Short-length routes in low-cost networks *via* Poisson line patterns

David J. Aldous* and Wilfrid S. Kendall

Abstract

In designing a network to link n cities in a square of area n, one might be guided by the following two desiderata. First, the total network length should not be much greater than the length of the shortest network connecting all cities. Second, the average route length (taken over source-destination pairs) should not be much greater than the average straight-line distance. How small can we make these two differences? For typical configurations the shortest network length is order n and the average straight-line distance is order $n^{1/2}$, so it seems implausible that one can construct a network in which the first difference is $o(n)$ and the second difference is $o(n^{1/2})$. But in fact one can do better: for an arbitrary configuration one can construct a network where the first difference is $o(n)$ and the second difference is almost as small as $O(\log n)$. The construction is conceptually simple: over the minimum-length connected network (Steiner tree) superimpose a sparse stationary and isotropic Poisson line process. The key ingredient is a new result about the Poisson line process. Consider two points at distance r apart, and delete from the line process all lines which separate these two points. The resulting pattern of lines partitions the plane into cells; the cell containing the two points has mean boundary length $\approx 2r + \text{constant} \times \log r$. Turning to lower bounds we show that, under a weak equidistribution assumption, if the first difference is $o(n)$ then the second difference cannot be $O(\sqrt{\log n})$.

MSC 2000 subject classifications: Primary 60D05, 90B15

Key words and phrases: Buffon argument; excess statistic; mark distribution; spatial network; Poisson line process; ratio statistic; Slivynak theorem; Steiner tree; Vasershtein coupling; total variation distance

Short title: Lengths and costs in networks

1 Introduction

We start with a counter-intuitive observation and its motivation, which prompted us to probe more deeply into the underlying question.

*Research supported by N.S.F Grant DMS-0203062

Consider n points ("cities", say) in a square of area n. We are interested in both the worst-case setting where the city positions are arbitrary, and the average case setting where the city points are random, independent and uniformly distributed. Consider a connected network (a road network, say) of straight line segments linking these city points and perhaps other junction points. Recall that the minimum length connected network on a configuration of points $\mathbf{x}^n = \{x_1, \ldots, x_n\}$ is the *Steiner tree* $\mathrm{ST}(\mathbf{x}^n)$.

It is well known and straightforward to prove (cf. Steele 1997, Yukich 1998) that in both the worst case and the average case the total network length $\mathrm{len}(\mathrm{ST}(\mathbf{x}^n))$ grows as order $O(n)$. When designing a network, it is reasonable to regard total network length as a "cost". A natural corresponding "benefit" would be the existence (in some average sense) of short routes between city points. Let $\ell(x_i, x_j)$ be the route-length (length of shortest path) between points x_i and x_j in a given network, and let $\mathrm{dist}(x_i, x_j) = |x_i - x_j|$ denote Euclidean distance (so $\ell(x_i, x_j) \geq \mathrm{dist}(x_i, x_j)$). A good network should possess the following

> **Short routes property:** For typical pairs (i, j), the route-length $\ell(x_i, x_j)$ between city points x_i and x_j is not much larger than the Euclidean distance $\mathrm{dist}(x_i, x_j)$.

A first take on a statistic to measure this property for a connected network $G(\mathbf{x}^n)$ is the *ratio statistic*, based on averaging the ratios of network route-lengths *versus* Euclidean distances. Consider a network $G(\mathbf{x}^n)$ to be the configuration of city points $\mathbf{x}^n = \{x_1, \ldots, x_n\}$ together with a collection of line segments which combine to connect every city x_i to every other city x_j.

Definition 1 (Ratio statistic)**.** Let $\mathrm{average}_{(i,j)}$ denotes the average over all distinct pairs (i, j). Then

$$\mathrm{ratio}(G(\mathbf{x}^n)) \quad = \quad \underset{(i,j)}{\mathrm{average}} \frac{\ell(x_i, x_j)}{d(x_i, x_j)} - 1 \quad \geq \quad 0\,. \tag{1}$$

Consider a network $G(\mathbf{x}^n)$ based on n uniform random points $\mathbf{x}^n \subset [0, \sqrt{n}]^2$, having (say) twice the total length of the Steiner tree. Initially we speculated that in this case the expectation $\mathbb{E}\left[\mathrm{ratio}(G(\mathbf{x}^n))\right]$ would converge to some strictly positive constant as $n \to \infty$. However this intuition is wrong (see section 5.3):

> **Counterintuitive observation:** On well-dispersed configurations, it is possible to construct networks whose total lengths are greater than the corresponding Steiner tree lengths by only an asymptotically negligible factor, but for which the ratio statistic converges to zero as total network length converges to infinity.

Motivation for these considerations arises from analysis of real-world networks. Consider for example the "core" part of the U.K. rail network linking the 40 largest cities. The real network has a certain total length and a certain value for some statistic R devised to capture the "short routes" property. Even though the real network evolved via a complex historical process, one can study whether it is close to optimal, in the sense of whether its value of R is close to the minimum possible value of R over all possible networks of the same total length. So the issue arises of what statistic R best captures the imprecisely expressed "short routes" property, and one can investigate this issue by theoretical study of different statistics in the random points model. We interpret the counterintuitive observation above as implying that the ratio$(\cdot)$ statistic of Definition 1 is probably *not* a good choice of statistic, because we prove this observation by constructing networks which are approximately optimal by this criterion and yet are plainly rather different from many plausible real-world networks. What *is* a good choice of statistic will be discussed in a companion paper, along with the U.K. rail example.

Informally, the counter-intuitive observation suggests that we can construct networks for configurations of n points which have total network length exceeding that of the Steiner tree by just $o(n)$, and such that the average excess of network distance over Euclidean distance is $o(n^{1/2})$ (bearing in mind that average Euclidean distance for "evenly spread out" configurations should be $O(n^{1/2})$). In fact much more is true: the observation holds on an additive scale at almost $O(\log n)$, even in "worst case" scenarios:

Definition 2 (Excess average length for a network)**.** The *excess route length* for a network $G(\mathbf{x}^n)$ is

$$\text{excess}\,(G(\mathbf{x}^n)) \quad = \quad \underset{(i,j)}{\text{average}}\,(\ell(x_i, x_j) - \text{dist}(x_i, x_j))\ . \tag{2}$$

Theorem 3 (Upper bound on minimum excess network length)**.** *For each n let $\boldsymbol{x}^n$ be an arbitrary configuration of n city points in a square of area n.*

(a) Let $w_n \to \infty$. There exist networks $G(\boldsymbol{x}^n)$ connecting up the cities such that

(i) $\text{len}(G(\boldsymbol{x}^n)) - \text{len}(\text{ST}(\boldsymbol{x}^n)) = o(n)$;

(ii) $\text{excess}(G(\boldsymbol{x}^n)) = o(w_n \log n)$.

(b) Let $\varepsilon > 0$. There exist networks $G(\boldsymbol{x}^n)$ connecting up the cities such that

(i) $\text{len}(G(\boldsymbol{x}^n)) - \text{len}(\text{ST}(\boldsymbol{x}^n)) \le \varepsilon n$;

(ii) $\text{excess}(G(\boldsymbol{x}^n)) = O(\log n)$.

This result is proved in Sections 2 and 3. The idea is to build a hierarchical network. At small scales routes use the underlying Steiner tree. At large scales, routes use a sparse collection of randomly oriented lines (a realization of a stationary and isotropic *Poisson line process*); this is the key ingredient that permits an excess of at most $O(\log(n))$ (Section 2). We believe that only these two scales are needed, but to simplify analysis (so as to avoid non-elementary analysis of Steiner trees) we introduce an intermediate scale consisting of a widely-spaced grid. Thus a route from an originating city navigates through the

Steiner tree to a grid line and then along the grid line to a line of the Poisson line process, and then navigates in the reverse sense down to the destination city. (For technical reasons we also introduce occasional small rectangles to permit circumnavigation around Steiner tree "hot-spots" (Section 3)). The key ingredient in the analysis is a calculation concerning the Poisson line process, which has separate interest as a result in stochastic geometry (Theorem 7 below). Consider two points at distance r apart, and delete from the line process all lines which separate these two points. The resulting pattern of lines partitions the plane into cells; the cell containing the two points has mean boundary length which for large r is asymptotic to $2r + \text{constant} \times \log r$.

For *lower bounds* it is necessary to impose some condition on the empirical distribution of the city points in $\mathbf{x}^n$, since if all the city points concentrate on a line then the excess is zero! We need a quantitative condition on equidistribution of city points over a region, formalized via the following truncated *Vasershtein coupling* scheme.

Definition 4 (Quantitative equidistribution condition)**.** Let $\mathbf{x}^n$ be a configuration in the plane, μ^n be a probability measure on the plane, and $L_n > 0$. Say *$\boldsymbol{x}^n$ is L_n-equidistributed as* μ^n if there exists a coupling of random variables (X_n, Y_n) such that

(a) X_n has uniform distribution on the finite point-set $\mathbf{x}^n$,

(b) Y_n has distribution μ^n,

(c) $\mathbb{E}\left[\min\left(1, \frac{|X_n - Y_n|}{L_n}\right)\right] \to 0$ as $n \to \infty$.

