Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications ▮ (▮▮▮▮) ▮▮▮–▮▮▮
stochastic processes and their applications
www.elsevier.com/locate/spa

Bounds to the normal for proximity region graphs

Larry Goldstein[a], Tobias Johnson[a], Raphaël Lachièze-Rey[b,*]

[a] University of Southern California, United States
[b] Université Paris Descartes, Sorbonne Paris Cité, France

Received 7 July 2016; received in revised form 10 April 2017; accepted 5 July 2017
Available online xxxx

Abstract

In a proximity region graph $\mathcal{G}$ in $\mathbb{R}^d$, two distinct points x, y of a point process μ are connected when the 'forbidden region' $S(x, y)$ these points determine has empty intersection with μ. The Gabriel graph, where $S(x, y)$ is the open disk with diameter the line segment connecting x and y, is one canonical example. When μ is a Poisson or binomial process, under broad conditions on the regions $S(x, y)$, bounds on the Kolmogorov and Wasserstein distances to the normal are produced for functionals of $\mathcal{G}$, including the total number of edges and the total length. Variance lower bounds, not requiring strong stabilization, are also proven to hold for a class of such functionals.

MSC: 60D05; 52A22; 60F05
Keywords: Forbidden region graph; Berry–Esseen bounds; Stabilization; Poisson functionals

1. Introduction

The family of graphs that we study here, all with vertex sets given by a locally finite point process μ in $\mathbb{R}^d$, is motivated by two canonical examples considered in [1], the Gabriel graph and the relative neighborhood graph. Two distinct points x and y of μ are connected by an edge in the Gabriel graph if and only if there does not exist any point z of the process μ lying in the open disk whose diameter is the line segment connecting x and y. The relative neighborhood graph has an edge between x and y if and only if there does not exist a point z of μ such that

$$\max(\|x - z\|, \|z - y\|) < \|x - y\|,$$

* Corresponding author.
E-mail address: raphael.lachieze-rey@parisdescartes.fr (R. Lachièze-Rey).

http://dx.doi.org/10.1016/j.spa.2017.07.002

that is, if and only if there is no point z of μ that is closer to either x or y than these points are to each other.

These two examples are special cases of 'proximity graphs' as defined in [3], where distinct points x and y of μ are connected if and only if a region $S(x, y)$ determined by x and y contains no points of μ, that is, when $\mu \cap S(x, y) = \emptyset$. As $S(x, y)$ must be free of points of μ in order for x and y to be joined, we call $S(x, y)$ the 'forbidden region' determined by x and y. In particular, with $B(x, r)$ and $B^o(x, r)$ denoting the closed and open ball of radius r centered at x, respectively, the forbidden regions of the Gabriel graph are given by

$$S(x, y) = B^o((x + y)/2, \|y - x\|/2), \tag{1}$$

and those of the relative neighborhood graph by

$$S(x, y) = B^o(x, \|y - x\|) \cap B^o(y, \|x - y\|). \tag{2}$$

It is easy to check that the forbidden regions $S(x, y)$ of the Gabriel graph are contained in those of the relative neighbor graph, and hence edges of latter are also edges of former.

We refer to the graphs formed in this manner also as 'forbidden region graphs'. Indeed, when coining the label 'proximity graphs' in [3], one reads that 'this term could be misleading in some cases'. Indeed, forbidden region graphs may depend on 'non-proximate' information, such as the graph considered in Example 5 of [3], whose forbidden region $S(x, y)$ is the infinite strip bounded by the two parallel hyperplanes containing x and y, each perpendicular to $y - x$. Allowing forbidden regions to depend on larger sets of points and to be determined by more complex rules yield well studied graphs with additional structure, including the Minimum Spanning Tree and the Delaunay triangulation, see [1].

For a forbidden region graph $\mathcal{G}$ and a Poisson or binomial point process μ in some bounded measurable 'viewing window' denoted $\mathbb{X}$ in the sequel, ensuring that the graph and functional $L(\mu)$ in (3) are finite, we study the distribution of

$$L(\mu) = \sum_{\{x,y\} \subseteq \mu, x \neq y} \mathbf{1}(\mu \cap S(x, y) = \emptyset)\psi(x, y), \tag{3}$$

for some $\psi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ satisfying $\psi(x, y) = \psi(y, x)$. For instance, taking $\psi(x, y) = \|x - y\|^\alpha$ for some $\alpha \geq 0$, for $\alpha = 0$ and $\alpha = 1$ the value of $L(\mu)$ is the number of edges and the total length of $\mathcal{G}$, respectively.

Recall that the Kolmogorov distance between random variables U and V is defined as

$$d_K(U, V) = \sup_{t \in \mathbb{R}} |\mathbb{P}(U \leqslant t) - \mathbb{P}(V \leqslant t)|,$$

and the Wasserstein distance as

$$d_W(U, V) = \sup_{h \in \mathrm{Lip}_1} |\mathbb{E}[h(U) - h(V)]|,$$

where Lip_1 stands for the class of 1-Lipschitz functions $\mathbb{R} \to \mathbb{R}$. Theorem 2, our main result, is a bound on the normal approximation of L in $d(\cdot, \cdot)$, denoting either the Wasserstein or Kolmogorov metric, that holds under broad conditions on the forbidden regions and underlying point process. Its immediate corollary, in conjunction with the variance lower bound of Theorem 4, provides the following result for the two motivating examples just introduced.

Corollary 1. *Let $\mathbb{X} = B(0, 1)$, and suppose that η_t is either a Poisson process with intensity t on $\mathbb{X}$, or a binomial process of t independent and uniformly distributed points on $\mathbb{X}$, and let*