Crossing Numbers and Stress of Random Graphs

Markus Chimani[1(✉)] , Hanna Döring[2], and Matthias Reitzner[2]

[1] Theoretical CS, Institute of Computer Science, Uni Osnabrück,
Osnabrück, Germany
markus.chimani@uni-osnabrueck.de
[2] Stochastics, Institute of Mathematics, Uni Osnabrück, Osnabrück, Germany
{hanna.doering,matthias.reitzner}@uni-osnabrueck.de

Abstract. Consider a random geometric graph over a random point process in $\mathbb{R}^d$. Two points are connected by an edge if and only if their distance is bounded by a prescribed distance parameter. We show that projecting the graph onto a two dimensional plane is expected to yield a constant-factor crossing number (and rectilinear crossing number) approximation. We also show that the crossing number is positively correlated to the stress of the graph's projection.

1 Introduction

An undirected abstract graph G_0 consists of vertices and edges connecting vertex pairs. An *injection* of G_0 into $\mathbb{R}^d$ is an injective map from the vertices of G_0 to $\mathbb{R}^d$, and edges onto curves between their corresponding end points but not containing any other vertex point. For $d \geq 3$, we may assume that distinct edges do not share any point (other than a common end point). For $d = 2$, we call the injection a *drawing*, and it may be necessary to have points where curves *cross*. A drawing is *good* if no pair of edges crosses more than once, nor meets tangentially, and no three edges share the same crossing point. Given a drawing D, we define its crossing number $\mathrm{cr}(D)$ as the number points where edges cross. The crossing number $\mathrm{cr}(G_0)$ of the graph itself is the smallest $\mathrm{cr}(D)$ over all its good drawings D. We may restrict our attention to the *rectilinear crossing number* $\overline{\mathrm{cr}}(G_0)$, where edge curves are straight lines; note that $\overline{\mathrm{cr}}(G_0) \geq \mathrm{cr}(G_0)$.

The crossing number and its variants have been studied for several decades, see, e.g., [30], but still many questions are widely open. We know the crossing numbers only for very few graph classes; already for $\mathrm{cr}(K_n)$, i.e., on complete graphs with n vertices, we only have conjectures, and for $\overline{\mathrm{cr}}(K_n)$ not even them. Since deciding $\mathrm{cr}(G_0)$ is NP-complete [15] (and $\overline{\mathrm{cr}}$ even ∃ℝ-complete [4]), several attempts for approximation algorithms have been undertaken. The problem does not allow a PTAS unless P = NP [6]. For general graphs, we currently do not know whether there is an α-approximation for any constant α. However, we can achieve constant ratios for dense graphs [14] and for bounded pathwidth graphs [3]. Other strong algorithms deal with graphs of maximum bounded

degree and achieve either slightly sublinear ratios [13], or constant ratios for further restrictions such as embeddability on low-genus surfaces [16–18] or a bounded number of graph elements to remove to obtain planarity [7,9,10,12].

We will make use of the *crossing lemma*, originally due to [2,25][1]: There are constants[2] $d \geq 4, c \geq \frac{1}{64}$ such that any abstract graph G_0 on n vertices and $m \geq dn$ edges has $\mathrm{cr}(G_0) \geq cm^3/n^2$. In particular for (dense) graphs with $m = \Theta(n^2)$, this yields the asymptotically tight maximum of $\Theta(m^2)$ crossings.

Random Geometric Graphs (RGGs). We always consider a *geometric* graph G as input, i.e., an abstract graph G_0 together with a straight-line injection into $\mathbb{R}^d$, for some $d \geq 2$; we identify the vertices with their points. For a 2-dimensional plane L, the postfix operator $|_L$ denotes the projection onto L.

Given a set of points V in $\mathbb{R}^d$, the *unit-ball graph* (*unit-disk graph* if $d = 2$) is the geometric graph using V as vertices that has an edge between two points iff balls of radius 1 centered at these points touch or overlap. Thus, points are adjacent iff their distance is ≤ 2. In general, we may use arbitrary threshold distances $\delta > 0$. We are interested in *random* geometric graphs *(RGGs)*, i.e., when using a Poisson point process to obtain V for the above graph class 2.

Stress. When drawing (in particular large) graphs with straight lines in practice, *stress* is a well-known and successful concept, see, e.g., [5,20,21]: let G be a geometric graph, d_0, d_1 two distance functions on vertex pairs—(at least) the latter of which depends on an injection—and w weights. We have:

$$\mathrm{stress}(G) := \sum_{v_1, v_2 \in V(G), v_1 \neq v_2} w(v_1, v_2) \cdot (d_0(v_1, v_2) - d_1(v_1, v_2))^2. \qquad (1)$$

In a typical scenario, G is injected into $\mathbb{R}^2$, d_0 encodes the graph-theoretic distances (number of edges on the shortest path) or some given similarity matrix, and d_1 is the Euclidean distance in $\mathbb{R}^2$. Intuitively, in a drawing of 0 (or low) stress, the vertices' geometric distances d_1 are (nearly) identical to their "desired" distance according to d_0. A typical weight function $w(v_1, v_2) := d_0(v_1, v_2)^{-2}$ softens the effect of "bad" geometric injections for vertices that are far away from each other anyhow. It has been observed *empirically* that low-stress drawings *tend* to be visually pleasing and to have a low number of crossings, see, e.g., [8,22]. While it may seem worthwhile to approximate the crossing number by minimizing a drawing's stress, there is no sound mathematical basis for this approach.

There are different ways to find (close to) minimal-stress drawings in 2D [5]. One way is multidimensional scaling, cf. [20], where we start with an injection of an abstract graph G_0 into some high-dimensional space $\mathbb{R}^d$ and asking for a projection of it onto $\mathbb{R}^2$ with minimal stress. It should be understood that Euclidean distances in a unit-ball graph in $\mathbb{R}^d$ by construction closely correspond to the graph-theoretic distances. In fact, for such graphs it seems reasonable to

---

[1] Incidentally, the lemma allows an intriguingly elegant proof using stochastics [1].
[2] The currently best constants $d = 7, c = \frac{1}{20}$ are due to [19].

use the distances in $\mathbb{R}^d$ as the given metrics d_0, and seek an injection into $\mathbb{R}^2$—whose resulting distances form d_1—by means of projection.