A sufficient condition for the following result is that $\mathbf{x}^n$ is L_n-equidistributed as the uniform distribution on the square of area n, for some $L_n = o(\sqrt{\log n})$. The purpose of introducing the *non*-uniform distribution μ^n in Definition 4 is to permit us to express Theorem 5 below in terms of weaker and more local conditions: for example a consequence of Theorem 5(b) is that we may replace the *uniform* reference distribution by any distribution μ on $[0,1]^2$ with a continuous density component, rescaled to produce a distribution μ^n on $[0, n^{1/2}]^2$.

Theorem 5 (Lower bound on minimum excess network length)**.** *Let $\boldsymbol{x}^n$ be a configuration of city points in a square $[0, \sqrt{n}]^2$. Let $L_n = o(\sqrt{\log n})$. Suppose either*

(a) $\boldsymbol{x}^n$ is L_n-equidistributed as the uniform distribution on the square of area n;

or (more generally)

(b) for some fixed ρ and ε, there is a subcollection $\boldsymbol{y}^{k(n)}$ of $k(n)$ city points, all lying in a disk D_n of area $\pi\rho n$, such that $k(n) > \pi\rho n\varepsilon$, and such that $\boldsymbol{y}^{k(n)}$ is L_n-equidistributed as the uniform distribution on D_n.

Let $G(\boldsymbol{x}^n)$ be a network based on the full collection of n city points. If $\text{len}(G(\boldsymbol{x}^n))/n$ remains bounded as $n \to \infty$, then

$$\text{excess}(G(\boldsymbol{x}^n)) \quad = \quad \Omega(\sqrt{\log n})\,. \tag{3}$$

Configurations $\mathbf{x}^n$ produced by independent uniform sampling from $[0, \sqrt{n}]^2$ satisfy the conditions of this theorem (see Remark 15). The proof of the theorem is given in Section 4, and exploits a tension between the two following facts:

(a) A short route between x_i and x_j must run approximately parallel to the Euclidean geodesic, and hence will tend to make almost orthogonal intersections with random segments perpendicular to this geodesic.

(b) On the other hand, the equidistribution condition means that two city points x_i and x_j randomly chosen from the subcollection must be nearly independent uniform draws from D_n, which permits the derivation of *upper bounds* on the probability of nearly orthogonal intersections of the form given in fact (a).

Finally note that the assumption $\text{len}(G(\mathbf{x}^n)/n$ remains bounded as $n \to \infty$ in the lower bound is weaker than the corresponding assumption $\text{len}(G(\mathbf{x}^n)) - \text{len}(\text{ST}(\mathbf{x}^n)) \leq \varepsilon n$ in the upper bound, but we are unable to improve (3) under the stronger assumption.

2 The Poisson line process network

Our upper bound on minimal excess $(G(\mathbf{x}^n))$ is based on a result from stochastic geometry (Theorem 7 below) which is of independent interest.

Recall that a Poisson line process in the plane $\mathbb{R}^2$ is constructed as a Poisson point process whose points lie in the space which parametrizes the set of lines in the plane. We will consider only undirected lines, which will be parametrized by $(r, \theta) \in \mathbb{R} \times [0, \pi)$ where r is the signed distance from the line to a reference point and θ is the angle the line makes with a reference axis. A stationary and isotropic Poisson line process has intensity measure invariant under rotations and translations of $\mathbb{R}^2$: a stationary and isotropic Poisson line process Π of unit intensity is one for which the number of lines of Π hitting a unit segment has expectation 1 (further facts about Poisson line processes may be found in Stoyan et al. 1995, Chapter 8). We are interested in the cell containing two fixed points which is formed by the lines of Π that do not separate the two points, because this can be used as the efficient long-distance part of a network route between the two points (see Lemma 11). Theorem 7 establishes an asymptotic upper bound for the length of the mean cell perimeter in case of wide separation between the two points; we prepare for this by using a Buffon argument to derive an exact double-integral expression for the mean cell perimeter length:

Theorem 6 (Mean perimeter length)**.** *Let Π be a stationary and isotropic Poisson line process of unit intensity. Fix two points v_i, v_j which are distance m*

apart. Delete the lines of Π *which separate the two points* v_i*,* v_j*. The remaining line pattern partitions the plane: the cell* $\mathcal{C}(v_i, v_j)$ *containing the two fixed points has mean perimeter* $\mathbb{E}[\operatorname{len}\partial\mathcal{C}(v_i, v_j)] = 2m + J_m$*, where* J_m *is given by the double integral*

$$\begin{aligned} J_m &= \mathbb{E}[\operatorname{len}\partial\mathcal{C}(v_i, v_j)] - 2m \\ &= \frac{1}{2}\iint_{\mathbb{R}^2} (\phi - \sin\phi)\exp\left(-\tfrac{1}{2}(\eta - m)\right)\operatorname{Leb}(\mathrm{d}x)\,. \end{aligned} \tag{4}$$

Here $\eta = \eta(x)$ *is a sum of distances* $\operatorname{dist}(v_i, x) + \operatorname{dist}(v_j, x)$*, while* $\phi = \phi(x)$ *is the exterior angle at* x *of the triangle with vertices* x*,* v_i*,* v_j *(see Figure 1).*

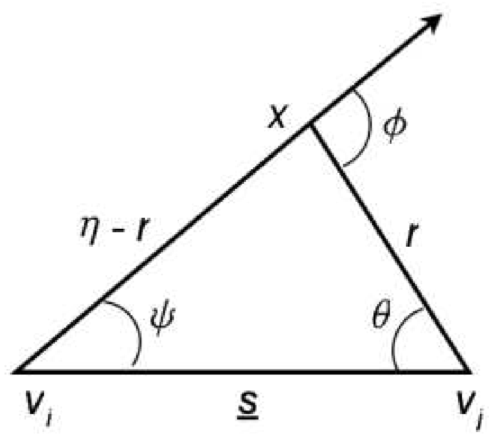

Figure 1: Illustration of definition of η and ϕ. Note that ϕ is the sum of the two interior angles ψ and θ.

Proof. This proof can be phrased in terms of measure-theoretic stochastic geometry, using the language of Palm distributions and Campbell measure. Since we deal only with constructions based on Poisson processes, we are able to adopt a less formal but more transparent exposition, for the sake of a wider readership.

Let $\underline{s}$ be the line segment of length m with end-points v_i, v_j. The idea of the proof is to measure $\mathbb{E}[\operatorname{len}\partial\mathcal{C}(v_i, v_j)]$ by computing the expected number of hits on $\partial\mathcal{C}(v_i, v_j)$ made by an *independent* homogeneous isotropic Poisson line process $\widetilde{\Pi}$, again of unit intensity. Each hit corresponds to one of the points in the *intersection point process* $\mathcal{X} = \{\iota(\ell, \widetilde{\ell}) : \ell \in \Pi, \widetilde{\ell} \in \widetilde{\Pi}\}$, where

$$\iota(\ell, \widetilde{\ell}) = \begin{cases} x & \text{if } \ell \cap \widetilde{\ell} = \{x\}\,, \\ \text{undefined} & \text{if } \ell, \widetilde{\ell} \text{ are parallel.} \end{cases} \tag{5}$$

Note that with probability 1 the intersection point $\iota(\ell, \widetilde{\ell})$ is defined for all $\ell \in \Pi$, $\widetilde{\ell} \in \widetilde{\Pi}$.

Not all points $x \in \mathcal{X}$ correspond to hits on $\partial\mathcal{C}(v_i, v_j)$. The condition for $x = \iota(\ell, \widetilde{\ell}) \in \mathcal{X}$ to be a hit on $\partial\mathcal{C}(v_i, v_j)$ is that either $\widetilde{\ell}$ hits $\underline{s}$ or x is not separated from $\underline{s}$ by any line from $\Pi \setminus \{\ell\}$. The Slivynak theorem (Stoyan, Kendall, and Mecke 1995, §4.4, example 4.3) implies that $\Pi \setminus \{\ell\}$ conditional on $\ell \in \Pi$ is itself a homogenous isotropic unit-rate Poisson line process; consequently if $\widetilde{\ell}$ does not hit $\underline{s}$ then the probability that $x = \iota(\ell, \widetilde{\ell}) \in \mathcal{X}$ is a hit on $\partial\mathcal{C}(v_i, v_j)$ is equal to the probability $p(x)$ that there is no line in Π which cuts both the segment from v_i to x and the segment from v_j to x (note that such a line would *not* cut the segment $\underline{s}$).

