# Scale-invariant random spatial networks

David J. Aldous

April 5, 2012

**Abstract**

Real-world road networks have an approximate scale-invariance property; can one devise mathematical models of random networks whose distributions are *exactly* invariant under Euclidean scaling? This requires working in the continuum plane. We introduce an axiomatization of a class of processes we call *scale-invariant random spatial networks*, whose primitives are routes between each pair of points in the plane. We prove that one concrete model, based on minimum-time routes in a binary hierarchy of roads with different speed limits, satisfies the axioms, and note informally that two other constructions (based on Poisson line processes and on dynamic proximity graphs) are expected also to satisfy the axioms. We initiate study of structure theory and summary statistics for general processes in this class.

*MSC 2010 subject classifications.* 60D05, 90B20

*Key words and phrases.* Poisson process, scale invariance, spatial network.

## 1 Introduction

Familiar web sites such as *Google maps* provide road maps on adjustable scale (zoom in or out) and a suggested route between any two specified addresses. Given $k$ addresses in a country, one could find the route for each of the $\binom{k}{2}$ pairs, and call the union of these routes the *subnetwork* (of the country's entire road network) spanning the $k$ points.

We abstract this idea by considering, for each pair of points $(z, z')$ in the plane, a random route $\mathcal{R}(z, z') = \mathcal{R}(z', z)$ between $z$ and $z'$. The collection of all routes (as $z$ and $z'$ vary) defines what one might call a continuum random spatial network, an idea we explain informally in this introduction (precise definitions will be given in section 2.2).

In particular, for each finite set $(z_1, \ldots, z_k)$ of points we get a random network $\mathbf{span}(z_1, \ldots, z_k)$, the *spanning subnetwork* linking the points, consisting of the union of the routes $\mathcal{R}(z_i, z_j)$. Mathematically natural structural properties we will impose on the distribution of such a process are
(i) translation and rotation invariance
(ii) scale-invariance.
For $0 < c < \infty$ the scaling map $\sigma_c : \mathbb{R}^2 \to \mathbb{R}^2$ takes $z$ to $cz$; we emphasize that (ii) means "naive Euclidean scaling", i.e. invariance under the action of $\sigma_c$, not any notion of "scaling exponent". For instance, scale-invariance implies that the route-length $D_r$ between points at (Euclidean) distance $r$ apart must scale as $D_r \stackrel{d}{=} rD_1$, where of course $1 \leq D_1 \leq \infty$. The setup so far does not exclude the possibility that routes are fractal, with infinite length, and such cases do in fact arise naturally in the tree-like models of section 8.7.2. But, envisaging road networks rather than some other physical structure, we restrict attention to the case $\mathbb{E}D_1 < \infty$. There is a rather trivial example, the *complete network* in which each $\mathcal{R}(z_1, z_2)$ is the straight line segment from $z_1$ to $z_2$, but the assumption "$\ell < \infty$" below will exclude this example.

Much of our study involves *sampled spanning subnetworks* $\mathcal{S}(\lambda)$, as follows. Write $\Xi(\lambda)$ for a Poisson point process in $\mathbb{R}^2$ of intensity $\lambda$, independent of the network. Then the points $\xi$ of $\Xi(\lambda)$, together with the routes $\mathcal{R}(\xi, \xi')$ for each pair of such points, form a random subnetwork we denote by $\mathcal{S}(\lambda)$. The distribution of $\mathcal{S}(\lambda)$ inherits the properties of translation- and rotation-invariance, and a form of scale-invariance described at (12). In particular we can define a constant $0 < \ell \leq \infty$ by

$$\ell = \text{mean length-per-unit-area of } \mathcal{S}(1)$$

(where "mean length-per-unit-area " is formalized by *edge-intensity* at (1)). In section 5.5 we note a crude lower bound $\ell \geq \frac{1}{4}$. We impose the property

$$\ell < \infty.$$

Regard $\ell$ as "normalized network length", for the purpose of comparing different networks.

Everything mentioned so far makes sense when only finite-dimensional distributions $\mathcal{R}(z_i, z_j)$ are specified. A first context in which we want to consider a process over the whole continuum concerns the following convenient abstraction of the notion of "major road". Write $\mathcal{R}_{(1)}(z_1, z_2)$ for the part of the route $\mathcal{R}(z_1, z_2)$ that is at distance $\geq 1$ from each of $z_1$ and $z_2$.

