Shape of shortest paths in random spatial networks

Alexander P. Kartun-Giles,1,* Marc Barthelemy,2,† and Carl P. Dettmann3,‡

1Max Plank Institute for Mathematics in the Sciences, Leipzig, Germany
2Institut de Physique Théorique, CEA, CNRS-URA 2306, Gif-sur-Yvette, France
3School of Mathematics, University of Bristol, Bristol BS8 1UG, United Kingdom

I. INTRODUCTION

Many complex systems assume the form of a spatial network [1,2]. Transport networks, neural networks, communication and wireless sensor networks, power and energy networks, and ecological interaction networks are all important examples where the characteristics of a spatial network structure are key to understanding the corresponding emergent dynamics.

Shortest paths form an important aspect of their study. Consider for example the appearance of bottlenecks impeding traffic flow in a city [3,4], the emergence of spatial small worlds [5,6], bounds on the diameter of spatial preferential attachment graphs [7–9], the random connection model [10–13], or in spatial networks generally [14,15], as well as geometric effects on betweenness centrality measures in complex networks [11,16] and navigability [17].

First-passage percolation (FPP) [18] attempts to capture these features with a probabilistic model. As with percolation [19], the effect of spatial disorder is crucial. Compare this to the elementary random graph [20]. In FPP one usually considers a deterministic lattice such as $\mathbb{Z}^d$ with independent, identically distributed weights, known as local passage times, on the edges. With a fluid flowing outward from a point, the question is as follows: What is the minimum passage time over all possible routes between this and another distant point, where routing is quicker along lower weighted edges?

More than 50 years of intensive study of FPP has been carried out [21]. This has led to many results such as the subadditive ergodic theorem, a key tool in probability theory, but also a number of insights in crystal and interface growth [22], the statistical physics of traffic jams [19], and key ideas of universality and scale invariance in the shape of shortest paths [23]. As an important intersection between probability and geometry, it is also part of the mathematical aspects of a theory of gravity beyond general relativity, and perhaps in the foundations of quantum mechanics, since it displays fundamental links to complexity, emergent phenomena, and randomness in physics [24,25].

A particular case of FPP is the topic of this article, known as Euclidean first-passage percolation (EFPP). This is a probabilistic model of fluid flow between points of a $d$-dimensional Euclidean space, such as the surface of a hypersphere. One studies optimal routes from a source node to each possible destination node in a spatial network built either randomly or deterministically on the points. Introduced by Howard and Newman much later in 1997 [26] and originally a weighted complete graph, we adopt a different perspective by considering edge weights given deterministically by the Euclidean distances between the spatial points themselves. This is in sharp contrast with the usual FPP problem, where weights are independent and identically distributed random variables.

Howard’s model is defined on the complete graph constructed on a point process. Long paths are discouraged by...
and weight the edges with their Euclidean length (see Fig. 1).

We then study the random length and transversal deviation of pairs of points according to given rules [28, 29] rather than the Euclidean length of the path is given by the sum of all edges length: $d(x, y) = |x - a| + |a - b| + |b - c| + |c - d| + |d - y|$.

The variant of EFPP we study is instead defined on a Poisson point process in an unbounded region (by definition, the number of points in a bounded region is a Poisson random variable), see, for example, Ref. [27]), but with links added between pairs of points according to given rules [28, 29] rather than the totality of the weighted complete graph. More precisely, the model we study in this paper is defined as follows. We take a random spatial network such as the random geometric graph constructed over a set of points denoted by circles here and the edges are denoted by lines. For a pair of nodes $(x, y)$ we look for the shortest path (shown here by a dotted line) where the length of the path is given by the sum of all edges length: $d(x, y) = |x - a| + |a - b| + |b - c| + |c - d| + |d - y|$.

In EFPP, we first construct a Poisson point process in $\mathbb{R}^d$ which forms the basis of an undirected graph. A fluid or current then flows outward from a single source at a constant speed with a travel time along an edge given by a power $\alpha > 1$ of the Euclidean length of the edge along which it travels [26]. See Fig. 2, where the model is shown on six different random spatial network models.

To highlight the difference between these results and our own, we have edge weights which are not independent random variables but interpoint distances. As far as we are aware, this has not been addressed directly in the literature. The reader eager to view the results can skip this section at first reading, apart from the definitions of II A; however, the remaining background is very helpful for appreciating the later discussion. In Sec. III we introduce the various spatial networks studied here, and in Sec. IV we present the numerical method and our new results on the EFPP model on random graphs. In particular, due to arguments based on scale invariance, we expect the appearance of power laws and universal exponents [23, Sec. 1]. We reveal the scaling exponents of the geodesics for the complete graph and for the network models studied here and also show numerical results about the travel-time and transversal deviation distribution. In particular, we find distinct exponents from the Kardar-Parisi-Zhang (KPZ) class (see, for example, Ref. [34] and references therein) which has wandering and fluctuation exponents $\xi = 2/3$ and $\chi = 1/3$, respectively. Importantly, we conjecture and numerically corroborate a Gaussian central limit theorem for the travel-time fluctuations, on the scale $t^{1/5}$ for the RGG and the other proximity graphs and $t^{3/5}$ for the Delaunay triangulation and other excluded region graphs, which is also distinct from KPZ where the Tracy-Widom distribution, and the scale $t^{1/3}$, is the famous outcome. Finally, in Sec. V we present some analytic ideas which help explain the distinction between universality classes. We then conclude and discuss some open questions in Sec. VI.