A classic counting argument from stochastic geometry then reveals that

$$p(x) \;=\; \exp\left(-\tfrac{1}{2}\left(\operatorname{dist}(v_i, x) + \operatorname{dist}(v_j, x) - m\right)\right) \quad = \quad \exp\left(-\tfrac{1}{2}(\eta - m)\right) . \quad (6)$$

Accordingly, if ν is the intensity of the point process $\mathcal{X}$ then we may compute the mean number of hits on $\partial\mathcal{C}(v_i, v_j)$ as

$$2m \;+\; \iint_{\mathbb{R}^2} \nu \, \mathbb{P}\left[\ell \not\pitchfork \underline{s}, \widetilde{\ell} \not\pitchfork \underline{s} \,|\, x = \iota(\ell, \widetilde{\ell}) \in \mathcal{X}\right] \exp\left(-\tfrac{1}{2}(\eta - m)\right) \operatorname{Leb}(\mathrm{d}x). \quad (7)$$

Here "$\ell \not\pitchfork \underline{s}$" stands for "the line ℓ does not hit $\underline{s}$" – noting that the conditioning in this context forces the Poisson line ℓ to pass through x but does not fix its orientation – and the summand $2m$ corresponds to the fact that hits of $\widetilde{\Pi}$ on $\underline{s}$ count as automatic hits on $\partial\mathcal{C}(v_i, v_j)$.

Condition on $x = \iota(\ell, \widetilde{\ell}) \in \mathcal{X}$ (which is to say, condition on there being Poisson lines $\ell \in \Pi$, $\widetilde{\ell} \in \widetilde{\Pi}$ both passing through x) and consider

(a) the angle ξ_1 of ℓ;

(b) the angle ξ_2 between ℓ and $\widetilde{\ell}$.

By isotropy of Π the random angle ξ_1 is $\operatorname{Uniform}(0, \pi)$. Conditional on ξ_1 and more generally on Π with an $\ell \in \Pi$ passing through x, the intersection of $\widetilde{\Pi}$ with ℓ is a Poisson point process on ℓ of unit intensity. Moreover if the intersection points are marked with angles of intersection ξ_2 then the mark ξ_2 has mark density $\frac{1}{2}\sin\xi_2$ over $\xi_2 \in [0, \pi)$ (consider the length of the silhouette of a portion of ℓ viewed at angle ξ_2). Hence the conditional distribution of ξ_2 for $x = \iota(\ell, \widetilde{\ell})$ has density $\frac{1}{2}\sin\xi_2$ over $\xi_2 \in [0, \pi)$, and so we can compute

$$\begin{aligned}
\mathbb{P}\left[\ell \not\pitchfork \underline{s}, \widetilde{\ell} \not\pitchfork \underline{s} \,|\, x = \iota(\ell, \widetilde{\ell})\right] \quad &= \quad \frac{1}{\pi} \int_0^{\pi - \theta - \psi} \left(1 - \int_{\theta + \psi - \xi_1}^{\pi - \xi_1} \frac{\sin\xi_2}{2} \mathrm{d}\xi_2\right) \mathrm{d}\xi_1 \\
&= \quad \frac{\pi - \theta - \psi - \sin(\theta + \psi)}{\pi} \quad = \quad \frac{\phi - \sin\phi}{\pi} \quad (8)
\end{aligned}$$

where θ is the angle at v_j, and ψ is the angle at v_i, of the triangle formed by x, v_i, v_j; and ϕ is the exterior angle at x (see Figure 1).

Finally the intensity ν of $\mathcal{X}$ can be computed as $\frac{\pi}{2}$, for example by computing the mean number of hits of the unit disk by Π, then by computing the average length of the intersection of the disk with a line of Π conditional on that line hitting the disk. Thus

$$\begin{aligned}
J_m &= \mathbb{E}\left[\operatorname{len}(\partial\mathcal{C}(v_i,v_j))\right] - 2m \\
&= \nu \iint_{\mathbb{R}^2} \mathbb{P}\left[\ell \not\pitchfork \underline{\mathrm{s}}, \widetilde{\ell} \not\pitchfork \underline{\mathrm{s}} \,|x = \iota(\ell,\widetilde{\ell}) \in \mathcal{X}\right] \exp\left(-\tfrac{1}{2}(\eta - m)\right) \operatorname{Leb}(\mathrm{d}x) \\
&= \frac{1}{2} \iint_{\mathbb{R}^2} (\phi - \sin\phi) \exp\left(-\tfrac{1}{2}(\eta - m)\right) \operatorname{Leb}(\mathrm{d}x) \qquad (9)
\end{aligned}$$

as required. $\square$

We now state and prove the main result of this section: an $O(\log m)$ upper bound on the mean perimeter excess length J_m.

Theorem 7 (Asymptotic upper bound on mean perimeter length)**.** *The mean perimeter excess length J_m is subject to the following asymptotic upper bound:*

$$J_m \quad \leq \quad O(\log m) \qquad \textit{as } m \to \infty\,. \qquad (10)$$

Proof. Without loss of generality, place the points v_i and v_j at $(-\frac{m}{2},0)$ and $(\frac{m}{2},0)$. The double integral in (4) possesses mirror symmetry in each of the two axes, so we can write

$$\begin{aligned}
J_m &= 2 \iint_{[0,\infty)^2} (\phi - \sin\phi) \exp\left(-\tfrac{1}{2}(\eta - m)\right) \operatorname{Leb}(\mathrm{d}x) \\
&= 2 \int_0^{\pi/2} \int_0^{\frac{m}{2}\sec\theta} (\phi - \sin\phi) \exp\left(-\tfrac{1}{2}(\eta - m)\right) r \,\mathrm{d}r\,\mathrm{d}\theta + \\
&\qquad + 2 \int_{\pi/2}^{\pi} \int_0^{\infty} (\phi - \sin\phi) \exp\left(-\tfrac{1}{2}(\eta - m)\right) r \,\mathrm{d}r\,\mathrm{d}\theta \qquad (11)
\end{aligned}$$

(using polar coordinates (r,θ) about the second point v_j located at $(\frac{m}{2},0)$). The integrand in the second summand is dominated by $\pi \exp\left(-\frac{r}{2}\right) r$, which is integrable over $(r,\theta) \in (0,\infty) \times (\frac{\pi}{2},\pi)$. (In this region geometry shows that $\eta - m > r(1-\cos\theta) \geq r$.) Thus we can apply Lebesgue's dominated convergence theorem to deduce that the second summand is $O(1)$ as $m \to \infty$, hence may be neglected.

In fact we can also show that part of the first summand generates an $O(1)$ term: the dominated convergence theorem can be applied for any $\varepsilon \in (0,\pi/2]$ to show that

$$2 \int_0^{\pi/2} \int_\varepsilon^{\frac{m}{2}\sec\theta} (\phi - \sin\phi) \exp\left(-\tfrac{1}{2}(\eta - m)\right) r \,\mathrm{d}r\,\mathrm{d}\theta \quad = \quad O(1)\,,$$

since the integrand is dominated by $\pi \exp\left(-\frac{r}{2}(1-\cos\theta)\right) r$ over the region $(r,\theta) \in (0,\infty) \times (\varepsilon, \frac{\pi}{2})$ (in this region geometry shows that $\eta - m > r(1-\cos\theta) >$

$r(1-\cos\varepsilon))$. Thus for fixed $\varepsilon \in (0, \frac{\pi}{2})$ as $m \to \infty$ we have the asymptotic expression

$$J_m \quad = \quad 2\int_0^{\varepsilon}\int_0^{\frac{m}{2}\sec\theta} (\phi - \sin\phi)\exp\left(-\tfrac{1}{2}(\eta - m)\right) r\,\mathrm{d}r\,\mathrm{d}\theta + O(1)\,.$$

Now in the region $(r,\theta) \in (0,\infty)\times(0,\varepsilon)$ we know $\phi < 2\theta < 2\varepsilon$, and moreover $\phi - \sin\phi$ is an increasing function of ϕ (so long as $\varepsilon < \frac{\pi}{4}$). Therefore there is a constant C_ε such that

$$\phi - \sin\phi \quad \leq \quad 2\theta - \sin(2\theta) \quad \leq \quad \frac{C_\varepsilon}{8}\,\frac{(2\theta)^2}{6} \quad \leq \quad C_\varepsilon \frac{1-\cos\theta}{3}\sin\theta\,.$$