Contribution. We consider RGGs for large t and investigate the mean, variance, and corresponding law of large numbers both for their rectilinear crossing number and their minimal stress when projecting them onto the plane. We also prove, for the first time, a positive correlation between these two measures.

While our technical proofs make heavy use of stochastic machinery (several details of which have to be deferred to the arXiv version [11]), the consequences are very algorithmic: We give a surprisingly simple algorithm that yields an *expected constant* approximation ratio for random geometric graphs even in the pure abstract setting. In fact, we can state the algorithm already now; the remainder of this paper deals with the proof of its properties and correctness:

Given a random geometric graph G in $\mathbb{R}^d$ (see below for details), we pick a *random* 2-dimensional plane L in $\mathbb{R}^d$ to obtain a straight-line drawing $G|_L$ that yields a crossing number approximation both for $\overline{\mathrm{cr}}(G_0)$ and for $\mathrm{cr}(G_0)$.

Throughout this paper, we prefer to work within the setting of a Poisson point process because of the strong mathematical tools from the Malliavin calculus that are available in this case. It is straightforward to de-Poissonize our results: this yields asymptotically the same results—even with the same constants—for n uniform random points instead of a Poisson point process; we omit the details.

2   Notations and Tools from Stochastic Geometry

Let $W \subset \mathbb{R}^d$ be a convex set of volume $\mathrm{vol}_d(W) = 1$. Choose a Poisson distributed random variable n with parameter t, i.e., $\mathbb{E}n = t$. Next choose n points $V = \{v_1, \ldots, v_n\}$ independently in W according to the uniform distribution. Those points form a Poisson point process V in W of intensity t. A Poisson point process has several nice properties, e.g., for disjoint subsets $A, B \subset W$, the sets $V \cap A$ and $V \cap B$ are independent (thus also their size is independent). Let $V_{\neq}^k$, $k \geq 1$, be the set of all ordered k-tuples over V with pairwise distinct elements. We will consider V as the vertex set of a geometric graph G for the distances parameter $(\delta_t)_{t>0}$ with edges $E = \{\{u,v\} \mid u, v \in V, u \neq v, \|u-v\| \leq \delta_t\}$, i.e., we have an edge between two distinct points if and only if their distance is at most δ_t. Such *random geometric graphs (RGG)* have been extensively investigated, see, e.g., [27,29], but nothing is known about the stress or crossing number of its underlying abstract graph G_0.

A *U-statistic* $U(k, f) := \sum_{\mathbf{v} \in V_{\neq}^k} f(\mathbf{v})$ is the sum over $f(\mathbf{v})$ for all k-tuples $\mathbf{v}$. Here, f is a measurable non-negative real-valued function, and $f(\mathbf{v})$ only depends on $\mathbf{v}$ and is independent of the rest of V. The number of edges in G is a U-statistic as $m = \frac{1}{2} \sum_{v,u \in V, v \neq u} \mathbb{1}(\|v-u\| \leq \delta_t)$. Likewise, the stress of a geometric graph as well as the crossing number of a straight-line drawing is a U-statistic, using 2- and 4-tuples of V, respectively. The well-known multivariate Slivnyak-Mecke formula tells us how to compute the expectation $\mathbb{E}_V$ over all realizations of the Poisson process V; for U-statistics we have, see [31, Cor. 3.2.3]:

$$\mathbb{E}_V \sum_{(v_1,\ldots,v_k)\in V_{\neq}^k} f(v_1,\ldots,v_k) = t^k \int_{W^k} f(v_1,\ldots,v_k)\, dv_1 \cdots dv_k. \qquad (2)$$

We already know $\mathbb{E}_V n = \mathbb{E}_V |V| = t$. Solving the above formula for the expected number of edges, we obtain

$$\mathbb{E}_V m = \mathbb{E}_V |E| = \frac{\kappa_d}{2}\, t^2 \delta_t^d + O(t^2 \delta_t^{d+1}\, \mathrm{surf}(W)), \qquad (3)$$

where $\kappa_d = \mathrm{vol}_d(B_d)$ is the volume of the unit ball B_d in $\mathbb{R}^d$, and $\mathrm{surf}(W)$ the surface area of W. For n and m, central limit theorems and concentration inequalities are well known as $t \to \infty$, see, e.g., [27,29].

The expected degree $\mathbb{E}_V \deg(v)$ of a typical vertex v is approximately of order $\kappa_d\, t\, \delta_t^d$ (this can be made precise using Palm distributions). This naturally leads to three different asymptotic *regimes* as introduced in Penrose's book [27]:

– in the *sparse regime* we have $\lim_{t\to\infty} t\, \delta_t^d = 0$, thus $\mathbb{E}_V \deg(v)$ tends to zero;
– in the *thermodynamic regime* we have $\lim_{t\to\infty} t\, \delta_t^d = c > 0$, thus $\mathbb{E}_V \deg(v)$ is asymptotically constant;
– in the *dense regime* we have $\lim_{t\to\infty} t\, \delta_t^d = \infty$, thus $\mathbb{E}_V \deg(v) \to \infty$.

Observe that in standard graph theoretic terms, the *thermodynamic regime* leads to *sparse graphs*, i.e., via (3) we obtain $\mathbb{E}_V m = \Theta(t) = \Theta(\mathbb{E}_V n)$. Similarly, the dense regime—together with $\delta_t \to c$—leads to *dense graphs*, i.e., $\mathbb{E}_V m = \Theta(t^2) = \Theta((\mathbb{E}_V n)^2)$. Recall that to employ the crossing lemma, we want $m \geq 4n$. Also, the lemma already shows that any good (straight-line) drawing of a dense graph G_0 already gives a constant-factor approximation for $\mathrm{cr}(G_0)$ (and $\overline{\mathrm{cr}}(G_0)$). In the following we thus assume a constant $0 < c \leq t\, \delta_t^d$ and $\delta_t \to 0$, i.e., $m = o(n^2)$.

