Nearest-neighbour Markov point processes on graphs with Euclidean edges

Marie-Colette van Lieshout

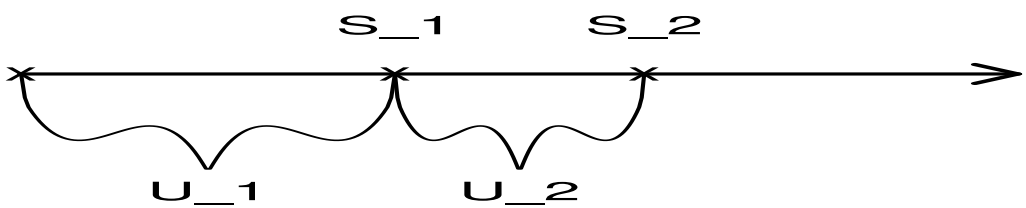
CWI and University of Twente
The Netherlands

Renewal process

Let U_i, $i = 1, 2, \ldots$, be i.i.d. **inter-arrival times** with p.d.f. π, c.d.f. F. Set $S_0 \equiv 0$ and define

$$S_i = S_{i-1} + U_i$$

for $i \in \mathbb{N}$. Those S_i that fall in $(0, T]$ form a **renewal process** X with time horizon T.



Stochastic intensity

The stochastic intensity

$$h^*(t)\,dt = \mathbb{P}(\text{ point in } dt | X \cap (0,t))$$

of a renewal process is

$$h^*(t) = h(V_{t-})$$

where V_t is the backward recurrence time at t, i.e. the time elapsed since the last event before or at time t, and

$$h(t) = \frac{\pi(t)}{1 - F(t)}$$

is the hazard rate of π.

Disadvantage:

- h^* is asymmetric in that it depends on the **past** of the process;
- h^* does not generalise easily to networks without chronology.

Probability distribution

The probability distribution of X (and of any **simple point process** on the half-line) can be specified by giving

- a (discrete) probability mass function q_n, $n \in \mathbb{N}_0$, for the number of points in $[0, T]$,
- for each $n \in \mathbb{N}$, a symmetric p.d.f. p_n for the locations of the points given there are n of them,

which combine in (un-normalised) **Janossy densities**

$$j_n(t_1, \ldots, t_n) = n! q_n p_n(t_1, \ldots, t_n)$$

for $n \in \mathbb{N}$, with $j_0 = q_0$.

Interpretation

Janossy densities

- are invariant under permutations of the t_i;
- $j_n(dt_1, \ldots, dt_n)$ is the probability of finding exactly n points in $[0, T]$, one each in $dt_1, \ldots, dt_n$.

Upon **normalisation**,

$$\begin{aligned} p(\{t_1, \ldots, t_n\}) &= e^T n! q_n p_n(t_1, \ldots, t_n) \\ &= e^T j_n(t_1, \ldots, t_n) \end{aligned}$$

is a probability density w.r.t. the law of the unit rate Poisson process on $[0, T]$.

Renewal process: Probability density

Order the points chronologically:

$$t_1 < \cdots < t_n.$$

Then

$$\begin{aligned} p(\{t_1, \ldots, t_n\}) &= e^T \exp\left[-\int_0^T h^*(t)\,dt\right] \prod_{i=1}^n h^*(t_i) \\ &= e^T (1 - F(T - t_n)) \prod_{i=1}^n \pi(t_i - t_{i-1}) \end{aligned}$$

w.r.t. the law of a unit rate Poisson process on $[0, T]$. By convention, an empty product is one and $t_0 = 0$.

Papangelou conditional intensity

Consider a point t such that

$$t_1 < \cdots < t_{i-1} < t < t_i < \cdots < t_n$$

and define

$$\lambda(t|\{t_1, \ldots, t_n\}) = \frac{p(\{t, t_1, \ldots, t_n\})}{p(\{t_1, \ldots, t_n\})}$$

provided $p(\{t_1, \ldots, t_n\}) > 0$.

For a **renewal process**,

$$\lambda(t|\{t_1, \ldots, t_n\}) = \frac{\pi(t - t_{i-1})\, \pi(t_i - t)}{\pi(t_i - t_{i-1})}$$

with suitable modifications at the end points.

Important observation: $\lambda(t|\{t_1, \ldots, t_n\})$ depends only on the **nearest neighbours** t_{i-1} and t_i of t.