Hence

$$\begin{aligned}
&2\int_0^{\varepsilon}\int_0^{\frac{m}{2}\sec\theta} (\phi - \sin\phi)\exp\left(-\tfrac{1}{2}(\eta - m)\right) r\mathrm{d}r\mathrm{d}\theta \\
&\quad\leq\quad \tfrac{2}{3}C_\varepsilon \int_0^{\varepsilon}\int_0^{\frac{m}{2}\sec\theta} (1-\cos\theta)\sin\theta\exp\left(-\tfrac{r}{2}(1-\cos\theta)\right) r\mathrm{d}r\mathrm{d}\theta \\
&\quad=\quad \tfrac{8}{3}C_\varepsilon \int_0^{\varepsilon}\left(\int_0^{\frac{m}{4}(\sec\theta-1)} e^{-s}s\mathrm{d}s\right)\frac{\sin\theta\,\mathrm{d}\theta}{1-\cos\theta} \qquad \text{(using } s = \tfrac{r}{2}(1-\cos\theta))\\
&\quad\leq\quad \tfrac{8}{3}C_\varepsilon \int_0^{\frac{m}{4}(\sec\varepsilon-1)}\left(\int_0^{v} e^{-s}s\mathrm{d}s\right)\frac{1}{1+4v/m}\,\frac{\mathrm{d}v}{v} \qquad \text{(using } v = \tfrac{m}{4}(\sec\theta-1))\\
&\quad\leq\quad \tfrac{8}{3}C_\varepsilon \log\left(\tfrac{m}{4}(\sec\varepsilon - 1)\right) + O(1)\,.
\end{aligned}$$

□

Remark 8. More careful analysis yields useful $o(1)$-asymptotics: in fact as $m \to \infty$ it can be shown that

$$J_m \quad = \quad \tfrac{8}{3}\log m + \tfrac{8}{3}\left(\gamma + \tfrac{55}{24}\right) + o(1)\,, \tag{12}$$

where γ is the Euler-Mascheroni constant. These $o(1)$-asymptotics show very good agreement with simulation: see for example the simulation reported in the legend of Figure 2.

3 A low-cost network with short routes

In this section we prove Theorem 3: for a given configuration $\mathbf{x}^n \subset [0,\sqrt{n}]^2$ we construct networks $G(x^n)$ for which both $\text{len}(G(\mathbf{x}^n)) - \text{len}(\text{ST}(\mathbf{x}^n))$ and $\text{excess}(G(\mathbf{x}^n)$ are small. The network is constructed by augmenting the Steiner tree network $\text{ST}(\mathbf{x}^n)$ in a hierarchical manner. Working from the largest scale downwards, we construct

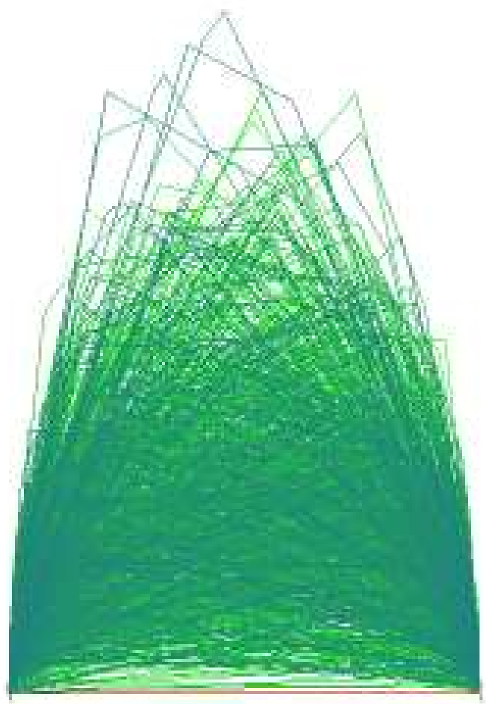

Figure 2: Simulation of semi-perimeters for 1000 independent cells for unit-rate Poisson line process, with city points located at distance 10^8 units apart. The figure is subject to vertical exaggeration: y-axis is scaled at 10^4 times x-axis. Empirical mean excess semi-perimeter is 27.63 with standard error ± 0.28, *versus* predicted mean excess semi-perimeter 27.462 (using $o(1)$-asymptotics).

1. a *stationary and isotropic Poisson line process* Π *of intensity* η, where η will be small: note that this can be constructed from a unit intensity process by scaling. A simple computation (Stoyan et al. 1995, §8.4) shows that the mean total length of the intersection of the resulting line pattern with $[0,\sqrt{n}]^2$ equals $\pi\eta n$.

2. A *medium-scale rectangular grid with cell side-length* $s_n \sim (\log n)^{1/3}$. Total length of this grid in $[0,\sqrt{n}]^2$ is bounded above by

$$2(1+\tfrac{\sqrt{n}}{s_n})\sqrt{n} \quad = \quad o(n)\,.$$

3. The *Steiner tree* $\mathrm{ST}(\boldsymbol{x}^n)$.

4. A small number (at most $n/2$) of *small hot-spot cells* based on a small-scale rectangular grid with cell side-length $t_n \sim \frac{1}{(\log n)^{1/6}}$. A cell in this

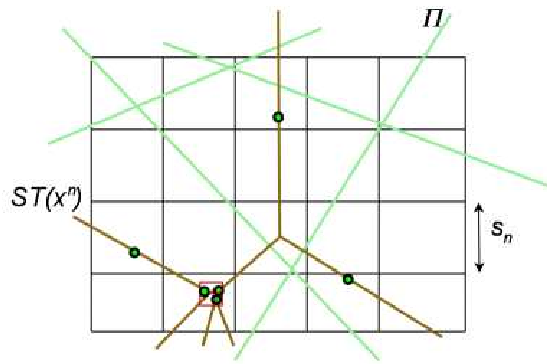

Figure 3: Illustration of construction of network to deliver an upper bound on mean excess route-length. City points are indicated by small circles. In this figure there is just one hot-spot cell.

grid is described as a *hot-spot cell* if it contains two or more city points. These hot-spot cells are used to by-pass regions where the Steiner tree might become complicated and expensive in terms of network traversal. We add further small segments connecting each hot-spot cell perimeter to city points within the hot-spot cell. Total length of these additions can be bounded by

$$4\frac{n}{2}t_n + n\frac{t_n}{2} \quad = \quad o(n)\,.$$

Thus the mean excess length of this augmented network is $o(n) + \pi\eta n$. The construction is illustrated in Figure 3. Note that we can choose s_n and t_n such that $n^{1/2}/s_n$ and s_n/t_n are integers, so that the small-scale lattice is a refinement of the medium-scale lattice, which itself refines the square $[0, \sqrt{n}]^2$.

3.1 Worst-case results for Steiner trees

We first record two elementary results on Steiner trees. The first result bounds the length of a Steiner tree in terms of the square-root of the number of points (for the planar case).

Lemma 9. *Consider a configuration $\boldsymbol{x}^k$ of k points in a square of side r: there is a constant C_1 not depending on k or r such that*

$$\operatorname{len}\left(\operatorname{ST}(\boldsymbol{x}^k)\right) \quad \leq \quad C_1\sqrt{k}r\,. \tag{13}$$

Proof. See Steele (1997, §2.2). □

The second result provides a local bound on length contributed by a larger Steiner tree in a small square containing a fixed number of points.

Lemma 10. *Consider the Steiner tree* $\mathrm{ST}(\boldsymbol{x}^n)$ *for an* arbitrary *configuration* $\boldsymbol{x}^n$ *in the plane. Let* G *be the restriction of the network* $\mathrm{ST}(\boldsymbol{x}^n)$ *to a fixed open square of side-length* t*. Suppose* k *points* $x_1, \ldots, x_k$ *of the configuration* $\boldsymbol{x}^n$ *lie within the square. Then*

$$\mathrm{len}(G) \quad \leq \quad t\left(4 + C_1\sqrt{k+1}\right) . \tag{14}$$

Proof. Let $y_1, \ldots, y_m$ be the locations at which $\mathrm{ST}(\mathbf{x}^n)$ crosses into the interior of the square. (Note: $m = 0$ is possible if $\{x_1, \ldots, x_k\} = \mathbf{x}^n$: in this case choose y_1 arbitrarily from the perimeter of the square.) Then

$$\begin{aligned}
\mathrm{len}(G) \quad &\leq \quad \mathrm{len}(\mathrm{ST}(\{x_1, \ldots, x_k, y_1, \ldots, y_m\})) && \text{by minimality of } \mathrm{ST}(\mathbf{x}^n), \\
&\leq \quad \mathrm{len}(\mathrm{ST}(\{x_1, \ldots, x_k, y_1\})) + 4t && \text{using square perimeter,} \\
&\leq \quad t\left(4 + C_1\sqrt{k+1}\right) && \text{using the previous lemma.}
\end{aligned}$$

□

3.2 Route-lengths in the medium-large network

The part of the construction involving the medium-scale grid and the Poisson line process is useful in variant problems, so we separate out the following estimate involving these ingredients.