Conceptually, we want to study an edge-process $\mathcal{E}$ viewed as the union of $\mathcal{R}_{(1)}(z_1, z_2)$ over all pairs $(z_1, z_2)$. To formalize this directly would require some notion of "regularity" for a realization, for instance some notion of a.e. continuity of routes $\mathcal{R}(z_1, z_2)$ as $z_1$ and $z_2$ vary. But we can avoid this complication by first considering only $z_1, z_2$ in $\Xi(\lambda)$ and then letting $\lambda \to \infty$. After defining $\mathcal{E}$ in this way, we can define

$$p(1) := \text{mean length-per-unit-area of } \mathcal{E}$$

and impose the requirement

$$p(1) < \infty.$$

If a process of random routes $\mathcal{R}(z, z')$ satisfies the properties we have described (as stated precisely in section 2.2), then we will call it a *scale-invariant random spatial network* (SIRSN). As the choice of name suggests, it is the scale-invariance that makes such processes of mathematical interest; in section 1.5 we briefly discuss its plausibility for real-world networks.

We do not know any closely related previous work. We will discuss one related area of theory (discrete random spatial networks; section 1.2) and one area of application (fast algorithms for shortest routes:; section 1.4). Several more distantly related topics are mentioned in section 8.7.

## 1.1 Outline of paper

The purpose of this paper is to initiate study of SIRSNs, with three emphases. First, we give a careful formulation of an axiomatic setup for SIRSNs, with discussion of possible alternatives (section 2). Second, it is not obvious that SIRSNs exist at all! We give details of one construction in section 3. That construction envisages a square lattice of freeways, with "speed level $j$" freeways spaced $2^j$ apart, and the routes are the minimum time paths. Being based on the discrete lattice makes some estimates technically straightforward, but completing the details of proof requires surprisingly intricate arguments. This construction is somewhat artificial in not naturally having all the desired invariance properties, so these need to be forced by external randomization. We briefly mention two other constructions (in section 4.1 based on a weighted Poisson line process representing the different-level freeways, and in section 4.2 based on a dynamic construction of random points and roads added accoding to a deterministic rule) which intuitively seem more natural but for which we have been unable to complete all the details of a proof.

Third, in sections 5 - 6 we begin developing some general theory from the axiomatic setup. Of course *scale-invariance* is a rather weak assumption, loosely analogous to *stationarity* for a stochastic process, so one cannot expect sharp results holding throughout this general class of process. Our general results might be termed "structure theory" and concern existence and uniqueness issues for singly- and doubly-infinite geodesics, continuity of routes $\mathcal{R}(z_1, z_2)$ as a function of $(z_1, z_2)$, numbers of routes connecting disjoint subsets, and bounds on the parameters $\mathbb{E}D_1, \ell, p(1)$. One feature worth emphasis is that (very loosely analogous to *entropy rate* for a stationary process) the quantity $p(1)$ is a non-obvious statistic of a SIRSN, but turns out to be central in the foundational setup, in the structure theory, and in conceptual interpretation as a model for road networks. The latter is best illustrated by the "algorithms" story in sections 1.4 and 6.4.

Being a new topic there are numerous open problems, both conceptual and technical, stated in a final discussion section 8.

Before starting technical material, sections 1.2 - 1.5 give further verbal discussion of background to the topic.

## 1.2 Discrete spatial networks

Traditional models of (deterministic or random) spatial networks start with a *discrete* set of points and then assign linking edges via some rule, e.g. the random geometric graph [20] or proximity graphs [15], surveyed in [5]. One specific motivation for the present work was as a second attempt to resolve a paradox – more accurately, an unwelcome feature of a naive model – in the discrete setting, observed in [8]. In studying the trade-off between total network length and the effectiveness of a network in providing short routes between discrete cities, one's first thought might be to measure the latter by the average, over *all* pairs $(x, y)$, of the ratio

$$\text{(route-length for } x \text{ to } y)/(\text{Euclidean distance from } x \text{ to } y)$$

instead of averaging over pairs at Euclidean distance $\approx r$ to get our $\mathbb{E}D_r$. But it turns out that (in the $n \to \infty$ limit of a network on $n$ points) one can make this ratio tend to 1 for a network whose length is only $1+o(1)$ times the length of the Steiner tree, by simply superimposing on the Steiner tree a sparse Poisson line process. Such "theoretically optimal" networks are completely unrealistic, so there must be something wrong with the optimization criteria. What's wrong is that the networks are ineffective for small $r$. One way to get a non-trivial tradeoff in the $n \to \infty$ limit was described in [5]: using the statistic $\max_r r^{-1}\mathbb{E}D_r$ as the measure of effectiveness leads to more realistic-looking networks. In the discrete setting a network model cannot be precisely scale-invariant, but such considerations prompted investigation of continuum models which are assumed to be scale-invariant, so that $r^{-1}\mathbb{E}D_r$ is constant.