CWI

Nearest-neighbour Markov point process

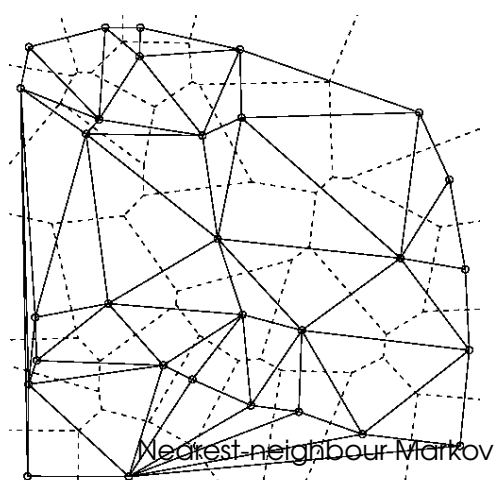
Baddeley and Møller (1989).

Point process X on $[0, T]$ with p.d.f. p is **nearest-neighbour Markov** if whenever $p(\{t_1, \ldots, t_n\}) > 0$,

(a) $p(\mathbf{y}) > 0$ for all $\mathbf{y} \subseteq \{t_1, \ldots, t_n\}$;

(b) for all $t \notin \{t_1, \ldots, t_n\}$, $\lambda(t|\{t_1, \ldots, t_n\})$ depends only on u, its nearest neighbours in $\{t_1, \ldots, t_n\}$ and their mutual nearest neighbour relations.

Note: in higher dimensions, points are nearest neighbours (**Delaunay neighbours**) if their Voronoi cells overlap.



Hammersley–Clifford theorem

A point process X with p.d.f. p is nearest-neighbour Markov iff $p(\{t_1, \ldots, t_n\})$ can be factorised up to normalisation as

$$\begin{cases} \prod_{i=1}^n \gamma(\{t_i\}) \prod_{i=2}^n \gamma(\{t_{i-1}, t_i\}) & \text{if } \gamma(\{t_i, t_j\}) > 0 \text{ for all distinct } t_i, t_j \\ 0 & \text{otherwise} \end{cases}$$

where $t_1 < \cdots < t_n$ and $\gamma \geq 0$ is an **interaction function**:
if $\gamma(\{t_1, \ldots, t_n\}) > 0$,

- $\gamma(\mathbf{y}) > 0$ for subsets $\mathbf{y}$ of $\{t_1, \ldots, t_n\}$;
- if $\gamma(\{t_{i-1}, t, t_i\}) > 0$, $t_{i-1} < t < t_i$, then also $\gamma(\{t, t_1, \ldots, t_n\}) > 0$.

Example: if $\pi > 0$, the renewal process is hereditary and

$$\begin{aligned} \gamma(\{t_i\}) &= \pi(t_i)(1 - F(T - t_i)), \\ \gamma(\{t_{i-1}, t_i\}) &= \frac{\pi(t_i - t_{i-1})}{\pi(t_i)(1 - F(T - t_{i-1}))}. \end{aligned}$$

CWI

Objective in this talk

Aim: develop analogues of **renewal processes** on networks, exploiting their one-dimensional nature.

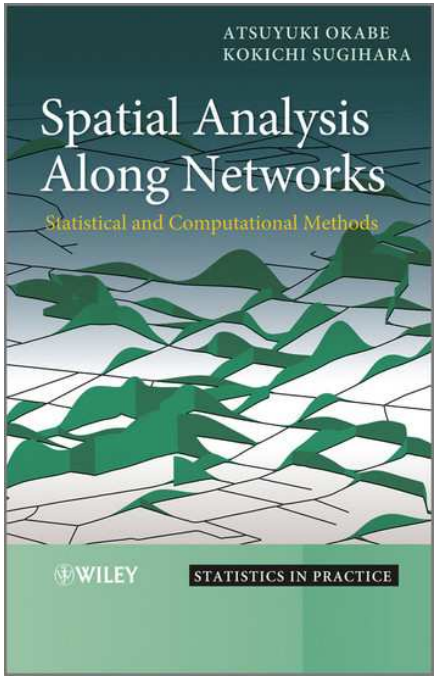
To do so, we study the **Delaunay relation** on networks, which define **nearest-neighbour Markov** point process in the sense of Baddeley and Møller (1989) by a probability density of the form

$$p(\mathbf{x}) \propto 1\{\gamma(\mathbf{x}) > 0\} \prod_{x_i \in \mathbf{x}} \gamma(\{x_i\}) \prod_{i<j : x_i \sim_{\mathbf{x}} x_j} \gamma(\{x_i, x_j\}),$$

where $\gamma \geq 0$ is an interaction function, at least on a tree. In the general case, a **local Delaunay relation** is proposed.