The Slivnyak-Mecke formula is a classical tool to compute expectations and will thus be used extensively throughout this paper. Yet, suitable tools to compute variances came up only recently. They emerged in connection with the development of the Malliavin calculus for Poisson point processes [23,26]. An important operator for functions $g(V)$ of Poisson point processes is the *difference* (also called *add-one-cost*) operator,

$$D_v g(V) := g(V \cup \{v\}) - g(V),$$

which considers the change in the function value when adding a single further point v. We know that there is a Poincaré inequality for Poisson functionals [23, 32], yielding the upper bound in (4) below. On the other hand, the isometry property of the Wiener-Itô chaos expansion [24] of an (square integrable) L^2-function $g(V)$ leads to the lower bound in (4):

$$t \int_W (\mathbb{E}_V D_v g(V))^2\, dv \quad \leq \quad \mathbb{V}\mathrm{ar}_V\, g(V) \quad \leq \quad t \int_W \mathbb{E}_V (D_v g(V))^2\, dv. \qquad (4)$$

Often, in particular in the cases we are interested in in this paper, the bounds are sharp in the order of t and often even sharp in the occurring constant. This is due to the fact that the Wiener-Itô chaos expansion, the Poincaré inequality, and the lower bound are particularly well-behaved for Poisson U-statistics [28].

3 Rectilinear Crossing Number of an RGG

Let $\mathcal{L}$ be the set of all two-dimensional linear planes and $L \in \mathcal{L}$ be a random plane chosen according to a (uniform) Haar probability measure on $\mathcal{L}$. The drawing $G_L := G|_L$ is the projection of G onto L. Let $[u, v]$ denote the segment between vertex points $u, v \in V$ if their distance is at most δ_t and $\emptyset$ otherwise. The rectilinear crossing number of G_L is a U-statistic of order 4:

$$\overline{\mathrm{cr}}(G_L) = \frac{1}{8} \sum_{(v_1, v_2, v_3, v_4) \in V_{\neq}^4} \mathbb{1}([v_1, v_2]|_L \cap [v_3, v_4]|_L \neq \emptyset).$$

Keep in mind that even for the best possible projection we only obtain $\min_{L \in \mathcal{L}} \overline{\mathrm{cr}}(G|_L) \geq \overline{\mathrm{cr}}(G_0)$. To analyze $\mathbb{E}_V \min_{L \in \mathcal{L}} \overline{\mathrm{cr}}(G|_L)$ is more complicated than $\mathbb{E}_{L,V} \overline{\mathrm{cr}}(G|_L)$; fortunately, we will not require it.

3.1 The Expectation of the Rectilinear Crossing Numbers

For the expectation with respect to the underlying Poisson point process the Slivnyak-Mecke formula (2) gives

$$\mathbb{E}_V \overline{\mathrm{cr}}(G_L) = \frac{1}{8} t^4 \int_W \underbrace{\int_{W^3} \mathbb{1}([v_1, v_2]|_L \cap [v_3, v_4]|_L \neq \emptyset) \, dv_4 dv_3 dv_2}_{=:I_W(v_1)} \, dv_1.$$

Let c_d be the constant given by the expectation of the event that two independent edges cross. In this paper's arXiv version [11, Appendix A], we prove in Proposition 15 that $c_d \leq 2\pi\kappa_d^2$, that $\frac{I_W(v_1)}{\delta_t^{2d+2}}$ is bounded by c_d times the volume of the maximal $(d-2)$-dimensional section of W, and that

$$\lim_{\delta_t \to 0} \frac{I_W(v_1)}{\delta_t^{2d+2}} = c_d \mathrm{vol}_{d-2}((v_1 + L^{\perp}) \cap W), \tag{5}$$

where $L^{\perp}$ is the $d-2$ dimensional hyperplane perpendicular to L. Using the dominated convergence theorem of Lebesgue and Fubini's theorem we obtain

$$\lim_{t \to \infty} \frac{\mathbb{E}_V \overline{\mathrm{cr}}(G_L)}{t^4 \delta_t^{2d+2}} = \frac{1}{8} c_d \int_W \mathrm{vol}_{d-2}((v_1 + L^{\perp}) \cap W) \, dv_1$$

$$= \frac{1}{8} c_d \int_{W|_L} \int_{(v_1^L + L^{\perp}) \cap W} \mathrm{vol}_{d-2}((v_1^L + L^{\perp}) \cap W) \, dv_1^{L^{\perp}} dv_1^L$$

$$= \frac{1}{8} c_d \underbrace{\int_{W|_L} \mathrm{vol}_{d-2}((v_1^L + L^{\perp}) \cap W)^2 \, dv_1^L}_{=:I^{(2)}(W, L)}.$$

Theorem 1. *Let G_L be the projection of an RGG onto a two-dimensional plane L. Then, as $t \to \infty$ and $\delta_t \to 0$,*

$$\mathbb{E}_V \overline{\mathrm{cr}}(G_L) = \frac{1}{8} c_d\, t^4 \delta_t^{2d+2}\, I^{(2)}(W, L) + o(\delta_t^{2d+2} t^4).$$

For unit-disk graphs, i.e., $d = 2$, the choice of L is unique and the projection superfluous. There the expected crossing number is asymptotically $\frac{c_2}{8} t^4 \delta_t^6$ and thus of order $\Theta(m^3/n^2)$ which is asymptotically optimal as witnessed by the crossing lemma. In general, the expectation is of order

$$t^4 \delta_t^{2d+2} = \Theta\left(\frac{m^3}{n^2} \left(\frac{m}{n^2}\right)^{\frac{2-d}{d}}\right).$$

The extra factor m/n^2 can be understood as the probability that two vertices are connected via an edge, thus measures the "density" of the graph.