Lemma 11. *Let* $n^{1/2}/s_n$ *be an integer. Consider the superposition of the rectangular grid with cell side-length* s_n *and the Poisson line process of intensity* η*, intersected with the square* $[0, n^{1/2}]^2$*. Let* v_i, v_j *be vertices of the grid. Then*

$$\mathbb{E}\left[\textit{route-length } v_i \textit{ to } v_j\right] \leq \mathrm{dist}(v_i, v_j) + C_2 \tfrac{1}{\eta} \log(\eta\sqrt{2n})$$

for an absolute constant C_2*.*

Proof. Let $\mathcal{C}(v_i, v_j)$ be the cell of Π containing v_i and v_j (having deleted lines from Π which separate v_i from v_j). Let $R(v_i, v_j)$ be the rectangle bounded by v_i and v_j; then by convexity the route-length from v_i to v_j is bounded above by

$$\frac{1}{2} \mathrm{len}\, \partial \left(R(v_i, v_j) \cap \mathcal{C}(v_i, v_j)\right) \quad \leq \quad \frac{1}{2} \mathrm{len}\, \partial \mathcal{C}(v_i, v_j) \,,$$

whose mean value can be computed by recognizing that the Poisson line process is a rescaled version of a homogeneous isotropic unit rate Poisson line process. Hence by scaling the asymptotic upper bound of Theorem 7 we have

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{2} \mathrm{len}\, \partial \left(R(v_i, v_j) \cap \mathcal{C}(v_i, v_j)\right)\right] - \mathrm{dist}(v_i, v_j) \quad &\leq \quad O\left(\frac{1}{\eta} \log\left(\eta\, \mathrm{dist}(v_i, v_j)\right)\right) \\
&= \quad O\left(\frac{1}{\eta} \log\left(\eta\sqrt{2n}\right)\right) .
\end{aligned}$$

□

3.3 Navigating the augmented network

We now explain how to move from points of $\mathbf{x}^n$ up to a vertex of the medium-scale grid.

Given $x_i \in \mathbf{x}^n$, if this is in one of the hot-spot cells then move to the perimeter of the hot-spot cell and thence to a suitable point of departure on the perimeter, with route-length at most $\frac{5}{2}t_n$. Now move along the Steiner tree within the relevant medium-scale grid box to the box perimeter; however by-pass all hot-spot cells. There are $(s_n/t_n)^2 = \left((\log n)^{1/3}(\log n)^{1/6}\right)^2 = \log n$ small squares each of which involves a route-length of either $2t_n$ (if a hot-spot box which will be by-passed) or $t_n(4+C_1\sqrt{2})$ (if not, by Lemma 10). Hence the total trip to the medium-scale grid box perimeter (including emergence from the initial hot-spot, if required) has length at most

$$\tfrac{5}{2}t_n + t_n(4+C_1\sqrt{2})\times s_n^2/t_n^2 \quad \sim \quad \tfrac{5}{2}t_n + (4+C_1\sqrt{2})\times(\log n)^{5/6} \quad = \quad o(\log n)\,.$$

Furthermore the route length from perimeter to vertex of medium-scale grid box is at most $\frac{1}{2}s_n \sim \frac{1}{2}(\log n)^{1/3} = o(\log n)$. So for each x_i there is a medium-scale grid vertex v_i for which route-length from x_i to v_i is $o(\log n)$. Combining with Lemma 11 and noting that the medium-scale grid geometry forces $\operatorname{dist}(v_i, v_j) \leq \operatorname{dist}(x_i, x_j) + 2\frac{s_n}{\sqrt{2}}$, we find

$$\mathbb{E}\left[\text{route-length from } x_i \text{ to } x_j\right] - \operatorname{dist}(x_i,x_j) \leq \sqrt{2}s_n + o(\log n) + C_2\tfrac{1}{\eta}\log\left(\eta\sqrt{2n}\right).$$

Averaging over the city points of $\mathbf{x}^n$, it follows that the dominant contribution comes from the cell semi-perimeters, and indeed

$$\mathbb{E}\left[\operatorname{excess}(G(\mathbf{x}^n))\right] \quad \leq \quad O\left(\tfrac{1}{\eta}\log\left(\eta\sqrt{2n}\right)\right)\,.$$

The two different results of Theorem 3 follow by choosing η to behave in two different ways:

(a) either $\eta \to 0$, $\eta w_n \to \infty$,

(b) or $\eta = \varepsilon > 0$.

4 A lower bound on average excess route-length

In this section we prove Theorem 5. The proof is divided into four parts. Firstly (Subsection 4.1) we show how to reduce the problem to an analogous case in which the excess is computed for two random city points drawn independently and uniformly from the whole disk D_n given in condition (b) of the theorem. Then (Subsection 4.2) we show that the network geodesic must run almost parallel to the Euclidean geodesic if the excess is small. On the other hand (Subsection 4.3) we can use the uniformity of the two random city points to control the extent to which network segments can run both close to and nearly parallel to the Euclideang geodesic. Finally (Subsection 4.4) we use the opposing estimates of Subsections 4.2 and 4.3 to derive a proof of the theorem using the method of contradiction.

4.1 Reduction to case of a pair of uniformly random city points

First we indicate how condition (a) of Theorem 5 implies condition (b). Under condition (a) we can use the coupling between X_n and Y_n to show that $\#\{\mathbf{x}^n \cap D_n\}/n \to \pi\rho$: therefore for large n the number of city points in D_n is approximately $\pi\rho n$. On the other hand the same coupling can be used to bound the total variation distance between the two conditional distributions $\mathcal{L}(Y_n|X_n \in D_n)$ and $\mathcal{L}(Y_n|Y_n \in D_n) = \text{Uniform}(D_n)$, and to show that this bound tends to zero. We can then use rejection sampling techniques to couple $\mathcal{L}(Y_n|X_n \in D_n)$ and $\text{Uniform}(D_n)$ so that the truncated Vasershtein distance tends to zero; as the distance is a metric we can combine this coupling with the (conditioned) coupling of $\mathcal{L}(X_n|X_n \in D_n)$ and $\mathcal{L}(Y_N|X_n \in D_n)$ to obtain a coupling which satisfies condition (b).

We now note that it is sufficient to consider the analogous result for a configuration $\mathbf{x}^n$ of n city points in the disk D_n. For then we can apply the result to the lesser configuration $\mathbf{y}^{k(n)}$ (for $k(n)$ as given in condition (b) of Theorem 5) and obtain

$$\text{excess}(G(\mathbf{y}^{k(n)})) \quad = \quad \Omega(\sqrt{\log k(n)}) = \Omega(\sqrt{\log \pi\rho n\varepsilon}) \quad = \quad \Omega(\sqrt{\log n})\,,$$

while

$$\begin{aligned}\text{excess}(G(\mathbf{y}^{k(n)})) \quad &= \quad \frac{n(n-1)}{k(n)(k(n)-1)}\,\text{excess}\,(G(\mathbf{x}^n)) \\ &\leq \quad \frac{1}{\pi\rho\varepsilon(\pi\rho\varepsilon - 1/n)}\,\text{excess}\,(G(\mathbf{x}^n))\ ,\end{aligned}$$

from which Theorem 5 follows.

We therefore consider $\mathbf{x}^n \subset D_n$ being L_n-equidistributed as the uniform distribution on D_n. So by definition there is a coupling (X_1, Y_1) (here we omit dependence on n) where X_1 has uniform distribution on $\mathbf{x}^n$, Y_1 has uniform distribution on D_n and

$$\Delta_n \quad = \quad \mathbb{E}\left[\min\left(1, \frac{|X_1 - Y_1|}{L_n}\right)\right] \quad \to \quad 0 \text{ as } n \to \infty. \tag{15}$$

Write (X_2, Y_2) for an independent copy of X_1, Y_1. In the definition of *excess* it makes no asymptotic difference if we allow $j = i$ in $\text{average}_{(i,j)}$, so we may take

$$\text{excess}(G(\mathbf{x}^n)) \quad = \quad \mathbb{E}\left[\ell(X_1, X_2) - \text{dist}(X_1, X_2)\right]. \tag{16}$$

Set

$$A_n \quad = \quad [|Y_1 - X_1| \leq L_n] \cap [|Y_2 - X_2| \leq L_n] \tag{17}$$

so that by Markov's inequality

$$\mathbb{P}[A_n] \quad \geq \quad 1 - 2\Delta_n. \tag{18}$$

Define $\ell(Y_1, Y_2)$ by supposing that Y_i is plumbed in to the network using a connection by a *temporary* line segment with endpoints Y_i and X_i. A direct computation shows that on A_n

$$\begin{aligned}
\ell(Y_1, Y_2) - \operatorname{dist}(Y_1, Y_2) \quad &\leq \\
(\ell(X_1, X_2) + |X_1 - Y_1| + |X_2 - Y_2|) - (\operatorname{dist}(X_1, X_2) &- |X_1 - Y_1| - |X_2 - Y_2|) \\
&\leq \quad \ell(X_1, X_2) - \operatorname{dist}(X_1, X_2) + 4L_n .
\end{aligned}$$

Consequently

$$\mathbb{E}\left[\ell(Y_1, Y_2) - |Y_1 - Y_2|; A_n\right] \quad \leq \quad \operatorname{excess}(G(\mathbf{x}^n)) + 4L_n \,. \tag{19}$$

By hypothesis $L_n = o(\sqrt{\log n})$, and so the proof of Theorem 5 reduces to showing that the left side (the excess for two random cities chosen uniformly in the disk) is $\Omega(\sqrt{\log n})$.