Networks

Okabe and Sugihara (2012).



Features:

- inherent network geometry;
- the network may not be closed under translations;
- focus on first and second order summary statistics, see Rakshit et al. (2017) and references therein.

Models on linear networks

Examples:

- Baddeley et al. (2017): Cox, Switzer-type and cell process.
- Anderes et al. (2016): expand the modelling framework in various directions:
  - a more general definition in terms of parametrised curves that may or may not overlap;
  - parametrisations naturally define a weighted shortest path distance;
  - investigation of alternative definitions of distance;
  - construction of log-Gaussian Cox processes in terms of a Gaussian process on the network specified by an isotropic covariance function.

CWI

Definition of network

Anderes, Møller and Rasmussen (2016).

A **graph with Euclidean edges** is a triple $G = (V, E = (e_i)_i, \Phi = (\phi_i)_i)$ s.t.

- (V, E) is a finite, simple connected graph (no loops/multiple edges);
- every edge $e_i = \{v_i^1, v_i^2\} \in E$, $v_i^1, v_i^2 \in V$, is parametrised by the inverse of a homeomorphism ϕ_i, and

$$\phi_i^{-1} : J_i \to \mathbb{R}^2$$

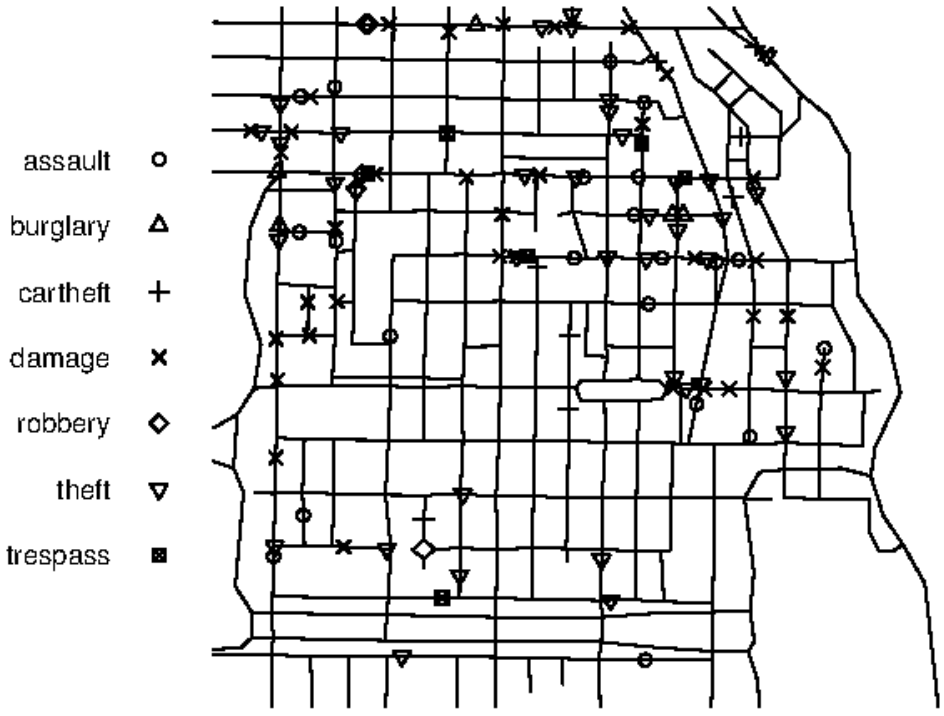
  is C^1 on an open interval $\emptyset \neq J_i \subseteq \mathbb{R}$ with endpoints $\phi_i(v_i^j)$, $j = 1, 2$, and has induced length $|\phi_i(v_i^1) - \phi_i(v_i^2)|$.

If the edge interiors are disjoint, define the **network**

$$L = V \cup \bigcup_{i=1}^{n(E)} \phi_i^{-1}(J_i)$$

equipped with weighted shortest path distance d_G.