II. BACKGROUND: FPP AND EFPP

In EFPP, we first construct a Poisson point process in $\mathbb{R}^d$ which forms the basis of an undirected graph. A fluid or current then flows outward from a single source at a constant speed with a travel time along an edge given by a power $\alpha > 1$ of the Euclidean length of the edge along which it travels [26]. See Fig. 2, where the model is shown on six different random spatial network models.

Developing FPP in this setting, Santalla et al. [35] recently studied the model on spatial networks, as we do here. Instead of EFPP, they weight the edges of the Delaunay triangulation, and also the square lattice, with independent and identically distributed variates, for example, $U(a, b)$ for $a, b > 0$, and proceed to numerically verify the existence of the KPZ class for the geodesics, see, e.g., Ref. [36]. Moreover, FPP on small-world networks and Erdős-Rényi random graphs are studied by Bhamidi, van der Hofstad, and Hooghiemstra in Ref. [37], who discuss applications in diverse fields such as magnetism [38], wireless ad hoc networks [10, 12, 39], competition in ecological systems [40], and molecular biology [41]. See also their work specifically on random graphs [42]. Optimal paths in disordered complex networks, where disorder is weighting the edges with independent and identically distributed random variables, is studied by Braunstein et al. [43] and later by Chen et al. [44]. We also point to the recent analytic results of Bakhtin and Wu, who have provided a good lower bound rate of growth of geodesic wandering, which in fact we find to be met with equality in the random geometric graph [45].

To highlight the difference between these results and our own, we have edge weights which are not independent random variables but interpoint distances. As far as we are aware, this has not been addressed directly in the literature.
SHAPE OF SHORTEST PATHS IN RANDOM SPATIAL NETWORKS

FIG. 2. Spatial networks, each built on a different realization of a simple, stationary Poisson point processes of expected $\rho = 1000$ points in the unit square $V = [-1/2, 1/2]^2$ but with different connection laws. The boundary points at time $t = 1/2$ of the first-passage process are shown in red, while their respective geodesics are shown in blue. (a) Hard RGG with unit disk connectivity. (b) Soft RGG with Rayleigh fading connection function $H(r) = \exp(-\beta r^2)$; (c) 7-NNG; (d) relative neighborhood graph, which is the lune-based $\beta$ skeleton for $\beta = 2$; (e) Gabriel graph, which is the lune-based $\beta$ skeleton for $\beta = 1$; and (f) the Delaunay triangulation.

A. First-passage percolation

Given independent and identically distributed weights, paths are sums of independent and identically distributed random variables. The lengths of paths between pairs of points can be considered to be a random perturbation of the plane metric. In fact, these lengths, and the corresponding transversal deviations of the geodesics, have been the focus of in-depth research over the since the late 1960s [21]. They exist as minima over collections of correlated random variables.

The travel times are conjectured (in the independent and identically distributed) case to converge to the Tracy-Widom distribution (TW), found throughout various models of statistical physics, see, e.g., Ref. [35, Sec. 1]. This links the model to random matrix theory, where $\beta$-TW appears as the limiting distribution of the largest eigenvalue of a random matrix in the $\beta$-hermite ensemble, where the parameter $\beta$ is 1, 2, or 4 [46].

The original FPP model is defined as follows. We consider vertices in the $d$-dimensional lattice $L^d = (Z^d, E^d)$, where $E^d$ is the set of edges. We then construct the function $\tau : E^d \to (0, \infty)$, which gives a weight for each edge and is usually assumed to be identically independently distributed random variables. The passage time from vertices $x$ to $y$ is the random variable given as the minimum of the sum of the $\tau$’s over all possible paths $P$ on the lattice connecting these points,

$$T(x, y) = \min_P \sum_P \tau(e).$$

This minimum path is a geodesic, and it is almost surely unique when the edge weights are continuous.

The average travel time is proportional to the distance

$$\mathbb{E}[T(x, y)] \sim |x - y|,$$

where here and in the following we denote the average of a quantity by $\mathbb{E}(\cdot)$ and where $a \sim b$ means $a$ converges to $Cb$ with $C$ a constant independent of $x, y$, as $|x - y| \to \infty$. More generally, if the ratio of the geodesic length and the Euclidean distance is less than a finite number $t$ (the maximum value of this ratio is called the stretch), then the network is a $t$ spanner [47]. Many important networks are $t$ spanners, including the Delaunay triangulation of a Poisson point process, which has $\pi/2 < t < 1.998$ [48,49]. The variance of the passage time over some distance $|x - y|$ is also important and scales as

$$\text{Var}[T(x, y)] \sim |x - y|^{2t}.$$
FIG. 3. Example Euclidean geodesics (blue) running between two end nodes of a simple, stationary Poisson point process (red). The maximal transversal deviation is shown (vertical black line). The Euclidean distance between the endpoints is the horizontal black line. The PPP density is equal for each model. (a) Hard RGG, (b) soft RGG with connectivity probability $H(r) = \exp(-r^2)$, (c) 7-NNG, (d) RNG, (e) GG, and (f) DT.

exponent $\xi$, i.e.,

$$E[D(x, y)] \sim |x - y|^\xi$$

for large $|x - y|$. Knowing $\xi$ informs us about the geometry of geodesics between two distant points. See Fig. 3 for an illustration of wandering on different networks.