3.2 The Variance of the Rectilinear Crossing Numbers

By the variance inequalities (4) for functionals of Poisson point processes we are interested in the moments of the difference operator of the crossing numbers:

$$\mathbb{E}_V D_v \overline{\mathrm{cr}}(G_L) = \frac{1}{8} \mathbb{E}_V \sum_{(v_2,\ldots,v_4) \in V_{\neq}^3} \mathbb{1}([v, v_2]|_L \cap [v_3, v_4]|_L \neq \emptyset) = \frac{1}{8} t^3 I_W(v) \quad (6)$$

$$\mathbb{E}_V (D_v \overline{\mathrm{cr}}(G_L))^2 = \mathbb{E}_V \left(\frac{1}{8} \sum_{(v_2,\ldots,v_4) \in V_{\neq}^3} \mathbb{1}([v, v_2]|_L \cap [v_3, v_4]|_L \neq \emptyset)\right)^2 \quad (7)$$

Plugging (7) into the Poincaré inequality (4) gives

$$\mathbb{Var}_V \overline{\mathrm{cr}}(G_L) \leq \frac{1}{64} t \int_W \mathbb{E}_V \left(\sum_{(v_2,\ldots,v_4) \in V_{\neq}^3} \mathbb{1}([v, v_2]|_L \cap [v_3, v_4]|_L \neq \emptyset)\right)^2 dv.$$

Using calculations from integral geometry (see this paper's arXiv version [11, Appendix B]), there is a constant $0 < c_d' \leq 2\pi\kappa_d c_d$ (given by the expectation of the event that two pairs of independent edges cross) such that

$$\mathbb{Var}_V \overline{\mathrm{cr}}(G_L) \leq \frac{1}{64}\left(c_d^2 + \frac{c_d'}{t\delta_t^d}\right) t^7 \delta_t^{4d+4} \int_W \mathrm{vol}_{d-2}((v + L^\perp) \cap W)^2 (1 + o(1)) dv$$

$$+ O(\max\{t^6 \delta_t^{4d+2},\, t^5 \delta_t^{3d+2},\, t^4 \delta_t^{2d+2}\}).$$

We use that $t\delta_t^d \geq c > 0$, assume $d \geq 3$, and use Fubini's theorem again.

$$\lim_{t \to \infty} \frac{\mathbb{Var}_V \overline{\mathrm{cr}}(G_L)}{t^7 \delta_t^{4d+4}} \leq \frac{1}{64}\left(c_d^2 + c_d' \lim_{t \to \infty} \frac{1}{t\delta_t^d}\right) \underbrace{\int_{W|_L} \mathrm{vol}_{d-2}((v + L^\perp) \cap W)^3\, dv}_{=:I^{(3)}(W,L)}.$$

On the other hand, (6) and the lower bound in (4) gives in our case

$$\mathbb{V}\mathrm{ar}_V \overline{\mathrm{cr}}(G_L) \geq t \int_W (\mathbb{E}_V D_v \overline{\mathrm{cr}}(G_L))^2 \, dv$$

$$\geq \frac{1}{64} t^7 \int_W I_W(v)^2 \, dv = \frac{1}{64} c_d^2 t^7 \delta_t^{4d+4} I^{(3)}(W, L)(1 + o(1)).$$

Thus our bounds have the correct order and, in the dense regime where $t\delta_t^d \to \infty$, are even sharp. Using $0 < c_d' \leq 2\pi\kappa_d c_d$ we obtain:

Theorem 2. *Let G_L be the projection of an RGG in $\mathbb{R}^d$, $d \geq 3$, onto a two-dimensional plane L. Then, as $t \to \infty$ and $\delta_t \to 0$,*

$$\frac{1}{64} c_d^2 I^{(3)}(W, L) \leq \lim_{t \to \infty} \frac{\mathbb{V}\mathrm{ar}_V \overline{\mathrm{cr}}(G_L)}{t^7 \delta_t^{4d+4}} \leq \frac{1}{64} \Big(c_d^2 + 2\pi\kappa_d c_d \lim_{t \to \infty} \frac{1}{t\delta_t^d}\Big) I^{(3)}(W, L).$$

Theorems 1 and 2 show for the standard deviation

$$\sigma(\overline{\mathrm{cr}}(G_L)) = \sqrt{\mathbb{V}\mathrm{ar}_V \overline{\mathrm{cr}}(G_L)} = \Theta(t^4 \delta_t^{2d+2} \, t^{-\frac{1}{2}}) = \Theta(\mathbb{E}_V \overline{\mathrm{cr}}(G_L) \, (\mathbb{E}_V n)^{-\frac{1}{2}}),$$

which is smaller than the expectation by a factor $(\mathbb{E}_V n)^{-\frac{1}{2}} = t^{-\frac{1}{2}}$. Or, equivalently, the coefficient of variation $\frac{\sigma(\overline{\mathrm{cr}}(G_L))}{\mathbb{E}_V \overline{\mathrm{cr}}(G_L)}$ is of order $t^{-\frac{1}{2}}$. As $t \to \infty$, our bounds on the expectation and variance together with Chebychev's inequality lead to

$$\mathbb{P}\left(\left|\frac{\overline{\mathrm{cr}}(G_L)}{t^4 \delta_t^{2d+2}} - \frac{\mathbb{E}_V \overline{\mathrm{cr}}(G_L)}{t^4 \delta_t^{2d+2}}\right| \geq \varepsilon\right) \leq \frac{\mathbb{V}\mathrm{ar}_V \overline{\mathrm{cr}}(G_L)}{t^8 \delta_t^{4d+4} \varepsilon^2} \to 0.$$

Corollary 3 (Law of Large Numbers). *For given L, the normalized random crossing number converges in probability (with respect to the Poisson point process V) as $t \to \infty$,*

$$\frac{\overline{\mathrm{cr}}(G_L)}{t^4 \delta_t^{2d+2}} \to \frac{1}{8} c_d I^{(2)}(W, L).$$

