points tightly together in combination with some points far away.

We next make these heuristics explicit.

2.2. Tightly clustered point patterns and the counterexample

Consider a non-empty open and bounded Borel set $W \subseteq \mathbb{R}^d$, $d \geq 1$, and assume, without loss of generality, that it contains the origin, i.e., $0 \in W$. Since W is open, there is a $y \in W$ which has distance $\|y\| = R > 0$ to the origin. Now consider an $m+1$ tuple $\psi = (0, \dots, 0, y)$. Note that due to the multiple points at the origin, this ensemble is not a point pattern and will be perturbed at a later stage.

Suppressing the dependence on ψ and W, let $S_y(h) = m\exp\{-R^2/(2h^2)\} + 1$ and $S_0(h) = m + \exp\{-R^2/(2h^2)\}$. Then, we have that

$$T_\kappa(h) \propto \widetilde{T}_\kappa(h) = h^d \left\{ mS_0(h)^{-1} + S_y(h)^{-1} \right\}.$$

Writing $a(h) = \exp\{-R^2/(2h^2)\}$,

$$\tilde{T}_\kappa(h) = h^d \left\{ \frac{m}{m + a(h)} + \frac{1}{1 + ma(h)} \right\},$$

so that

$$\frac{\partial}{\partial h}\{\widetilde{T}_\kappa(h)\} = h^{d-1}\left[d \left\{ \frac{m}{m + a(h)} + \frac{1}{1 + ma(h)} \right\} - \frac{R^2 ma(h)}{h^2} \left\{ \frac{1}{(m + a(h))^2} + \frac{1}{(1 + ma(h))^2} \right\} \right]. \tag{6}$$

It remains to transform the $m+1$ tuple $(0, \dots, 0, y)$ into a proper point pattern $(x_1, \dots, x_m, y)$. Since for fixed $h > 0$, the function

$$(x_i, x_j) \mapsto \exp\{-\|x_i - x_j\|^2/(2h^2)\}, \qquad x_i, x_j \in W,$$

and its h-derivative are continuous in the point locations, the derivative of $\tilde{T}_\kappa(h; \psi, W)$ is continuous under perturbations of the events in ψ. Hence, if the $(m+1)$-tuple $(0, \dots, 0, y)$ has strictly negative derivative, the same is true for all sufficiently small perturbations of the m points at the origin, and, since W is open, the perturbed points may be chosen in W as well. Recalling that $T_\kappa(h; \psi, W)$ tends to zero as $h \downarrow 0$ and to infinity as $h \to \infty$, the existence of a strictly negative derivative for some $h \in (0, \infty)$ implies that $T_\kappa(h; \psi, W)$ is neither monotonically increasing nor decreasing.

We summarize the conclusion of this counterexample in a theorem.

Theorem 1. *Given a non-empty, open and bounded set $W \subseteq \mathbb{R}^d$, $d \geq 1$, which contains the origin and $y \in W$ with $\|y\| = R > 0$, assume that the right-hand side of (6) is strictly negative. Then we can find $\varepsilon > 0$ such that, for every point pattern $\psi_\varepsilon = \{\tilde{x}_1, \dots, \tilde{x}_m, y\} \subseteq W$, where $\tilde{x}_i \in B(0, \varepsilon)$, $i = 1, \dots, m$, are distinct and $B(0, \varepsilon) \subseteq W$, we have*

$$\frac{\partial}{\partial h} T_\kappa(h; \psi_\varepsilon, W) < 0,$$

whereby $T_\kappa(h; \psi_\varepsilon, W)$ is not monotonically increasing in h.

The construction above should not be interpreted as yielding a single exceptional point pattern. Rather, since the derivative of $T_\kappa(h; \psi, W)$ for a fixed positive bandwidth h depends continuously on the locations of the points, the derivative remains strictly negative under sufficiently small perturbations of the point pattern. Consequently, non-monotonicity occurs for a whole neighbourhood of simple point patterns having the same qualitative geometry, namely a sufficiently tight and sufficiently large cluster together with a separated point. Hence, the phenomenon is robust with respect to perturbations of the locations and is not tied to the limiting configuration with coincident points.