The other exponent, $\chi$, informs us about the variance of their random length. Another way to see this exponent is to consider a ball of radius $R$ around any point. For large $R$, the ball has an average radius proportional to $R$ and the fluctuations around this average grow as $R^\chi$. With $\chi < 1$ the fluctuations die away $R \to \infty$, leading to the shape theorem, see, e.g., Ref. [21, Sec. 1].

1. Sublinear variance in FPP

According to Benjamini, Kalai, and Schramm, Var[$T(x, y)$] grows sublinearly with $|x - y|$ [50], a major theoretical step in characterizing their scaling behavior. With $C$ some constant which depends only on the distribution of edge weights and the dimension $d$, they prove that

$$\text{Var}[T(x, y)] \leq C|x - y|/\log |x - y|.$$  \hspace{1cm} (5)

The numerical evidence, in fact, shows this follows the non-typical scaling law $|x - y|^{2/3}$. Transversal fluctuations also scale as $|x - y|^{2/3}$ [21]. In this case, the fluctuations of $T$ are asymptotic to the TW distribution. According to recent results of Santalla et al. [51], curved spaces lead to similar fluctuations of a subtly different kind: If we embed the graph on the surface of a cylinder, then the distribution changes from the largest eigenvalue of the GUE, to GOE, ensembles of random matrix theory.

When we see a sum of random variables, it is natural to conjecture a central limit theorem, where the fluctuations of the sum, after rescaling, converge to the standard normal distribution in some limit, in this case as $|x - y| \to \infty$. Durrett writes in a review that “...novice readers would expect a central limit theorem being proved...however physicists tell us that in two dimensions, the standard deviation is of order $|x - y|^{1/3}$” (see Ref. [50, Section 1]). This suggests that one does not have a Gaussian central limit theorem for the travel-time fluctuations in the usual way. This has been rigorously proven [52–54].

2. Scaling exponents

A well-known result in the two-dimensional lattice case [55] is that $\chi = 1/3$, $\xi = 2/3$. Also, another belief is that $\chi = 0$ for dimensions $d$ large enough. Many physicists, see, for example, Refs. [55–61], also conjecture that independently from the dimension, one should have the so-called KPZ relation between these exponents

$$\chi = 2\xi - 1.$$  \hspace{1cm} (6)

This is connected with the KPZ universality class of random geometry, apparent in many physical situations, including traffic and data flows, and their respective models, including the corner growth model, ASEP, TASEP, etc. [19,62,63]. In particular, FPP is in direct correspondence with important problems in statistical physics [34] such as the directed polymer in random media (DPRM) and the KPZ equation, in which case the dynamical exponent $z$ corresponds to the wandering exponent $\xi$ defined for the FPP [35,64].
3. Bounds on the exponents

The situation regarding exact results is more complex [21,36]. The majority of results are based on the model on $\mathbb{Z}^d$. Kesten [65] proved that $\chi \leq 1/2$ in any dimension, and for the wandering exponent $\xi$, Licea et al. [66] gave some hints that possibly $\xi \geq 1/2$ in any dimension and $\xi \geq 3/5$ for $d = 2$.

Concerning the KPZ relation, Wehr and Aizenman [67] and Licea et al. [66] proved the inequality

$$\chi \geq (1 - d\xi)/2$$  \hspace{1cm} (7)

in $d$ dimensions. Newman and Piza [68] gave some hints that possibly $\chi \geq 2\xi - 1$. Finally, Chatterjee [36] proved Eq. (6) for $\mathbb{Z}^d$ with independent and identically distributed random edge weights, with some restrictions on distributional properties of the weights. These were lifted by independent work of Auffinger and Damron [21].

B. Euclidean first-passage percolation

Euclidean first-passage percolation [26] adopts a very different perspective from FPP by considering a fluid flowing along each of the links of the complete graph on the points at some weighted speed given by a function, usually a power, of the Euclidean length of the edge. We ask, between two points of the process at large Euclidean distance $|x - y|$, What is the minimum passage time over all possible routes?

More precisely, the original model of Howard and Newman goes as follows. Given a domain $\mathcal{V}$ such as the Euclidean plane, and $dx$ Lesbegue measure on $\mathcal{V}$, consider a Poisson point process $\mathcal{X} \subset \mathcal{V}$ of intensity $\rho dx$, and the function $\phi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ satisfying $\phi(0) = 0, \phi(1) = 1$, and strict convexity. We denote by $K_\mathcal{V}$ the complete graph on $\mathcal{X}$. We assign to edges $e = \{q, q'\}$ connecting points $q$ and $q'$ the weights $\tau(e) = \phi(|q - q'|)$, and a natural choice is

$$\phi(x) = x^\alpha, \quad \alpha > 1. \hspace{1cm} (8)$$

The reason for $\alpha > 1$ is that the shortest path is otherwise the direct link, so this introduces nontrivial geodesics.