Until now we fixed a plane L and computed the variance with respect to the random points V. Theorems 1 and 2 allow to compute the expectation and variance with respect to V *and* a randomly chosen plane L. For the expectation we obtain from Theorem 1 and by Fubini's theorem

$$\mathbb{E}_{L,V} \overline{\mathrm{cr}}(G_L) = \frac{1}{8} c_d \, t^4 \delta_t^{2d+2} \int_{\mathcal{L}} I^{(2)}(W, L) \, dL + o(t^4 \delta_t^{2d+2}), \tag{8}$$

as $t \to \infty$ and $\delta_t \to 0$, where dL denotes integration with respect to the Haar measure on $\mathcal{L}$. For simplicity we assume in the following that $\lim_{t\to\infty}(t\delta_t^d)^{-1} = 0$. We use the variance decomposition $\mathbb{V}\mathrm{ar}_{L,V} X = \mathbb{E}_L \mathbb{V}\mathrm{ar}_V X + \mathbb{V}\mathrm{ar}_L \mathbb{E}_V X$. By

$$\mathbb{E}_L \mathbb{V}\mathrm{ar}_V \overline{\mathrm{cr}}(G_L) = \frac{1}{64} c_d^2 \, t^7 \delta_t^{4d+4} \int_{\mathcal{L}} I^{(3)}(W, L) \, dL \; + o(t^7 \delta_t^{4d+4}), \quad \text{and}$$

$$\mathbb{V}\mathrm{ar}_L \mathbb{E}_V \overline{\mathrm{cr}}(G_L) = \mathbb{E}_L (\mathbb{E}_V \overline{\mathrm{cr}}(G_L))^2 - (\mathbb{E}_{L,V} \overline{\mathrm{cr}}(G_L))^2$$

$$= \frac{1}{64} c_d^2 \, t^8 \delta_t^{4d+4} \left[\int_{\mathcal{L}} I^{(2)}(W, L)^2 \, dL - \left(\int_{\mathcal{L}} I^{(2)}(W, L) dL \right)^2 \right] + o(t^8 \delta_t^{4d+4})$$

we obtain

$$\mathbb{V}\mathrm{ar}_{L,V} \overline{\mathrm{cr}}(G_L) = \frac{1}{64} c_d^2 \, t^8 \delta_t^{4d+4} \left[\int_{\mathcal{L}} I^{(2)}(W, L)^2 dL - \left(\int_{\mathcal{L}} I^{(2)}(W, L) dL \right)^2 \right]$$
$$+ o(t^8 \delta_t^{4d+4}). \tag{9}$$

Hölder's inequality implies that the term in brackets is positive as long as $I^{(2)}(W, L)$ is not a constant function.

3.3   The Rotation Invariant Case

If W is the ball B of unit volume and thus V is *rotation invariant*, then $I^{(2)}(B, L) = I^{(2)}(B)$ is a constant function independent of L, and the leading term in (9) is vanishing. From (8) we see that in this case the expectation is independent of L.

$$\mathbb{E}_V \overline{\mathrm{cr}}(G_L) = \mathbb{E}_L \mathbb{E}_V \overline{\mathrm{cr}}(G_L) = t^4 \delta_t^{2d+2} I^{(2)}(B) + o(t^4 \delta_t^{2d+2})$$

For the variance this implies $\mathbb{V}\mathrm{ar}_L \mathbb{E}_V \overline{\mathrm{cr}}(G_L) = 0$, and hence

$$\mathbb{V}\mathrm{ar}_{L,V} \overline{\mathrm{cr}}(G_L) = \mathbb{E}_L \mathbb{V}\mathrm{ar}_V \overline{\mathrm{cr}}(G_L) = \frac{1}{64} c_d^2 t^7 \delta_t^{4d+4} I^{(3)}(B) + o(t^7 \delta_t^{4d+4}).$$

In this case the variance $\mathbb{V}\mathrm{ar}_{L,V}$ is of the order t^{-1}—and thus surprisingly significantly—smaller than in the general case.

Theorem 4. *Let G_L be the projection of an RGG in the ball $B \subset \mathbb{R}^d$, $d \geq 3$, onto a two-dimensional uniformly chosen random plane L. Then*

$$\mathbb{E}_{L,V} \overline{\mathrm{cr}}(G_L) = \frac{1}{8} c_d \, t^4 \delta_t^{2d+2} I^{(2)}(B) + o(t^4 \delta_t^{2d+2}) \quad and$$

$$\mathbb{V}\mathrm{ar}_{L,V} \overline{\mathrm{cr}}(G_L) = \frac{1}{64} c_d^2 \, t^7 \delta_t^{4d+4} I^{(3)}(B) + o(t^7 \delta_t^{4d+4}),$$

as $t \to \infty$, $\delta_t \to 0$ and $t\delta_t^d \to \infty$.

Again, Chebychev's inequality immediately yields a law of large numbers which states that with high probability the crossing number of G_L in a random direction is very close to $\frac{1}{8} c_d \, t^4 \delta_t^{2d+2} I^{(2)}(B)$.

Corollary 5 (Law of Large Numbers). *Let G_L be the projection of an RGG in $B \subset \mathbb{R}^d$, $d \geq 3$, onto a random two-dimensional plane L. Then the normalized random crossing number converges in probability (with respect to the Poisson point process V and to L), as $t \to \infty$,*

$$\frac{\overline{cr}(G_L)}{t^4 \delta_t^{2d+2}} \;\rightarrow\; \frac{1}{8} c_d I^{(2)}(B).$$

As known by the crossing lemma, the optimal crossing number is of order $\frac{m^3}{n^2}$. In our setting this means that we are looking for the optimal direction of projection which leads to a crossing number of order $t^4 \delta_t^{3d}$, much smaller than the expectation $\mathbb{E}_V \overline{cr}(G_L)$. Chebychev's inequality shows that if $W = B$ it is difficult to find this optimal direction and to reach this order of magnitude; using $\delta_t \to 0$ in the last step we have:

$$\mathbb{P}_{L,V}\big(\overline{cr}(G_L) \leq ct^4 \delta_t^{3d}\big) \leq \mathbb{P}_{L,V}\big(|\overline{cr}(G_L) - \mathbb{E}_{L,V}\overline{cr}(G_L)| \geq \mathbb{E}_{L,V}\overline{cr}(G_L) - ct^4 \delta_t^{3d}\big)$$

$$\leq \frac{\mathbb{V}\mathrm{ar}_{L,V}\overline{cr}(G_L)}{(\mathbb{E}_{L,V}\overline{cr}(G_L) - ct^4 \delta_t^{3d})^2} = O(t^{-1}).$$

Hence a computational naïve approach of minimizing the crossing numbers by just projecting onto a sample of random planes seems to be expensive. This suggests to combine the search for an optimal choice of the direction of projection with other quantities of the RGG. It is a long standing assumption in graph drawing that there is a connection between the crossing number and the stress of a graph. Therefore the next section is devoted to investigations concerning the stress of RGGs.