Example: Chicago crime data



The Delaunay relation on a network

For a finite configuration $\mathbf{x} \subseteq L$ of distinct points define

$$x_i \sim_{\mathbf{x}} x_j \Leftrightarrow C(x_i|\mathbf{x}) \cap C(x_j|\mathbf{x}) \neq \emptyset$$

where

$$C(x_j|\mathbf{x}) = \{y \in L : d_G(y, x_j) \leq d_G(y, x) \text{ for all } x \in \mathbf{x}\}$$

is the Voronoi cell of x_j in $\mathbf{x}$.

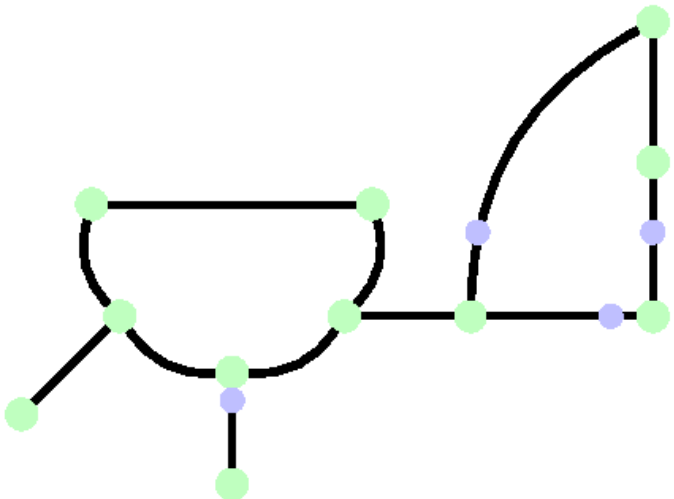
The $\mathbf{x}$–**neighbourhood** of a subset $\mathbf{y} \subseteq \mathbf{x}$ is defined as

$$N(\mathbf{y}|\mathbf{x}) = \{x \in \mathbf{x} : x \sim_{\mathbf{x}} y \text{ for some } y \in \mathbf{y}\}.$$

The configuration $\mathbf{y}$ is an $\mathbf{x}$-**clique** if for each $y_i, y_j \in \mathbf{y}$, $y_i \sim_{\mathbf{x}} y_j$.

Properties

- Voronoi cells are not necessarily line segments;
- points may have more than two neighbours;
- cliques need not be pairs.



The **clique indicator function** satisfies

- if $\chi(\mathbf{y}|\mathbf{x}) = 1$, then $\chi(\mathbf{y}|\mathbf{z}) = 1$ for all $\mathbf{y} \subseteq \mathbf{z} \subseteq \mathbf{x}$;
- if $\chi(\mathbf{y}|\mathbf{z}) = 0$, then $\chi(\mathbf{y}|\mathbf{x}) = 0$ for all $\mathbf{y} \subseteq \mathbf{z} \subseteq \mathbf{x}$.

The Delaunay relation on a tree

Definition

A graph with Euclidean edges (V, E, Φ) is a **tree** if the graph (V, E) is. Specifically, (V, E) has no **cycles**, i.e. no closed path $(v_0, v_1, \ldots, v_p, v_0)$, $v_i \in V$, over different edges having positive length ($p > 0$).