The first work on a Euclidean model of FPP concerned the Poisson-Voronoi tessellation of the $d$-dimensional Euclidean space by Vahidi-Asl and Wierman in 1992 [69]. This sort of generalization is a long term goal of FPP [21]. Much like the lattice model with independent and identically distributed weights, the model is rotationally invariant. The corresponding shape theorem, discussed in Ref. [21, Section 1], leads to a ball. The existence of bigeodesics (two paths, concatenated, which extend infinitely in two distinct directions from the origin, with the geodesic between the endpoints remaining unchanged), the linear rate of the local growth dynamics (the wetted region grows linearly with time), and the transversal fluctuations of the random path or surface are all summarized in Ref. [70].

**Bounds on the exponents**

Licea et al. [66] showed that for the standard first-passage percolation on $\mathbb{Z}^d$ with $d \geq 2$, the wandering exponent satisfies $\xi(d) \geq 1/2$ and specifically

$$\xi(2) \geq 3/5. \hspace{1cm} (9)$$

In Euclidean FPP, however, these bounds do not hold, and we have [71,72]

$$\frac{1}{d + 1} \leq \xi \leq \frac{3}{4} \hspace{1cm} (10)$$

and, for the wandering exponent,

$$\chi \geq \frac{1 - (d - 1)\xi}{2}. \hspace{1cm} (11)$$

Combining these different results then yields, for $d = 2$,

$$1/8 \leq \chi, \hspace{1cm} (12)$$

$$1/3 \leq \xi \leq 3/4. \hspace{1cm} (13)$$

Since the KPZ relation of Eq. (6) apparently holds in our setting, the lower bound for $\chi$ implies then a better bound for $\xi$, namely

$$\xi \geq \frac{3}{3 + d}. \hspace{1cm} (14)$$

which in the two-dimensional case leads to $\xi \geq 3/5$, the same result as in the standard FPP.

Also, the "rotational invariance" of the Poisson point process implies that the KPZ relation [Eq. (6)] is satisfied in each spatial network we study. We numerically corroborate this in Sec. IV. See, for example, Ref. [21, Section 4.3] for a discussion of the generality of the relation and the notion of rotational invariance.

C. EFPP on a spatial network

This is the model that we are considering here. Instead of taking, as in the usual EFPP, into account all possible edges with an exponent $\alpha > 1$ in Eq. (8), we allow only some edges between the points and take the weight proportional to their length (i.e., $\alpha = 1$ here). This leads to a different model but apparently universal properties of the geodesics. We therefore move beyond the weighted complete graph of Howard and Newman and consider a large class of spatial networks, including the random geometric graph (RGG), the $k$-nearest neighbor graph (NNG), the $\beta$ skeleton (BS), and the Delaunay triangulation (DT). We introduce them in Sec. III.

III. RANDOM SPATIAL NETWORKS

We consider in this study spatial networks constructed over a set of random points. We focus on the most straightforward case and consider a stationary Poisson point process in the $d$-dimensional Euclidean space, taking $d = 2$. This constitutes a Poisson random number of points, with expectation given by $\rho$ per unit area, distributed uniformly at random. We do not discuss here typical generalizations, such as to the Gibbs process, or Papangelou intensities [30].

First, we will consider the complete graph as in the usual EFPP, with edges weighted according to the details of Sec. II C. We will then consider the four distinct excluded region graphs defined below. Note that some of these networks actually obey inclusion relations, see, for example, Ref. [15]. We have

$$\text{RNG} \subset \text{GG} \subset \text{DT}, \hspace{1cm} (15)$$
where RNG stands for the relative neighborhood graph, GG the Gabriel graph, and DT the Delaunay triangulation. This nested relation trivially implies the following inequality:

\[ \xi_{\text{RNG}} \geq \xi_{\text{GG}} \geq \xi_{\text{DT}} \]  
(16)

as adding links can only decrease the wandering exponent. We are not aware of a similar relation for \( \chi \). We will also consider three distinct proximity graphs such as the hard and soft RGG and the \( k \)-nearest-neighbor graph.

### A. Proximity graphs

The main idea for constructing these graphs is that two nodes have to be sufficiently near in order to be connected.

#### 1. Random geometric graph

The usual random geometric graph is defined in Ref. [29] and was introduced by Gilbert [73] who assumes that points are randomly located in the plane and have each a communication range \( r \). Two nodes are connected by an edge if they are separated by a distance less than \( r \).

We also have the following variant: the soft random geometric graph [10,74,75]. This is a graph formed on \( \mathcal{X} \subset \mathbb{R}^d \) by adding an edge between distinct pairs of \( \mathcal{X} \) with probability \( H(|x-y|) \), where \( H : \mathbb{R}^+ \rightarrow [0,1] \) is called the connection function, and \( |x-y| \) is Euclidean distance.

We focus on the case of Rayleigh fading, where, with \( \gamma > 0 \) a parameter and \( \eta > 0 \) the path exponent, the connection function, with \( |x-y| > 0 \), is given by

\[ H(|x-y|) = \exp(-\gamma |x-y|^\eta) \]  
(17)

and is otherwise zero. This choice is discussed in Ref. [32, Section 2.3].

This graph is connected with high probability when the mean degree is proportional to the logarithm of the number of nodes in the graph. For the hard RGG, this is given by \( \rho \pi r^2 \), and otherwise the integral of the connectivity function over the region visible to a node in the domain, scaled by \( \rho \) [75]. As such, the graph must have a very large typical degree to connect.