4   The Stress of an RGG

According to (1) we define the stress of G_L as

$$\mathrm{stress}(G, G_L) := \frac{1}{2} \sum_{(v_1, v_2) \in V_{\neq}^2} w(v_1, v_2)(d_0(v_1, v_2) - d_L(v_1, v_2))^2,$$

where $w(v_1, v_2)$ a positive weight-function and d_0 resp. d_L are the distances between v_1 and v_2, resp $v_1|_L$ and $v_2|_L$. As $\overline{cr}(G)$, stress is a U-statistic, but now of order two. Using the Slivnyak-Mecke formula, it is immediate that

$$\mathbb{E}_V \mathrm{stress}(G, G_L) = \frac{1}{2} t^2 \underbrace{\int_{W^2} w(v_1, v_2)(d_0(v_1, v_2) - d_L(v_1, v_2))^2 dv_1 dv_2}_{=:S^{(1)}(W,L)}.$$

For the variance, the Poincaré inequality (4) implies

$$\mathbb{V}\text{ar}_V\text{stress}(G, G_L) \leq t \int\limits_W \mathbb{E}_V \left(D_v(\text{stress}(G, G_L))\right)^2 dv$$

$$= \frac{1}{4}t \int\limits_W \mathbb{E}_V \left(\sum_{v_1 \in V} w(v, v_1)(d_0(v, v_1) - d_L(v, v_1))^2\right)^2 dv$$

$$= \frac{1}{4}t^3 \underbrace{\int\limits_{W^3} \prod_{i=1}^{2} \left(w(v, v_i)(d_0(v, v_i) - d_L(v, v_i))^2\right) dv_1 dv_2 dv}_{=:S^{(2)}(W,L)}$$

$$+ \frac{1}{4}t^2 \int\limits_{W^2} w(v, v_1)^2 (d_0(v, v_1) - d_L(v, v_1))^4 \, dv_1 dv.$$

Hence the standard deviation of the stress is smaller than the expectation by a factor $t^{-\frac{1}{2}}$ and thus the stress is concentrated around its mean. Again the computation of the lower bound for the variance in (4) is asymptotically sharp.

$$\mathbb{V}\text{ar}_V\text{stress}(G, G_L) \geq t \int\limits_W (\mathbb{E}_V D_v(\text{stress}(G, G_L)))^2 dv$$

$$= \frac{1}{4}t \int\limits_W \left(\mathbb{E}_V \sum_{v_1 \in V} w(v, v_1)(d_0(v, v_1) - d_L(v, v_1))^2\right)^2 dv = \frac{1}{4}t^3 S^{(2)}(W, L).$$

Theorem 6. *Let G_L be the projection of an RGG in $\mathbb{R}^d$, $d \geq 3$, onto a two-dimensional plane L. Then*

$$\mathbb{E}_V\text{stress}(G, G_L) = \frac{1}{2} t^2 \, S^{(1)}(W, L) \quad and$$

$$\mathbb{V}\text{ar}_V\text{stress}(G, G_L) = \frac{1}{4} t^3 \, S^{(2)}(W, L) + O(t^2).$$

The discussions from Sects. 3.2 and 3.3 lead to analogous results for the stress of the RGG. Using Chebychev's inequality we could derive a law of large numbers. Taking expectations with respect to a uniform plane L we obtain:

$$\mathbb{E}_{L,V}\text{stress}(G, G_L) = \frac{1}{2}t^2 \int\limits_{\mathcal{L}} S^{(1)}(W, L)dL,$$

$$\mathbb{V}\text{ar}_{L,V}\text{stress}(G, G_L) = \frac{1}{4}t^4 \left[\int\limits_{\mathcal{L}} S^{(1)}(W, L)^2 dL - \left(\int\limits_{\mathcal{L}} S^{(1)}(W, L)dL\right)^2\right] + O(t^3).$$

Again, the term in brackets is only vanishing if $W = B$. In this case

$$\mathbb{V}\text{ar}_{L,V}\text{stress}(G, G_L) = \mathbb{E}_L \mathbb{V}\text{ar}_V\text{stress}(G, G_L) = \frac{1}{4}t^3 S^{(2)}(B) + O(t^2).$$

5   Correlation Between Crossing Number and Stress

It seems to be widely conjectured that the crossing number and the stress should be positively correlated. Yet it also seems that a rigorous proof is still missing. It is the aim of this section to provide the first proof of this conjecture, in the case where the graph is a random geometric graph.

Clearly, by the definition of $\overline{\mathrm{cr}}$ and stress we have

$$D_v \, \overline{\mathrm{cr}}(G_L) \geq 0 \text{ and } D_v \, \mathrm{stress}(G, G_L) \geq 0,$$

for all v and all realizations of V. Such a functional F satisfying $D_v(F) \geq 0$ is called increasing. The Harris-FKG inequality for Poisson point processes [23] links this fact to the correlation of $\overline{\mathrm{cr}}(G_L)$ and $\mathrm{stress}(G, G_L)$.

Theorem 7. *Because* stress *and* $\overline{\mathrm{cr}}$ *are increasing we have*

$$\mathbb{E}_V \overline{\mathrm{cr}}(G_L)\mathrm{stress}(G, G_L) \geq \mathbb{E}_V \overline{\mathrm{cr}}(G_L) \, \mathbb{E}_V \mathrm{stress}(G, G_L),$$

and thus the correlation is positive.