4.2 Near-parallelism for case of small excess

We now substantiate our previous remark that the network geodesic must run almost parallel to the Euclidean geodesic if the excess is small.

It is convenient to situate the disk D_n in the complex plane $\mathbb{C}$ in order to have a compact notation for rotations. For $t > 0$ we define Z_t and Φ by

$$\begin{aligned}
\exp(i\Phi) \quad &= \quad \frac{Y_2 - Y_1}{|Y_2 - Y_1|} \,, \\
Z_t \quad &= \quad Y_1 + t \times \exp(i\Phi) \,.
\end{aligned} \tag{20}$$

Let $\gamma : [0, \ell(Y_1, Y_2)] \to \mathbb{C}$ be the unit-speed network geodesic running from Y_1 to Y_2 (using the temporary plumbing to move from Y_1 to X_1 and then again from Y_2 to X_2). Then (bearing in mind that $|\gamma'(t)| = 1$)

$$\ell(Y_1, Y_2) \quad = \quad \int_0^{\ell(Y_1,Y_2)} |\gamma'(s)| \,\mathrm{d}s \quad \geq \quad \int_0^{\operatorname{dist}(Y_1,Y_2)} |\gamma'(\tau(t))| \,\tau'(t) \mathrm{d}t \,, \tag{21}$$

where $\tau(t)$ is the first time s at which $\langle \gamma(s) - Y_1, \exp(i\Phi) \rangle = t$. (Note that τ' can be infinite, but only at a countable number of points.) This and the following constructions are illustrated in Figure 4.

Defining $\theta(t)$ by $\sec\theta(t) = \tau'(t)$, and using $\sec\theta \geq 1 + \frac{1}{2}\theta^2$, we deduce

$$\ell(Y_1, Y_2) \quad \geq \quad \operatorname{dist}(Y_1, Y_2) + \frac{1}{2} \int_0^{\operatorname{dist}(Y_1,Y_2)} \theta(t)^2 \mathrm{d}t \,. \tag{22}$$

Furthermore we can use Pythagoras and the geodesic property of Euclidean line segments to show the following. Let $H(t)$ be the maximum $|r|$ for which, for some s,

$$\gamma(s) \quad = \quad Z_t + ir \exp(i\Phi) \,.$$

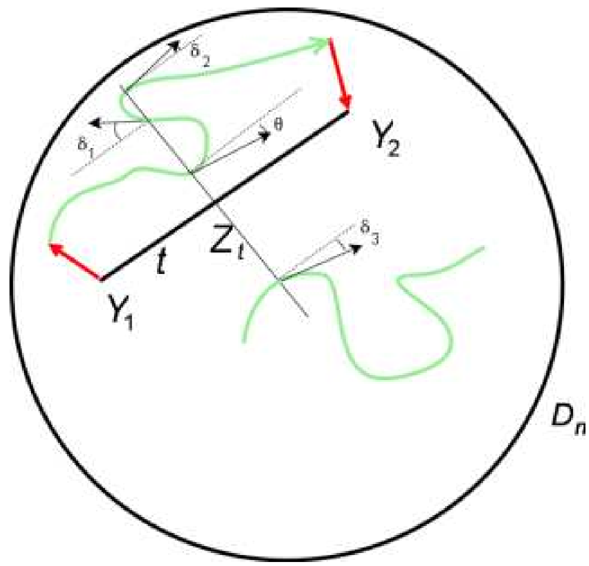

Figure 4: Illustration of construction of Y_1, Y_2, and Z_t. The angles $\theta(t)$ and δ_1, δ_2, ... are computed using the angles of incidence of network segments on the perpendicular running through Z_t; $\Upsilon_{t,\chi}$ is the minimum of absolute values of all such angles of points of intersection within $\sqrt{2t\chi + \chi^2}$ of Z_t.

If the excess for the network geodesic from Y_1 to Y_2 is bounded above by $\ell(Y_1, Y_2) - \operatorname{dist}(Y_1, Y_2) \leq \chi$ then $H(t) \leq \sqrt{2t\chi + \chi^2}$.

Let $\Upsilon_{t,\chi}$ be the smallest $|\delta|$ such that some network segment intersects the perpendicular $\{Z_t + ir \exp i\Phi : r \in \mathbb{R}\}$ at angle $\pi/2 + \delta$ and at distance at most $\sqrt{2t\chi + \chi^2}$ from Z_t (thus δ is the angle of incidence of this network segment on the perpendicular). If $\ell(Y_1, Y_2) - \operatorname{dist}(Y_1, Y_2) \leq \chi$ and $\operatorname{dist}(Y_1, Y_2) \geq \kappa\sqrt{\rho n}$, we can use (22) to deduce

$$\ell(Y_1, Y_2) - \operatorname{dist}(Y_1, Y_2) \quad \geq \quad \frac{1}{2}\int_0^{\kappa\sqrt{\rho n}} \Upsilon_{t,\chi}^2 \mathrm{d}t - \frac{1}{2}\left(\frac{\pi^2}{4}\right) \times \left(|X_1 - Y_1| + |X_2 - Y_2|\right) .$$

(The second summand allows for the temporary plumbing in of connections X_1Y_1 and X_2Y_2, for which the angle $\theta(t) \in (0, \frac{\pi}{2})$ is not controlled by permanent network segments). So introduce the event

$$B_{\kappa,\chi} \quad = \quad [\ell(Y_1, Y_2) - \operatorname{dist}(Y_1, Y_2) \leq \chi \,, \operatorname{dist}(Y_1, Y_2) \geq \kappa\sqrt{\rho n}] \qquad (23)$$

and recall the event $A_n = \cap_{i=1}^2 [|Y_i - X_i| \leq L_n]$. Taking expectations, we deduce

$$\begin{aligned}
&\mathbb{E}\left[\ell(Y_1, Y_2) - \operatorname{dist}(Y_1, Y_2)\ ;\ B_{\kappa,\chi} \cap A_n\right] \\
&\qquad\geq \frac{1}{2}\int_0^{\kappa\sqrt{\rho n}} \mathbb{E}\left[\Upsilon_{t,\chi}^2\ ;\ B_{\kappa,\chi} \cap A_n\right] \mathrm{d}t - \frac{\pi^2}{4} L_n\,.
\end{aligned}$$

Using integration by parts to replace the expectation by a probability,

$$\begin{aligned}
&\mathbb{E}\left[\ell(Y_1, Y_2) - \operatorname{dist}(Y_1, Y_2)\ ;\ B_{\kappa,\chi} \cap A_n\right] + \frac{\pi^2}{4} L_n \\
&\qquad\geq \int_0^{\kappa\sqrt{\rho n}} \int_0^\infty \mathbb{P}\left[[\Upsilon_{t,\chi} > u] \cap B_{\kappa,\chi} \cap A_n\right] u \,\mathrm{d}u\mathrm{d}t \\
&\qquad= \int_0^{\kappa\sqrt{\rho n}} \int_0^\infty \left(\mathbb{P}\left[B_{\kappa,\chi} \cap A_n\right] - \mathbb{P}\left[[\Upsilon_{t,\chi} \leq u] \cap B_{\kappa,\chi} \cap A_n\right]\right) u \,\mathrm{d}u\mathrm{d}t \\
&\qquad\geq \int_0^{\kappa\sqrt{\rho n}} \int_0^\infty \max\left(\mathbb{P}\left[B_{\kappa,\chi} \cap A_n\right] - \mathbb{P}\left[\Upsilon_{t,\chi} \leq u\right], 0\right) u \,\mathrm{d}u\mathrm{d}t\,. \qquad (24)
\end{aligned}$$

Note that from the definitions of $B_{\kappa,\chi}$ and A_n, using (18), (19) and Markov's inequality

$$1 - \mathbb{P}\left[B_{\kappa,\chi} \cap A_n\right] \leq 2\Delta_n + \mathbb{P}\left[\operatorname{dist}(Y_1, Y_2) \geq \kappa\sqrt{\rho n}\right] + \frac{\operatorname{excess}(G(\mathbf{x}^n)) + 4L_n}{\chi}. \qquad (25)$$

To make progress we need to find an upper bound for $\mathbb{P}\left[\Upsilon_{t,\chi} \leq u\right]$ and this is the subject of the next section.

4.3 Upper bounds using uniform random variables

Firstly we compute an upper bound on the joint density of the quantities Z_t and Φ from the previous section, illustrated in Figure 5.