#### 2. \( k \)-Nearest-neighbor graph

For this graph, we connect points to their \( k \in \mathbb{N} \) nearest neighbors. When \( k = 1 \), we obtain the nearest-neighbor graph (1-NNG), see, e.g., Ref. [76, Section 3]. The model is notably different from the RGG because local fluctuations in the density of nodes do not lead to local fluctuations in the degrees. The typical degree is much lower than the RGG when connected [76] though still remains disconnected on a random, countably infinite subset of the \( d \)-dimensional Euclidean space, since isolated subgraph exist. For large-enough \( k \), the graph contains the RGG as a subgraph. See Sec. VB for further discussion.

### B. Excluded region graphs

The main idea here for constructing these graphs is that two nodes will be connected if some region between them is empty of points. See Fig. 4 for a depiction of the geometry of the lens regions for \( \beta \) skeletons.

---

FIG. 4. The geometry of the lune-based \( \beta \) skeleton for (a) \( \beta = 1/2 \), (b) \( \beta = 1 \), and (c) \( \beta = 2 \). For \( \beta < 1 \), nodes within the intersection of two disks each of radius \(|x-y|/2\beta\) preclude the edges between the disk centers, whereas for \( \beta > 1 \), we use instead radii of \( \beta|x-y|/2 \). Thus, whenever two nodes are nearer each other than any other surrounding points, they connect and otherwise do not.

### IV. NUMERICAL RESULTS

#### A. Numerical setup

Given the models in the previous section, we numerically evaluate the scaling exponents \( \chi \) and \( \xi \), as well as the distribution of the travel-time fluctuations. We now describe the numerical setup. With density of points \( \rho > 0 \), and a small
relative neighborhood graph). The point process density $\rho$ distinguished with different colours (green and blue), as are EFPP on the complete graph, the DT, and the two $\beta$ skeletons (Gabriel graph and relative neighborhood graph). The point process density $\rho$ points per unit area is given for each model.

We then increase $w$, in steps of three units of distance, and repeat until we have statistics of all $w$ to the limit of computational feasibility. This varies slightly among models. The RGGs are more difficult to simulate due to their known connectivity constraint where vertex degrees must approach infinity, see, e.g., Ref. [29, Chapter 1]. Thus we cannot simulate connected graphs to the same limits of Euclidean span as with the other models.

We are then able to relate the mean and standard deviation of the passage time, as well as the mean wandering, to $w$, at various $\rho$, and for each model. For example, the left hand plots in Fig. 5 show that the typical passage time $ET(x, y)$ grows linearly with $w$, i.e., grows linearly with $w$, for all networks [10,14,15]. The standard error is shown but is here not clearly distinguishable from the symbols.

We ensure $h$ is large enough to stop the geodesics hitting the boundary, so we use a domain of height equal to the mean deviation $\mathbb{E}D(w)$, plus six standard deviations.

The key computational difficulty here is the memory requirement for large graphs, of which all $N$ are stored simultaneously, and mapped in parallel on a Linux cluster over a function which measures the path statistic. This parallel processing is used to speed up the computation of the geodesics lengths and wandering.

B. Scaling exponents

The results are shown in Fig. 5. These plots, shown in loglog, reveal a power-law behavior of $T$ and $D$, and the linear growth of typical travel time with Euclidean span. We then compute the exponents to two significant figures using a nonlinear model fit, based on the model $a|x - y|^b$, and then determine the parameters $a$, $b$ using the quasi-Newton method in Mathematica 11.

Our numerical results suggest that we can distinguish two classes of spatial network models by the scaling exponents...
of their Euclidean geodesics. The proximity graphs (hard and soft RGG and $k$-NNG) are in one class, with exponents

\[ \chi_{\text{RGG,NNG}} = 0.20 \pm 0.01, \]  
\[ \xi_{\text{RGG,NNG}} = 0.60 \pm 0.01, \]  

whereas the excluded region graphs (the $\beta$ skeletons and Delaunay triangulation), and Howard’s EFPP model with $\alpha > 1$, are in another class with

\[ \chi_{\text{DT,}\beta\text{-skel,EFPP}} = 0.40 \pm 0.01, \]  
\[ \xi_{\text{DT,}\beta\text{-skel,EFPP}} = 0.70 \pm 0.01. \]  

Clearly, the KPZ relation of Eq. (6) is satisfied up to the numerical accuracy which we are able to achieve. We corroborate that this is independent of the density of points and connection range scaling, given the graphs are connected. The exponents hold asymptotically, i.e., large interpoint distances. Thus we conjecture

\[
\begin{align*}
\text{Var}[T(x,y)] &\sim |x-y|^{2/5}, \\
\mathbb{E}[D(x,y)] &\sim |x-y|^{7/10},
\end{align*}
\]

for the proximity graphs (the DT and the $\beta$ skeletons for all $\beta$), and, for the RGGs and the $k$-NNG,

\[
\begin{align*}
\text{Var}[T(x,y)] &\sim |x-y|^{2/5}, \\
\mathbb{E}[D(x,y)] &\sim |x-y|^{3/5}.
\end{align*}
\]

We summarize these new results in Table I. It is surprising that these exponents are apparently rational numbers. In Bernoulli continuum percolation, for example, the threshold connection range for percolation is not known but not thought to be rational, as it is with bond percolation on the integer lattice [29, Chapter 10]. Exact exponents are not necessarily expected in the continuum setting of this problem, which suggests there is more to be said about the classification of first-passage process via this method.