We immediately obtain that the covariance is positive and is of order at most

$$\mathbb{C}\mathrm{ov}_V\big(\overline{\mathrm{cr}}(G_L), \mathrm{stress}(G_L)\big) \leq \sqrt{\mathbb{V}\mathrm{ar}_V \overline{\mathrm{cr}}(G_L) \, \mathbb{V}\mathrm{ar}_V \mathrm{stress}(G, G_L)}$$

$$\leq \frac{1}{16}c_d\Big(1 + \frac{2\pi\kappa_d}{c_d} \lim_{t \to \infty} \frac{1}{t\delta_t^d}\Big)^{\frac{1}{2}} t^5 \delta_t^{2d+2} I^{(3)}(W, L)^{\frac{1}{2}} S^{(2)}(W, L)^{\frac{1}{2}} + o(t^5 \delta_t^{2d+2}).$$

In [11, Appendix C] we use Mehler's formula to prove a lower bound:

$$\mathbb{C}\mathrm{ov}_V\big(\overline{\mathrm{cr}}(G_L), \mathrm{stress}(G, G_L)\big) \geq \frac{t^5}{16} \int_{W^2} I_W(v)w(v, v_1)(d_0(v, v_1) - d_L(v, v_1))^2 dv_1 dv.$$

We combine this bound with (5), divide by the standard deviations from Theorems 2 and 6 and obtain the asymptotics for the correlation coefficient:

Theorem 8. *Let G_L be the projection of an RGG in $\mathbb{R}^d$, $d \geq 3$, onto a two-dimensional plane L. Then*

$$\lim_{t \to \infty} \mathbb{C}\mathrm{orr}_V\big(\overline{\mathrm{cr}}(G_L), \mathrm{stress}(G, G_L)\big)$$

$$\geq \frac{\int_{W^2} \mathrm{vol}_{d-2}((v + L^\perp) \cap W)w(v, v_1)(d_0(v, v_1) - d_L(v, v_1))^2 dv_1 dv}{(1 + \frac{2\pi\kappa_d}{c_d} \lim_{t \to \infty} \frac{1}{t\delta_t^d})^{\frac{1}{2}} I^{(3)}(W, L)^{\frac{1}{2}} S^{(2)}(W, L)^{\frac{1}{2}}}.$$

It can be shown that this bound is even tight and asymptotically gives the correct correlation coefficient.

5.1   The Rotation Invariant Case

In principle the bounds for the covariance in the Poisson point process V given above can be used to compute covariance bounds in L and V when L is not fixed but random. For this we could use the covariance decomposition

$$\mathbb{Cov}_{L,V}(X,Y) = \mathbb{E}_L \mathbb{Cov}_V(X,Y) + \mathbb{Cov}_L(\mathbb{E}_V X, \mathbb{E}_V Y).$$

Here we concentrate again on the case when $W = B$ is the ball of unit volume and thus V is rotation invariant. Then $\mathbb{Cov}_L(\mathbb{E}_V \overline{\mathrm{cr}}(G_L), \mathbb{E}_V \mathrm{stress}(G_L)) = 0$, and as an immediate consequence of Theorem 8 we obtain

Corollary 9. *Let G_L be the projection of an RGG in $B \subset \mathbb{R}^d$, $d \geq 3$, onto a two-dimensional random plane L. Then the correlation between the crossing number and the stress of the RGG is positive with*

$$\lim_{t \to \infty} \mathbb{Corr}_{L,V}(\overline{\mathrm{cr}}(G_L), \mathrm{stress}(G, G_L))$$

$$\geq \frac{\int_{B^2} \mathrm{vol}_{d-2}(v + L^{\perp}) \cap B) w(v, v_1)(d_0(v, v_1) - d_L(v, v_1))^2 dv_1 dv}{(1 + \frac{2\pi\kappa_d}{c_d} \lim_{t \to \infty} \frac{1}{t\delta_t^d})^{\frac{1}{2}} I^{(3)}(B)^{\frac{1}{2}} S^{(2)}(B)^{\frac{1}{2}}}.$$

In particular, the correlation does not vanish as $t \to \infty$. This gives the first proof we are aware of, that there is a strict positive correlation between the crossing number and the stress of a graph. Hence, at least for RGGs, the method to optimize the stress to obtain good crossing numbers can be supported by rigorous mathematics.

6   Consequences and Conclusion

Apart from providing precise asymptotics for the crossing numbers of drawings of random geometric graphs, the main findings are the positive covariance and the non-vanishing correlation between the stress and the crossing number of the drawing of a random geometric graph. Of interest would be whether $\mathbb{Cov}_L(\overline{\mathrm{cr}}(G_L), \mathrm{stress}(G, G_L)) > 0$ for arbitrary graphs G. Yet there are simple examples of graphs G where this is wrong. Yet we could ask in a slightly weaker form whether at least $\mathbb{E}_V \mathbb{Cov}_L(\overline{\mathrm{cr}}(G_L), \mathrm{stress}(G_L)) > 0$, but we have not been able to prove that.

We may coarsely summarize the gist of all the above findings algorithmically in the context of crossing number approximation, ignoring precise numeric terms that can be found above. We yield the first (expected) crossing number approximations for a rich class of randomized graphs:

Corollary 10. *Let G be a random geometric graph in $\mathbb{R}^2$ (unit-disk graph) as defined above.* With high probability, *the number of crossings in its natural straight-line drawing is at most a* constant factor *away from $\mathrm{cr}(G_0)$ and $\overline{\mathrm{cr}}(G_0)$.*

Corollary 11. *Let G be a random geometric graph in $\mathbb{R}^d$ (unit-ball graph) as defined above. We obtain a straight-line drawing D by projecting it onto a randomly chosen 2D plane.* With high probability, *the number of crossings in D is at most a* factor α *away from* $\mathrm{cr}(G_0)$ *and* $\overline{\mathrm{cr}}(G_0)$. *Thereby, α is only dependent on the graph's density.*

Corollary 12. *Let G be a random geometric graph and use its natural distances in $\mathbb{R}^d$ as input for stress minimization. The stress is positively correlated to the crossing number. Loosely speaking, a drawing of G with close to minimal stress is expected* to yield a close to minimal number of crossings.