Lemma 12. *Suppose Y_1, Y_2 are independent uniformly distributed random points in a disk D of radius $\sqrt{\rho n}$ and centre 0 in the complex plane $\mathbb{C}$. With Z_t and Φ defined as in (20), the joint density of Z_t and Φ is given over $\mathbb{C} \times [0, 2\pi)$ by*

$$\mathbb{I}\left[z - t e_\phi \in D\right] \frac{(t + s(z,\phi))^2}{2\pi^2\rho^2 n^2} \operatorname{Leb}(\mathrm{d}z)\,\mathrm{d}\phi\,, \qquad (26)$$

where $e_\phi = e^{i\phi}$ is the unit vector making angle ϕ with a reference x-axis, and $s(z,\phi)$ is the distance from z to the disk boundary ∂D in the direction ϕ (thus in particular $z + s(z,\phi)e_\phi$ is on the disk boundary).

Proof. Express the joint density for Y_1, Y_2 as a product of a uniform density over D for Y_1 and polar coordinates r, ϕ about Y_1 for Y_2:

$$\mathbb{I}\left[y_1 \in D\right] \frac{\operatorname{Leb}(\mathrm{d}y_1)}{\pi\rho n} \mathbb{I}\left[y_1 + r e^{i\phi} \in D\right] \frac{r\,\mathrm{d}r\,\mathrm{d}\phi}{\pi\rho n}\,.$$

Obtain the result by integrating out the r variable and transforming the y_1 variable to z by $z = y_1 + t e^{i\phi}$. □

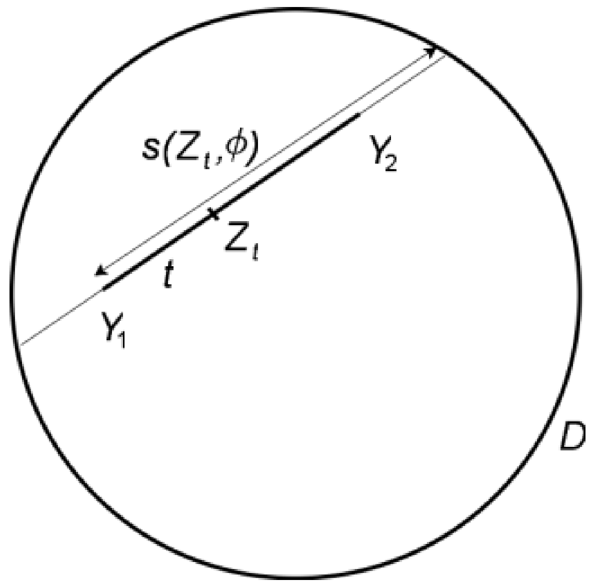

Figure 5: Illustration of construction in Lemma 12.

Corollary 13. *The density for* Z_t *and* $\Phi \pmod{\pi}$ *is*

$$
\begin{aligned}
f(z,\phi) \quad &= \\
&\left(\mathbb{I}\left[z - te_\phi \in D\right] \frac{(t + s(z,\phi))^2}{2} + \mathbb{I}\left[z + te_\phi \in D\right] \frac{(t + s(z, \pi + \phi))^2}{2}\right) \times \\
&\qquad \times \mathbb{I}\left[0 \le \phi < \pi\right] \frac{\operatorname{Leb}(\mathrm{d}z)\,\mathrm{d}\phi}{\pi^2 \rho^2 n^2} \,. \qquad (27)
\end{aligned}
$$

with an upper bound

$$
f(z,\phi) \quad \le \quad 4 \times \mathbb{I}\left[0 \le \phi < \pi\right] \frac{\operatorname{Leb}(\mathrm{d}z)\,\mathrm{d}\phi}{\pi^2 \rho n} \,. \qquad (28)
$$

Proof. Equation (27) follows immediately from adding the two expressions from Equation (26) for $\phi \pmod{\pi}$. The upper bound follows by noting

1. the maximum will occur when $z - te_\phi$ runs along a diameter;

2. furthermore when one of $z \pm te_\phi$ lies on the disk boundary;

3. and furthermore when $z = 0$ is located at the centre of the disk (so $t = s(z, \pm\phi) = \sqrt{\rho n}$).

□

Now consider the line segment $\mathcal{S}_{t,\chi}$ centred at Z_t, with end-points given by the pair $\pm i\sqrt{2t\chi+\chi^2}\exp(i\Phi)$; and consider the rose-of-directions empirical measure of angles made by intersections of network edges with this segment:

$$\mathcal{R}_{t,\chi}(A) \quad = \quad \#\{\text{ network intersections on } \mathcal{S}_{t,\chi} \text{ with angle of incidence lying in } A\} \tag{29}$$

(here angles are measured modulo π, and $A \subseteq [0,\pi)$). We may apply a Buffon-type argument to bound $\mathbb{E}[\mathcal{R}_{t,\chi}(A)]$ using Inequality (28). Consider the contribution to the expectation from a fixed line segment of the network of length ℓ: the result of disintegrating the integral expression for this according to the value of ϕ is an integral of $f(z,\phi)$ with respect to z over a region formed by intersecting the disk with a parallelogram of base side-length ℓ and height $2\sqrt{2t\chi+\chi^2}\sin\alpha$ (here the angle α depends implicitly on ϕ). Of course the integral vanishes if $\phi \notin A$. Thus Inequality (28) yields a bound

$$\mathbb{E}[\mathcal{R}_{t,\chi}(A)] \quad \leq \quad \frac{4\,\mathrm{len}(G(\mathbf{x}^n))}{\pi^2\rho n} \times \int_A 2\sqrt{2t\chi+\chi^2}\sin\alpha \mathrm{d}\alpha.$$

For constant χ, the event $[\Upsilon_{t,\chi} \leq u]$ is the event $[\mathcal{R}_{t,\chi}(\frac{\pi}{2}-u,\frac{\pi}{2}+u) \geq 1]$ and so

$$\begin{aligned}\mathbb{P}[\Upsilon_{t,\chi} \leq u] \quad &\leq \quad \mathbb{E}\left[\mathcal{R}_{t,\chi}(\tfrac{\pi}{2}-u,\tfrac{\pi}{2}+u)\right] \\ &\leq \quad \frac{16}{\pi^2\rho}\frac{\mathrm{len}(G(\mathbf{x}^n))}{n}\sqrt{2t\chi+\chi^2} \times u\,. \end{aligned} \tag{30}$$

4.4 Calculations

We have assembled the ingredients for the proof of Theorem 5, and now perform the calculations to get a quantitative lower bound.

Choose constants (as explained later) κ and $\chi = \chi_n$ such that for sufficiently large n (assumed in what follows)

$$\mathbb{P}[B_{\kappa,\chi} \cap A_n] \geq 2^{-1/3}. \tag{31}$$

Combine (19) and (24) (and the fact that $\pi^2/4 < 3$) to get

$$\mathrm{excess}(G(\mathbf{x}^n)) + 7L_n \geq \int_0^{\kappa\sqrt{\rho n}} \int_0^\infty \max\left(2^{-1/3} - \mathbb{P}[\Upsilon_{t,\chi} \leq u], 0\right) u \,\mathrm{d}u \,\mathrm{d}t.$$

By (30) and hypothesis of Theorem 5, there exists a constant B such that

$$\mathbb{P}[\Upsilon_{t,\chi} \leq u] \quad \leq \quad \sqrt{\frac{B}{12}}\sqrt{2t\chi+\chi^2} \times u.$$

Applying the formula $\int_0^\infty \max(0,\alpha-\beta u)u\,\mathrm{d}u = \frac{\alpha^3}{6\beta^2}$ we see

$$\mathrm{excess}(G(\mathbf{x}^n)) + 7L_n \geq \frac{1}{B}\int_0^{\kappa\sqrt{\rho n}} \frac{1}{2t\chi+\chi^2}\,\mathrm{d}t = \frac{\log(\kappa\sqrt{\rho n}+\frac{\chi}{2}) - \log\frac{\chi}{2}}{2\chi B}. \tag{32}$$

Recall this holds under the assumption that χ_n and κ satisfy (31). To finish we turn to an argument by contradiction: that is, suppose that (passing to a subsequence if necessary) $\mathrm{excess}(G(\mathbf{x}^n)) = o(\sqrt{\log n})$. By hypothesis $L_n = o(\sqrt{\log n})$. Inspecting (25) we see that we can choose some $\chi_n = o(\sqrt{\log n})$ and some small $\kappa > 0$ such that (31) holds. But then (32) takes the form

$$o(\sqrt{\log n}) \quad \geq \quad \frac{\Omega(\log n)}{o(\sqrt{\log n})},$$

which is impossible.

5 Closing remarks and supplements

5.1 Spatial network design

Within the realm of spatial network design, the closest work we know is that of Gastner and Newman 2006, who consider the similar notion of a *distribution network* for transporting material from one central vertex to all other vertices. They give a simulation study (their Figure 2) of a certain algorithm on random points, and comment

> Thus, it appears to be possible to grow networks that cost only a little more than the [minimum-length] network, but which have far less circuitous routes.

Our Theorem 3 provides a strong formalization of this idea.