<table>
<thead>
<tr>
<th>Network</th>
<th>$\xi$</th>
<th>$\chi$</th>
<th>Distribution of $T$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proximity graphs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soft RGG</td>
<td>3/5</td>
<td>1/5</td>
<td>Normal (Conj.)</td>
</tr>
<tr>
<td>Hard RGG</td>
<td>3/5</td>
<td>1/5</td>
<td>Normal (Conj.)</td>
</tr>
<tr>
<td>$k$-NNG</td>
<td>3/5</td>
<td>1/5</td>
<td>Normal</td>
</tr>
<tr>
<td><strong>Excluded region graphs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DT</td>
<td>7/10</td>
<td>2/5</td>
<td>Normal</td>
</tr>
<tr>
<td>GG</td>
<td>7/10</td>
<td>2/5</td>
<td>Normal</td>
</tr>
<tr>
<td>$\beta$ skeletons</td>
<td>7/10</td>
<td>2/5</td>
<td>Normal</td>
</tr>
<tr>
<td>RNG</td>
<td>7/10</td>
<td>2/5</td>
<td>Normal</td>
</tr>
<tr>
<td><strong>Euclidean FPP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With $\alpha = 3/2$</td>
<td>7/10</td>
<td>2/5</td>
<td>Normal</td>
</tr>
<tr>
<td>With $\alpha = 5/2$</td>
<td>7/10</td>
<td>2/5</td>
<td>Normal</td>
</tr>
</tbody>
</table>

C. Travel-time fluctuations

We see numerically that the travel-time distribution is a normal for most cases (see Fig. 6). We summarize these results in Table I and in Fig. 7 we show the skewness and kurtosis for the travel-time fluctuations, computed for the different networks. For a Gaussian distribution, the skewness is 0 and the kurtosis equal to 3, while the Tracy-Widom distribution displays other values.

We provide some detail of the distribution of $T$ for each model from the proximity class in Fig. 6. This is compared against four test distributions, the Gaussian orthogonal, unitary, and symplectic Tracy Widom distributions, and the standard normal distribution.

This makes the case of EFPP on spatial networks one of only a few special cases where Gaussian fluctuations in fact occur. Auffinger and Damron go into detail concerning each of the remaining cases in Ref. [21, Section 3.7]. One example, reviewed extensively by Chaterjee and Dey [36], is when geodesics are constrained to lie within thin cylinders, i.e., ignore paths which traverse too far, and thus select the minima from a subdomain. This result could shed some light on their questions, though in what way it is not clear.

We also highlight that Tracy-Widom is thought to occur in problems where matrices represent collections of totally uncorrelated random variables [77]. In the case of EFPP, we have the interpoint distances of a point process, which lead to spatially correlated interpoint distances, so the adjacency matrix does not contain independent and identically distributed values. This potentially leads to the loss of Tracy-Widom. However, we also see some cases of $N \times N$ large complex correlated Wishart matrices leading to TW for at least one of their eigenvalues and with convergence at the scale $N^{2/3}$ [78].

D. Transversal fluctuations

The transversal deviation distribution results appear beside our evaluation of the scaling exponents, in Fig. 8. All the models produce geodesics with the same transversal fluctuation distribution, despite distinct values of $\xi$. The fluctuations are also distinct from the Brownian bridge (a geometric Brownian motion constrained to start and finish at two fixed position vectors in the plane), running between the midpoints of the boundary arcs [19]. It is a key open question to provide some
FIG. 6. Travel-time distributions for the DT [(a)–(c)], RNG [(d)–(f)], and Gabriel [(g)–(i)] graphs, compared with the GUE and GSE Tracy-Widom ensembles, and the Gaussian distribution. The point process density $\rho$ points per unit area is given for each model. The slight skew of the TW distribution is not present in the data.

information about this distribution, as it is rarely studied in any FPP model, as far as we are aware of the literature. A key work is Kurt Johannson’s, where the wandering exponent is derived analytically in a variant of oriented first-passage percolation. One might ask whether a similar variant of EFPP might be possible [52].

V. DISCUSSION

The main results of our investigation are the new rational exponents $\chi$ and $\xi$ for the various spatial models, and the discovery of the unusual Gaussian fluctuations of the travel time. We found that for the different spatial networks the KPZ

FIG. 7. Skewness (a) and kurtosis (b), for the travel-time fluctuations, computed for each network model. For a Gaussian distribution, the skewness is 0 and the kurtosis equal to 3, values that we indicate by dashed black lines. The point process density $\rho$ points per unit area is given for each model. The Tracy-Widom distribution has only marginally different moments to the normal, also shown by dashed black lines, with labels added to distinguish each specific distribution (GOE, GUE, or GSE), as well as the Gaussian.
of questions that we think are important. Further, we have topics of further research which may shed light on the first-