Lemma

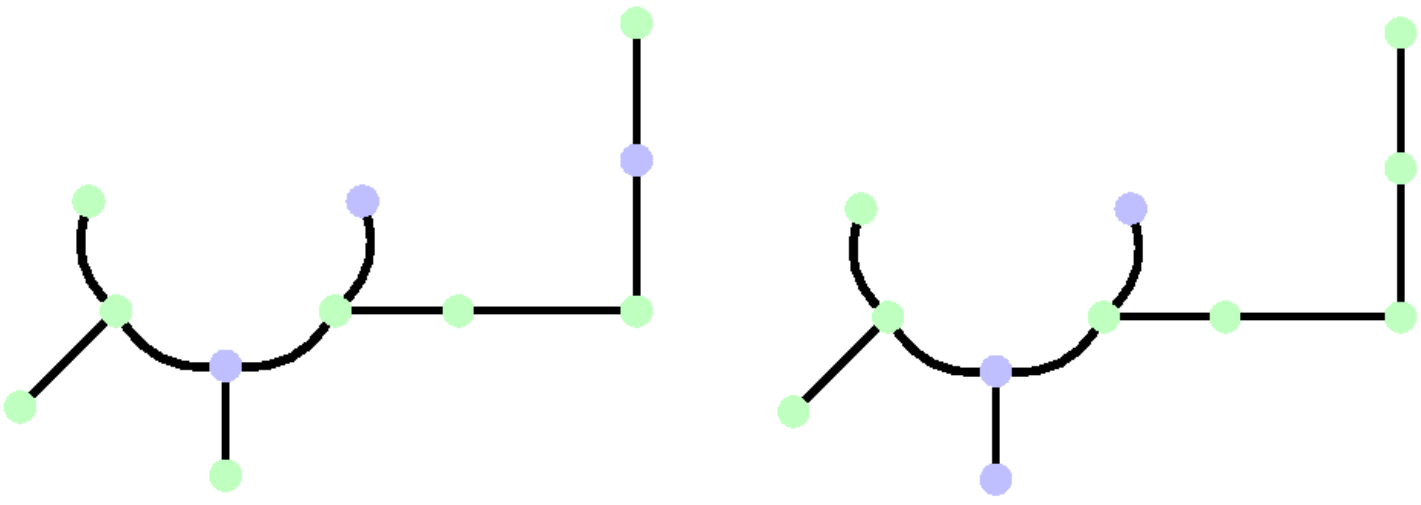
A graph with Euclidean edges is a tree iff there is exactly one path between any two points in its associated network L.

The Delaunay relation on a tree: Paths

Lemma

Let (V, E, Φ) be a tree, L its associated network and $\mathbf{y} = \{y_1, y_2, y_3\} \subseteq L$. Then there exist unique paths between the elements of $\mathbf{y}$ that

- either form a star with three stokes of strictly positive length emanating from a vertex $v \in L$,
- or combine into a single path.



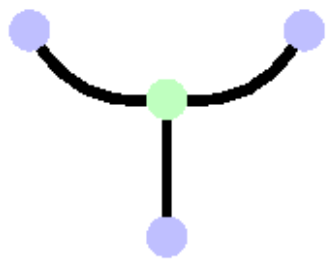
The Delaunay relation on a tree: Cliques

$\mathbf{y} \subseteq L$ is in **general position** if no three points lie on the boundary of the same d_G-ball.

Lemma

Clique sizes are at most two on the class of configurations in general position.

Moreover, for all $\mathbf{y} = \{y_1, y_2\} \subseteq \mathbf{x}$ with $\mathbf{x}$ in general position, $\chi(\mathbf{y}|\mathbf{x}) = 1$ iff the midpoint of $\mathbf{y}$ w.r.t. d_G along the unique path between y_1 and y_2 lies in $C(y_1|\mathbf{x}) \cap C(y_2|\mathbf{x})$.



The blue points form a clique iff the stokes are equally long.

Consistency conditions

Baddeley and Møller (1989).

For all finite point configurations $\mathbf{y} \subseteq \mathbf{z} \subseteq L$ and all $u, v \in L$ such that $u, v \notin \mathbf{z}$,

(C1) $\chi(\mathbf{y}|\mathbf{z}) \neq \chi(\mathbf{y}|\mathbf{z} \cup \{u\})$ implies $\mathbf{y} \subseteq N(\{u\}|\mathbf{z} \cup \{u\})$;

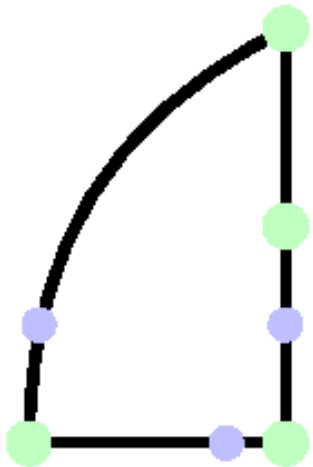
(C2) if $u \not\sim_{\mathbf{x}} v$ for $\mathbf{x} = \mathbf{z} \cup \{u, v\}$, then

$$\chi(\mathbf{y}|\mathbf{z} \cup \{u\}) + \chi(\mathbf{y}|\mathbf{z} \cup \{v\}) = \chi(\mathbf{y}|\mathbf{z}) + \chi(\mathbf{y}|\mathbf{x}).$$

Theorem

Let (V, E, Φ) be a graph with Euclidean edges that is a **tree** and L its associated network. Then the Delaunay relation satisfies (C1)–(C2) on the family of configurations **in general position**.