5.2 Fractal structure of the Steiner tree on random points

Longstanding statistical physics interest in continuum limits of various discrete two-dimensional self-avoiding walks arising in probability models, *eg*

- uniform self-avoiding walks on the lattice,
- paths within uniform spanning trees in the lattice,
- paths within minimum spanning trees in the lattice,

has recently been complemented by spectacular successes of rigorous theory (Lawler, Schramm, and Werner 2004). It is conjectured that routes in Steiner trees on random points have similar fractal properties (Read 2005): route-length between points at distance n should grow as n^γ for some $\gamma > 1$. However, as our construction shows, such results have little relevance to spatial network design.

5.3 The counterintuitive observation

The counterintuitive observation following Definition 1 follows quickly from the work of Theorem 3. Suppose the configuration $\mathbf{x}^n$ is well-dispersed, in the weak sense that for some $\gamma \in (0,1)$ we find the number of city point pairs within $n^{\gamma/2}$ of each other is $o\left(\binom{n}{2} n^{\gamma-1}\right)$ (certainly this is the case for most patterns generated by uniform random sampling from $[0,\sqrt{n}]^2$). Consider a network $G(\mathbf{x}^n)$ produced by augmenting the Steiner tree according to the construction in the proof of Theorem 3. Using the properties of this construction, the following can be shown

$$\begin{aligned}
\mathbb{E}\left[\text{ratio}\left(G(\mathbf{x}^n)\right)\right] &= \mathbb{E}\left[\operatorname*{average}_{(i,j)} \frac{\ell(x_i,x_j)}{\text{dist}(x_i,x_j)} - 1\right] \\
&\leq \text{constant } \times o(n^{\gamma-1}) + (1 - o(n^{\gamma-1}))\left(\frac{O(\log\sqrt{2n})}{n^{\gamma/2}}\right) \\
&\leq O\left(\max\left(\frac{1}{n^{1-\gamma}}, \frac{\log n}{n^{\gamma/2}}\right)\right).
\end{aligned}$$

5.4 Derandomization

Theorem 3 is a purely deterministic assertion, though our proof used randomization (supplied by the Poisson line process). It seems intuitively plausible that one could give a purely deterministic proof, say by replacing the Poisson line process with a suitable sparse set of deterministically positioned lines having a dense set of orientations.

5.5 Quantifying equidistribution

The classical equidistribution property

> the empirical distribution of $\{n^{-1/2}x_i^n, 1 \leq i \leq n\}$ converges weakly to the uniform distribution on $[0,1]^2$

is equivalent (by a straightforward argument) to the property

> $\mathbf{x}^n$ is L_n-equidistributed as the uniform distribution on the square of area n, for *some* $L_n = o(n^{1/2})$.

Replacing one sequence of L_n by a slower-growing sequence makes equidistribution a stronger assumption, and so our assumption in Theorem 5(a) (equidistribution for some $L_n = o(\log^{1/2} n)$) is stronger than the classical equidistribution property. Indeed Theorem 5 fails under the classical equidistribution property, as the following example shows.

Example 14. Let $L_n = n^\gamma$ for some $\gamma \in \left(\frac{3}{8}, \frac{1}{2}\right)$. There exist networks $G(\mathbf{x}^n)$ which are L_n-equidistributed as the uniform distribution on the square of area n, for which $\text{len}(G(\mathbf{x}^n)) = o(n)$ whilst $\text{excess}(G(\mathbf{x}^n)) \to 0$.

For example: partition $[0,n^{1/2}]^2$ into subsquares of side $L_n/\log n$, construct the complete graph on all centers of such subsquares, allocate the n points evenly amongst subsquares and position them arbitrarily close to the centers.

Remark 15. Sample the configuration $\mathbf{x}^n$ independently and uniformly from $[0, \sqrt{n}]^2$. Let $L_n \to \infty$, perhaps arbitrarily slowly. Then the probability that the configuration $\mathbf{x}^n$ is L_n-equidistributed with the uniform distribution converges to 1. This follows by dviding $[0, \sqrt{n}]^2$ into cells of side-length asymptotic to $L_n/\sqrt{2}$, by conditioning on $\mathbf{x}^n$, and by "blurring" the points of $\mathbf{x}^n$ by replacing each point $x \in \mathbf{x}^n$ by an independent draw taken uniformly from the cell containing x. Then a uniform random draw $\widetilde{Y}_n$ of one of the blurred points can be coupled to lie within L_n of a uniform random draw X_n from the finite configuration $\mathbf{x}^n$. A simple argument using the Binomial distribution then shows that the total variation distance between $\widetilde{Y}_n$ and Uniform($[0, \sqrt{n}]^2$) tends to zero; it follows that X_n can be coupled to a Uniform($[0, \sqrt{n}]^2$) random variable Y_n so that

$$\mathbb{E}\left[\min\left(1, \frac{|X_n - Y_n|}{L_n}\right) | \, \mathbf{x}^n\right] \quad \to \quad 0\,,$$

where the convergence takes place in probability.

5.6 Poisson line process networks

Remark 8 indicates that more can be said about the mean semi-perimeter

$$\tfrac{1}{2}\,\mathbb{E}\left[\operatorname{len}(\partial\mathcal{C}(v_i, v_j))\right]\,,$$

and this will be returned to in later work. For example, consider the network formed entirely from a Poisson line pattern. If the pattern is conditioned to contain points v_i, v_j then the perimeter $\partial\mathcal{C}(v_i, v_j)$ will be close to providing a genuine network geodesic.

Note that questions about $\mathcal{C}(v_i, v_j)$ bear a family resemblance to the D.G.Kendall conjecture about the asymptotic shape of large cells in a Poisson line pattern. However $\mathcal{C}(v_i, v_j)$ is the result of a very explicit conditioning and hence explicit and rather complete answers can be obtained by direct methods, in contrast to the striking work on resolving the conjecture about large cells (Miles 1995; Kovalenko 1997; Kovalenko 1999; Hug, Reitzner, and Schneider 2004).

5.7 An open question

In the random points model we can pose a more precise question. Over choices of network G subject to the constraint

$$\mathbb{E}\left[\operatorname{len}(G(\mathbf{x}^n)) - \operatorname{len}(\mathrm{ST}(\mathbf{x}^n))\right] \quad = \quad o(n)\,,$$

or the constraint

$$\mathbb{E}\left[\operatorname{len}(G(\mathbf{x}^n)\right] \quad = \quad O(n)\,,$$

what is the minimum value of $\mathbb{E}\left[\operatorname{excess}(G(\mathbf{x}^n))\right]$? Our results pin down this minimum value, in the latter case to the range $[\Omega\left(\sqrt{\log n}\right), O(\log n)]$ and in the former case the range $[\Omega\left(\sqrt{\log n}\right), o(w_n \log n)]$. But it remains an open question what should be the exact order of magnitude.

David J. Aldous, Department of Statistics, 367 Evans Hall # 3860,
U.C. Berkeley CA 94720, USA
Email: aldous@stat.berkeley.edu
URL: www.stat.berkeley.edu/users/aldous

Wilfrid Kendall, Department of Statistics,
University of Warwick, Coventry CV4 7AL, UK
Email: w.s.kendall@warwick.ac.uk
URL: http://www.warwick.ac.uk/go/wsk

References

Gastner, M. T. and M. E. J. Newman (2006). Shape and efficiency in spatial distribution networks. *Journal of Statistical Mechanics: Theory and Experiment 2006*(01), P01015.

Hug, D., M. Reitzner, and R. Schneider (2004). The limit shape of the zero cell in a stationary Poisson hyperplane tessellation. *Ann. Probab. 32*(1B), 1140–1167.

Kovalenko, I. (1997). A proof of a conjecture of David Kendall on the shape of random polygons of large area. *Kibernet. Sistem. Anal.* (4), 3–10, 187. English translation: *Cybernet. Systems Anal.* 33 461-467.

Kovalenko, I. N. (1999). A simplified proof of a conjecture of D. G. Kendall concerning shapes of random polygons. *J. Appl. Math. Stochastic Anal. 12*(4), 301–310.

Lawler, G. F., O. Schramm, and W. Werner (2004). Conformal invariance of planar loop-erased random walks and uniform spanning trees. *Ann. Probab. 32*(1B), 939–995.

Miles, R. E. (1995). A heuristic proof of a long-standing conjecture of D. G. Kendall concerning the shapes of certain large random polygons. *Adv. Appl. Probab. 27*(2), 397–417.

Read, N. (2005). Minimum spanning trees and random resistor networks in d dimensions. *Phys. Rev. E (3) 72*(3), 036114, 17.

Steele, J. M. (1997). *Probability theory and combinatorial optimization*, Volume 69 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

Stoyan, D., W. S. Kendall, and J. Mecke (1995). *Stochastic geometry and its applications* (Second ed.). Chichester: John Wiley & Sons. (First edition in 1987 joint with Akademie Verlag, Berlin).

Yukich, J. E. (1998). *Probability theory of classical Euclidean optimization problems*, Volume 1675 of *Lecture Notes in Mathematics*. Berlin: Springer-Verlag.