3

7

otherwise, we have where the “length of a nonedge” is the corresponding inter-

range, i.e., in terms of the longest edge, and shortest nonedge,

3]. So the computations used to produce these graphs and

relation holds and known bounds are satisfied. Also, due to

points per unit area is given for

Brownian bridge process between the same end points (red dashed

curve). The point process density

points at a fixed Euclidean distance,

DT, \(x-y\)=150, \(\rho=6\)

Gabriel, \(x-y\)=150, \(\rho=6\)

Hard RGG, \(x-y\)=150, \(\rho=6\)

RNG, \(x-y\)=150, \(\rho=6\)

Soft RGG, \(x-y\)=150, \(\rho=6\)

Brownian Bridge over [0,1]

FIG. 8. Transversal fluctuations of the geodesics in all models

colored points) and compared with the fluctuations of a continuous

Brownian bridge process between the same end points (red dashed

curve). The point process density \(\rho\) points per unit area is given for
each model.

in the necessary dense limit, so we are unable to verify the
fluctuations of either \(T\) or \(D\). However, we can see a skewness
and kurtosis for \(T(|x-y|)\) which are monotonically decreasing
with \(|x-y|\) toward the hypothesized limiting Gaussian
statistics, at least for the limited Euclidean span we can achieve.

Given that \(k\)-NNG is in the same class, we are left to
conjecture whether Gaussian fluctuations hold throughout all
the spatial models described in Sec. III. It remains an open
question to identify any exceptional models where this does
not hold.

B. Percolation and connectivity

If we choose two points at a fixed Euclidean distance,
then simulate a Poisson point process in the rest of the \(d\)
dimensional plane, construct the relevant graph, and consider
the probability that both points are in the giant component;
this is effectively a positive constant for reasonable distances,
assuming that we are above the percolation transition. At
small distances, the two events are positively correlated. Thus,
one can condition on this event and therefore, when simulat-
ing, discount results where the Euclidean geodesic does not
exist. This defines FPP on the giant component of a random
graph.

It is not clear from our experiments whether the rare
isolated nodes, or occasionally larger isolated clusters, either
in the RGGs or \(k\)-NNG, affect the exponents. One similar
model system would be the Lorentz gas: Put disks of constant
radius in the plane, perhaps at very low density, and seek the
shortest path between two points that does not intersect the
disks. The exponents \(\chi\) and \(\xi\) for this setting are not known
[19,79].

An alternative to giant component FPP would be to con-
tion on the two points being connected to each other. This
would be almost identical for the almost connected regime
but weird below the percolation transition. In that case the
event we condition on would have a probability decaying
exponentially with distance, and the point process would end
up being extremely special for the path to even exist. For
example, an extremely low-density RGG would be almost
empty apart from a path of points connecting the end points,
with a minimum number of hops.

C. Betweenness centrality

The extent to which nodes take part in shortest paths
throughout a network is known as betweenness centrality
[1,4]. We ask to what extent knowledge of wandering can
lead to understanding the centrality of nodes. The variant node
shortest path betweenness centrality is highest for nodes in
bottlenecks. Can this centrality index be analytically under-
stood in terms of the power-law scaling of \(D\)? Is the exponent
directly relevant to its large-scale behavior?

In order to illustrate more precisely this question, let \(G\)
be the graph formed on a point process \(X\) by joining pairs of
points with probability \(H(|x-y|)\). Consider \(\sigma_{xy}\) to be the
number of shortest paths in \(G\) which join vertices \(x\) and \(y\) in
\(G\) and \(\sigma_{xy}(z)\) to be the number of shortest paths which join \(x\)
to \(y\) in \(G\), but also run through \(z\); then, with \(\neq\) indicating a

\[\frac{3}{3+d} \leq \xi \leq \frac{3}{4}.\] (28)

It is surprising to find a large class of networks, in particular
the Delaunay triangulation, that displays an exponent \(\xi = 7/10\) and points to the question of the existence of another
class of graphs which display the theoretically maximal \(\xi = 3/4\).

Both immediately present a number of open questions and
topics of further research which may shed light on the first-
passage process on spatial networks. We list below a number of
questions that we think are important.

A. Gaussian travel-time fluctuations

We are not able to conclude that all the models in the
proximity graph class \(\chi = 3/5, \xi = 1/5\), have Gaussian fluc-
tuations in the travel time. This is for a technical reason. All
the models we study are either connected with probability 1,
such as the DT or \(\beta\) skeleton with \(\beta \leq 2\), or have a connection
probability which goes to 1 in some limit. We require con-
nected graphs, or paths do not span the boundary arcs, and
the exponents are not well defined.

Thus, the difficult models to simulate are the HRGG,
SRGG, and \(k\)-NNG, since these are in fact disconnected with
probability 1 without infinite expected degrees, i.e., the dense
limit of Penrose, see Ref. [29, Chapter 1], or with the fixed
degree of the \(k\)-NNG \(k = \Theta(\log n)\) and \(n \rightarrow \infty\) in a domain
with fixed density and infinite volume. Otherwise, we have
isolated vertices, or isolated subgraphs, respectively.