Our argument does not, however, provide a characterization of the point patterns for which $T_\kappa(h; \psi, W)$ is monotone, nor does it determine how frequently non-monotonicity occurs under particular point process models. The present result also does not address configurations with multiple clusters or multiple isolated points. These questions would require a more systematic investigation and are beyond the scope of this short note. We thus conclude that the present result shows that unrestricted monotonicity cannot hold, even for the Gaussian kernel, and it identifies one robust geometric mechanism by which it fails.

2.3. Examples

To see that the condition in Theorem 1 is not void, we present some examples in one and two dimensions.

Example 1. Let $d = 1$, $W = (-1, 2)$, $y = 1$ (so that $R = 1$) and $m = 25$. Set $h_0 = 1/(2\log 25)^{1/2} \approx 0.39$. Then $a(h_0) = 1/25$, and

$$\frac{dm}{m + a(h_0)} + \frac{d}{1 + ma(h_0)} + \log a(h_0) \left\{ \frac{2ma(h_0)}{(m + a(h_0))^2} + \frac{2ma(h_0)}{(1 + ma(h_0))^2} \right\}$$

$$\approx 1.50 - 1.62 < 0.$$

Example 2. Let $d = 2$, $W = (-1, 2)^2$, $y = (1, 1)$ (so that $R = \sqrt{2}$) and $m = 1000$. Fig. 1 plots the Stoyan–Grabarnik statistic $T_\kappa(h; \psi, W)$ as a function of h for the limiting configuration $\psi = (0, \dots, 0, y)$ used in the construction (left) and for ψ a realization of a Poisson point process with intensity 1000 on $W = (-1, 2)^2$. The graph in the left-hand panel is not monotonically increasing; the derivative is negative for $h \approx 0.38$. The derivative at 0.38 is -0.45. The right-hand panel represents a representative Poisson realization, where the graph increases monotonically from zero to infinity.

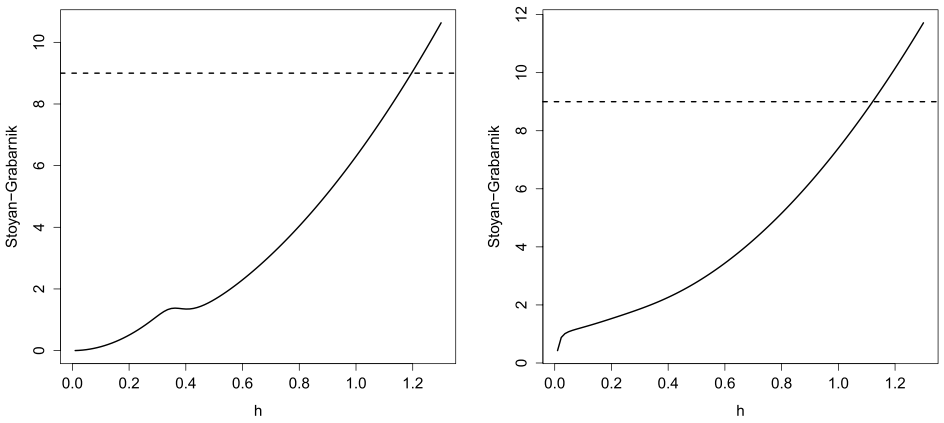


Fig. 1. The solid line is the graph of the function $h \to T_\kappa(h;\psi,W)$ for the limiting configuration $\psi = (0,\dots,0,y)$ used in the construction (left) and for ψ a realization of a Poisson point process with intensity 1000 on $W = (-1,2)^2$. The dotted line represents the cut-off point $|W| = 9$.

Funding

O. Cronie has been supported by the Swedish Research Council (2023-03320).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

Chiu, S.N., Stoyan, D., Kendall, W.S., Mecke, J., 2013. Stochastic geometry and its applications, third ed. Wiley.
Cronie, O., Van Lieshout, M.N.M., 2018. A non-model-based approach to bandwidth selection for kernel estimators of spatial intensity functions. Biometrika 105, 455–462.
Kresin, C., 2023. Applications and properties of point processes (Ph.D. thesis). University of California, Los Angeles.
Stoyan, D., Grabarnik, P., 1991. Second-order characteristics for stochastic structures connected with Gibbs point processes. Math. Nachr. 151, 95–100.
Wand, M.P., Jones, M.C., 1995. Kernel smoothing. Chapman and Hall.