References

1. Aigner, M., Ziegler, G.M.: Proofs from THE BOOK, 4th edn. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-00856-6
2. Ajtai, M., Chvátal, V., Newborn, M., Szemerédi, E.: Crossing-free subgraphs. In: Theory and Practice of Combinatorics, North-Holland Mathematics Studies, North-Holland, vol. 60, pp. 9–12 (1982)
3. Biedl, T.C., Chimani, M., Derka, M., Mutzel, P.: Crossing number for graphs with bounded pathwidth. In: Proceedings of International Symposium on Algorithms and Computation (ISAAC) 2017, pp. 13:1–13:13. LIPIcs 92 (2017)
4. Bienstock, D.: Some provably hard crossing number problems. Discrete Comput. Geom. **6**(5), 443–459 (1991)
5. Brandes, U., Pich, C.: An experimental study on distance-based graph drawing. In: Tollis, I.G., Patrignani, M. (eds.) GD 2008. LNCS, vol. 5417, pp. 218–229. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-00219-9_21
6. Cabello, S.: Hardness of approximation for crossing number. Discrete Comput. Geom. **49**(2), 348–358 (2013)
7. Cabello, S., Mohar, B.: Crossing number and weighted crossing number of near-planar graphs. Algorithmica **60**(3), 484–504 (2011). https://doi.org/10.1007/s00453-009-9357-5
8. Chimani, M., et al.: People prefer less stress and fewer crossings. In: Duncan, C., Symvonis, A. (eds.) Proceedings of International Symposium on Graph Drawing (GD) 2014. LNCS, vol. 8871, pp. 523–524. Springer, Heidelberg (2014)
9. Chimani, M., Hliněný, P.: Inserting multiple edges into a planar graph. In: Proceedings of International SYmposium on Computational Geometry (SoCG) 2016, pp. 30:1–30:15. LIPIcs 51 (2016)
10. Chimani, M., Hliněný, P., Mutzel, P.: Vertex insertion approximates the crossing number for apex graphs. Eur. J. Comb. **33**, 326–335 (2012)
11. Chimani, M., Döring, H., Reitzner, M.: Crossing numbers and stress of random graphs. eprint arXiv:1808.07558v1 (2018)
12. Chimani, M., Hliněný, P.: A tighter insertion-based approximation of the crossing number. J. Comb. Optim. **33**(4), 1183–1225 (2017)
13. Chuzhoy, J.: An algorithm for the graph crossing number problem. In: Proceedings of ACM Symposium on Theory of Computing (STOC) 2011, pp. 303–312. ACM (2011)
14. Fox, J., Pach, J., Suk, A.: Approximating the Rectilinear crossing number. In: Hu, Y., Nöllenburg, M. (eds.) GD 2016. LNCS, vol. 9801, pp. 413–426. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-50106-2_32

15. Garey, M.R., Johnson, D.S.: Crossing number is NP-complete. SIAM J. Algebr. Discrete Methods **4**(3), 312–316 (1983)
16. Gitler, I., Hliněný, P., Leaño, J., Salazar, G.: The crossing number of a projective graph is quadratic in the face-width. Electron. J. Comb. **29**, 219–233 (2007)
17. Hliněný, P., Chimani, M.: Approximating the crossing number of graphs embeddable in any orientable surface. In: Proceedings of ACM-SIAM Symposium on Discrete Algorithms (SODA) 2010, pp. 918–927 (2010)
18. Hliněný, P., Salazar, G.: Approximating the crossing number of toroidal graphs. In: Tokuyama, T. (ed.) ISAAC 2007. LNCS, vol. 4835, pp. 148–159. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-77120-3_15
19. de Klerk, E., Maharry, J., Pasechnik, D.V., Richter, R.B., Salazar, G.: Improved bounds for the crossing numbers of $K_{m,n}$ and K_n. SIAM J. Discrete Math. **20**(1), 189–202 (2006)
20. Klimenta, M., Brandes, U.: Graph drawing by classical multidimensional scaling: new perspectives. In: Didimo, W., Patrignani, M. (eds.) GD 2012. LNCS, vol. 7704, pp. 55–66. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-36763-2_6
21. Kobourov, S.G.: Force-directed drawing algorithms. In: Handbook of Graph Drawing and Visualization, pp. 383–408. CRC Press (2013)
22. Kobourov, S.G., Pupyrev, S., Saket, B.: Are crossings important for drawing large graphs? In: Duncan, C., Symvonis, A. (eds.) GD 2014. LNCS, vol. 8871, pp. 234–245. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-45803-7_20
23. Last, G., Penrose, M.: Fock space representation, chaos expansion and covariance, inequalities for general Poisson processes. Probab. Theory Relat. Fields **150**, 663–690 (2011)
24. Last, G., Penrose, M.: Lectures on the Poisson Process. Cambridge University Press, Cambridge (2017)
25. Leighton, F.T.: Complexity Issues in VLSI: Optimal Layouts for the Shuffle-exchange Graph and Other Networks. MIT Press, Cambridge (1983)
26. Peccati, G., Reitzner, M. (eds.): Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Itô Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7. Springer, Switzerland (2016). https://doi.org/10.1007/978-3-319-05233-5. Bocconi University Press, Milan
27. Penrose, M.: Random Geometric Graphs (Oxford Studies in Probability). Oxford University Press, Oxford (2004)
28. Reitzner, M., Schulte, M.: Central limit theorems for U-statistics of poisson point processes. Ann. Probab. **41**, 3879–3909 (2013)
29. Reitzner, M., Schulte, M., Thäle, C.: Limit theory for the Gilbert graph. Adv. Appl. Math. **88**, 26–61 (2017)
30. Schaefer, M.: Crossing Numbers of Graphs. CRC-Press, Boca Raton (2017)
31. Schneider, R., Weil, W.: Stochastic and Integral Geometry. Probability and Its Applications. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-78859-1
32. Wu, L.: A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. Probab. Theory Relat. Fields **118**, 427–438 (2000)