However, the \(k\)-NNG has typically shorter connection
range, i.e., in terms of the longest edge, and shortest nonedge,
where the “length of a nonedge” is the corresponding inter-
point distance between the disconnected vertices [76, Section
3]. So the computations used to produce these graphs and
then evaluate their statistical properties are significantly less
demanding. Thus, the HRGG is computationally intractable
sum over unordered pairs of vertices not including \( z \), define the betweenness centrality \( g(z) \) of some vertex \( z \) in \( G \) to be
\[
g(z) = \sum_{i\neq j \neq k} \frac{\sigma_{ij}(z)}{\sigma_{ij}}. \tag{29}
\]

In the continuous limit for dense networks we can discuss the betweenness centrality and we recall some of the results in Ref. [11]. More precisely, we define \( \chi_{xy}(z) \) as the indicator which gives unity whenever \( z \) intersects the shortest path connecting the \( d \)-dimensional positions \( x, y \in \mathcal{V} \). Then the normalized betweenness \( g(z) \) is given by
\[
g(z) = \frac{1}{\int_{\mathcal{V}^2} \chi_{xy}(0)dxdy} \int_{\mathcal{V}^2} \chi_{xy}(z)dxdy. \tag{30}
\]

Based on the assumption that there exists a single topological geodesic as \( \rho \to \infty \), and that this limit also results in an infinitesimal wandering of the path from a straight line segment, an infinite number of points of the process lying on this line segment intersect the topological geodesic as \( \rho \to \infty \), assuming the graph remains connected, and so \( \chi_{xy}(z) \) can then written as a \( \delta \) function of the transverse distance from \( z \) to the straight line from \( x \) to \( y \). The betweenness can then be computed and we obtain [11] [normalized by its maximum value at \( g(0) \)]
\[
g(\epsilon) = \frac{2}{\pi} (1 - \epsilon^2) E(\epsilon), \tag{31}
\]
where \( E(k) = \int_0^{\pi/2} d\theta [1 - k^2 \sin^2(\theta)]^{1/2} \) is the complete elliptic integral of the second kind. We have also rescaled such that \( \epsilon \) is in units of \( R \).

Take \( D(x, y) \) to be the maximum deviation from the horizontal of the Euclidean geodesic. Numerical simulations suggest that
\[
\mathbb{E} D(x, y) = C|\mathbf{x} - \mathbf{y}|^\xi, \tag{32}
\]
where the expectation is taken over all point sets \( \mathcal{X} \). The “thin cylinders” are given by a Heaviside \( \Theta \) function, so assume that the characteristic function is no longer a \( \delta \) spike but a wider cylinder,
\[
\chi_{xy}(z) = \Theta[D(x, y) - |z|], \tag{33}
\]
where \( z_\perp \) is the magnitude of the perpendicular deviation of the position \( z \) from \( \text{hull}(x, y) \). We then have that
\[
g(z) = \frac{1}{\int_{\mathcal{V}^2} \theta(D - \|z\|)dxdy} \int_{\mathcal{V}^2} \theta(D - |z|)dV \tag{34}
\]
(where \( \mathbf{0} \) is the transverse vector computed for the origin). This quantity is certainly difficult to estimate but provides a starting point for computing finite-density corrections to the result of Ref. [11].

The boundary of the domain is crucial in varying the centrality, which is something we ignore here. Without an enclosing boundary, such as with networks embedded into spheres or tori, the typical centrality at a position in the domain is uniform, since no point is clearly distinguishable from any other. This is discussed in detail in Ref. [11]. In fact, a significant amount of recent work on random geometric networks has highlighted the importance of the enclosing boundary [32, 74].

VI. CONCLUSIONS

We have shown numerically that there are two distinct universality classes in Euclidean first-passage percolation on a large class of spatial networks. These two classes correspond to the following two broad classes of networks: first, based on joining vertices according to critical proximity, such as in the RGG and the NNG, and, second, based on graphs which connect vertices based on excluded regions, as in the lune-based \( \beta \) skeletons or the DT. Heuristically, the most efficient way to connect two points is via the nearest neighbor, which suggests that \( \xi \) for proximity graphs should on the whole be smaller than for exclusion-based graphs, which is in agreement with our numerical observations.

The passage times show Gaussian fluctuations in all models, which we are able to numerically verify. This is a clear distinction between EFPP and FPP. After similar results of Chatterjee and Dey [36], it remains an open question why this happens and also of course how to rigorously prove it.

We also briefly discussed notions of the universality of betweenness centrality in spatial networks, which is influenced by the wandering of shortest paths. A number of open questions remain about the range of possible universal exponents which could exist on spatial networks, whose characterization would shed light on the interplay between the statistical physics of random networks, and their spatial counterparts, in way which could reveal deep insights about universality and geometry more generally.

All underlying data are reproduced in full within the paper.

ACKNOWLEDGMENTS

The authors thank Márton Balázs and Bálint Tóth for a number of very helpful discussions, as well as Ginestra Bianconi at QMUL, Jürgen Jost at MPI Leipzig, and the School of Mathematics at the University of Bristol, who provided generous hosting for APKG while carrying out various parts of this research. This work was supported by the EPSRC project “Spatially Embedded Networks” (Grant No. EP/N002458/1). A.P.K.G. was partly supported by the EPSRC project “Random Walks on Random Geometric Networks” (Grant No. EP/N508767/1).


[77] https://mathoverflow.net/questions/71306/when-should-we-expect-tracy-widom