Counterexample



Let $\mathbf{z}$ contain the blue points. Then $\chi(\mathbf{z}|\mathbf{z}) = 1$. Adding any $u \notin \mathbf{z}$ would **separate** two points of $\mathbf{z}$ ($\chi(\mathbf{z}|\mathbf{z} \cup \{u\}) = 0$) while u is **no neighbour** of the third point z ($z \notin N(\{u\}|\mathbf{z} \cup \{u\})$).

Adding any v s.t. $u \not\sim_{\mathbf{z}\cup\{u,v\}} v$,

$$\chi(\mathbf{z}|\mathbf{z} \cup \{u, v\}) = \chi(\mathbf{z}|\mathbf{z} \cup \{u\}) = \chi(\mathbf{z}|\mathbf{z} \cup \{v\}) = 0$$

although $\chi(\mathbf{z}|\mathbf{z}) = 1$.

Hammersley–Clifford theorem for trees

Consider (V, E, Φ) for which (V, E) is a tree. Let $p \geq 0$ be a function of finite configurations in general position.
Then p is **Markov wrt the Delaunay relation** iff $p(\mathbf{x})$ is proportional to

$$\begin{cases} \prod_{x_i \in \mathbf{x}} \gamma(\{x_i\}) \prod_{i<j: x_i \sim_{\mathbf{x}} x_j} \gamma(\{x_i, x_j\}), & \text{if } \gamma(\{x_i, x_j\}) > 0 \text{ for all distinct } x_i, x_j \\ 0, & \text{otherwise} \end{cases}$$

for some function $\gamma \geq 0$ (extended to $\mathbf{x}$ with $n(\mathbf{x}) \geq 3$ by $\gamma(\mathbf{x}) = 1\{\gamma(\mathbf{y}) > 0 \, \forall \mathbf{y} \subseteq \mathbf{x}, n(\mathbf{y}) \leq 2\}$) that is an **interaction function**: if $\gamma(\mathbf{x}) > 0$, then

- $\gamma(\mathbf{y}) > 0$ for all $\mathbf{y} \subseteq \mathbf{x}$;
- if $\gamma(N(\{u\}|\mathbf{x} \cup \{u\})) > 0$ then $\gamma(\mathbf{x} \cup \{u\}) > 0$.

CWI

The local Delaunay relation

Define $\sim_E$ on L as follows:

- points on edges are neighbours if their edges share a common vertex;
- vertices are neighbours of points on their incident (closed) edges;

Write $\sim_{\mathbf{x}}^{i,j}$ for the Delaunay relation restricted to

$$L \cap \left(\phi_i^{-1}(\bar{J}_i) \cup \phi_j^{-1}(\bar{J}_j)\right).$$

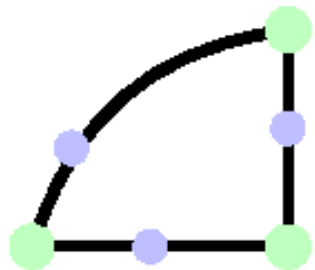
The **local Delaunay relation** $\sim_{\mathbf{z}}^{E}$ is defined as follows:

- for x, y on edges i, j that share a common vertex $x \sim_{\mathbf{z}}^{E} y$ iff $x \sim_{\mathbf{z}\cap\left(\phi_i^{-1}(\bar{J}_i)\cup\phi_j^{-1}(\bar{J}_j)\right)}^{i,j} y$;
- for x, y on the same (closed) edge e_i, $x \sim_{\mathbf{z}}^{E} y$ iff $x \sim_{\mathbf{z}\cap\left(\phi_i^{-1}(\bar{J}_i)\right)}^{i,i} y$.

Properties

Theorem

Let $G = (V, E, \Phi)$ be a graph with Euclidean edges and L its associated network. If G does not contain any triangles, then the local Delaunay relation $\sim_{\mathbf{x}}^{E}$ satisfies (C1)–(C2). By adding an artificial extra vertex in every triangle, (C1)–(C2) can be made to hold generally.



M.N.M. van Lieshout Nearest-neighbour Markov point processes on graphs with Euclidean edges. *Advances in Applied Probability*, **50:**1275–1293 (December 2018).