Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 126 (2016) 1306–1330

stochastic processes and their applications

www.elsevier.com/locate/spa

Concentration for Poisson functionals: Component counts in random geometric graphs

Sascha Bachmann

Institute for Mathematics, Osnabrück University, 49069 Osnabrück, Germany

Received 1 July 2015; received in revised form 6 November 2015; accepted 11 November 2015
Available online 17 November 2015

Abstract

Upper bounds for the probabilities $\mathbb{P}(F \geq \mathbb{E}F + r)$ and $\mathbb{P}(F \leq \mathbb{E}F - r)$ are proved, where F is a certain component count associated with a random geometric graph built over a Poisson point process on $\mathbb{R}^d$. The bounds for the upper tail decay exponentially, and the lower tail estimates even have a Gaussian decay.

For the proof of the concentration inequalities, recently developed methods based on logarithmic Sobolev inequalities are used and enhanced. A particular advantage of this approach is that the resulting inequalities even apply in settings where the underlying Poisson process has infinite intensity measure.

MSC: primary 60D05; secondary 05C80; 60C05

Keywords: Random graphs; Component counts; Concentration inequalities; Logarithmic Sobolev inequalities; Poisson point process

1. Introduction

Random geometric graphs have been studied extensively for some decades now. In the simplest version of these graphs, the vertices are given by a random set of points in $\mathbb{R}^d$ and two vertices are connected by an edge if their distance is less than a fixed positive real number. This model was introduced by E. N. Gilbert in [14], and since then many authors contributed

E-mail address: sascha.bachmann@uni-osnabrueck.de.

http://dx.doi.org/10.1016/j.spa.2015.11.004
0304-4149/

to various directions of research on random geometric graphs. For a historical overview on the topic we refer the reader to the book [25] by M. D. Penrose. Recent contributions are e.g. [11,20,21,26].

It is a well established fact that numerous real world phenomena can be modeled by means of a random geometric graph, like for example the spread of a disease or a fire (see e.g. [4,13]). Also, as communication networks such as wireless and sensor networks have become increasingly important in recent years, random geometric graphs have gained a considerable attention — since they provide natural models for these objects (see e.g. [10,15,24]).

Further applications arise from cluster analysis, where one aims to divide a given set of objects into groups (or clusters) such that objects within the same group are similar to each other (see e.g. [6,7] for further reading). If the objects are represented by points in $\mathbb{R}^d$, one way to perform this task is to built a geometric graph over the points and to take the connected components of the graph as the clusters. At this, a connected component of a graph G with vertex set V is an induced connected subgraph H of G with vertex set $V' \subseteq V$ such that for any $x \in V'$ and $y \in V \setminus V'$ there is no edge between x and y. For the purpose of statistical inference, a probabilistic theory for the connected components of the graph is needed.

Throughout the present work, the vertices of the considered random geometric graphs are given by a Poisson point process on $\mathbb{R}^d$. The class of random variables that is investigated in this paper includes a variety of quantities that are typically of interest in several of the applications described above. For example, one can consider the number of connected components of the graph with at most k (or alternatively with exactly k) vertices. Further random variables that are covered by our analysis are obtained by counting the number of components that are isomorphic to a fixed connected graph H. Early work on the latter quantities was done by R. Hafner in [16] and further related results are presented in [25].

The main contribution of the present paper is to establish new exponential upper bounds for the probabilities $\mathbb{P}(F \geq \mathbb{E}F + r)$ and $\mathbb{P}(F \leq \mathbb{E}F - r)$, where $\mathbb{E}F$ denotes the expectation of a component count F and $r > 0$ is a real number. Inequalities of this type are usually called *concentration inequalities*. In order to derive our estimates, we use and enhance a method that was recently developed by S. Bachmann and G. Peccati in [2]. The latter paper provides several refinements of a method for proving tail estimates for Poisson functionals (also known as the entropy-method), which is based on (modified) logarithmic Sobolev inequalities, and which was particularly studied in the seminal work by Wu [28], extending previous findings by Ané, Bobkov and Ledoux [1,5]. Combining Wu's modified logarithmic Sobolev inequality with the famous Mecke formula for Poisson processes, the authors of [2] were able to adapt concentration techniques for product space functionals, which were particularly developed by Boucheron, Lugosi and Massart [8], and also by Maurer [22], to the setting of Poisson processes. This approach adds a lot of flexibility to the entropy-method, and a remarkable feature of the obtained techniques is that they allow to deal with functionals build over Poisson processes with infinite intensity measure.

First applications for these techniques are worked out in [2] and also in [3], where concentration bounds for certain Poisson U-statistics with positive kernels are established. A crucial property that was exploited in the latter investigations is that adding a point to the Poisson process cannot decrease the value of the considered functionals. In principle, this monotonicity is not needed for the method suggested in [2] to be applied. However, due to somewhat more complicated objects that need to be controlled when dealing with non-monotonic functionals, the method has only been successfully used for monotonic quantities so far. Clearly, the component counts that are studied in the present paper are not monotonic. So, a particularly interesting aspect