We emphasize that our networks involve roads at definite positions in the plane. There is substantial recent literature, discussed in [10], involving quite different notions of random planar networks, based on identifying topologically equivalent networks.

### 1.3 Visualizing spanning subnetworks

Both construction and analysis of general SIRSNs are based on studying subnetworks $\mathbf{span}(z_1, \ldots, z_k)$ on fixed or (most often) random points. It is helpful to visualize what subnetworks look like – see Figure 1.

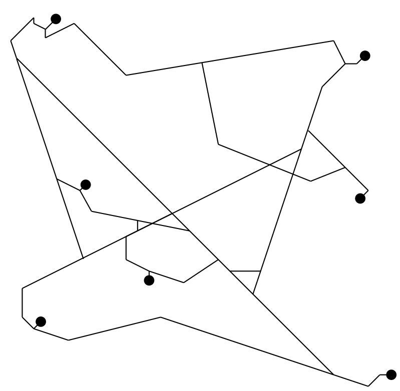

**Figure 1.** Schematic for the subnetwork of a SIRSN on 7 points $\bullet$

The qualitative appearance of Figure 1 is quite different from that of familiar spatial networks mentioned above, based on a discrete set of points, which could be viewed as abstractions of an inter-city road network, with cities as points. In contrast, we are abstracting the idea of the points $\bullet$ being individual street addresses a long way apart. The real-world route between two such street addresses will typically consist, in the middle, of roughly straight freeway segments but, nearing an endpoint, of a more jagged trajectory of shorter segments of lower-capacity roads; our setup and Proposition 13 imply the same behavior in our model.

## 1.4 Very fast shortest path algorithms

There is an interesting connection with the "shortest path algorithms" literature. Online mapping services and GPS devices require very quick computations of shortest routes. In this context, the U.S road network is represented as a graph on about 15 million street intersections (vertices) with edges (road segments) marked by distance (or typical driving time), and a given street address is recognized as being between two specific street intersections. Given a pair of (starting and destination) points, one wants to compute the shortest route. Neither of the two extremes – pre-compute and store the routes for all possible pairs; or use a classical Dijkstra-style algorithm for a given pair without any preprocessing – is practical. Bast et al (see [9] for an outline) find a set of about 10,000 intersections (which they call *transit nodes*) with the property that, unless the start and destination points are close, the shortest route goes via some transit node near the start and some transit node near the destination. Given such a set, one can pre-compute shortest routes and route-lengths between each pair of transit nodes; then answer a query by using the classical algorithm to calculate the route lengths from starting (and from destination) point to each nearby transit node, and finally minimizing over pairs of such transit nodes.

This idea is actually used commercially (and patented). Mathematical discussion was initiated by Abraham et al [1], who introduced the notion of *highway dimension*, defined as the smallest integer $h$ such that for every $r$ and every ball of radius $4r$, there exists a set of $h$ vertices such that every shortest route of length $> r$ within the ball passes through some vertex in the set. They discuss several algorithms whose performance can be analyzed in terms of highway dimension, and devise a particular model (a dynamic spanner construction on vertices given by an adversary) designed to have bounded highway dimension.

Now saying one can find $h$ independent of $r$ is a form of approximate scale-invariance, so the empirical fact that one can find transit nodes in the real-world road networks is a weak form of empirical scale-invariance. Within our model where precise scale-invariance is assumed, we can derive quantitative estimates relating to transit nodes – see section 6.4.

Incidently, the way we define edge-processes $\mathcal{E} = \mathcal{E}(\lambda, r)$ in terms of routes (mentioned in the Introduction and defined in section 2.2) is closely related to the notion of *reach* in the algorithmic literature [14].

### 1.5 Visualizing scale-invariance

Visualizing a photo of a road, scale-invariance seems implausible, because it implies existence of roads of arbitrarily large and arbitrarily small "sizes", however one interprets "size". But scale-invariance is not referring to the physical roads but to the process of "shortest routes", as in the discussion above. Figure 2 illustrates one aspect of scale-invariance. There is some number of crossing places (over the line) used by routes from one square to the other square. In our model, scale-invariance implies that the mean number of such crossings does not depend on the scale of the map. One could test this as a prediction about real-world road networks.

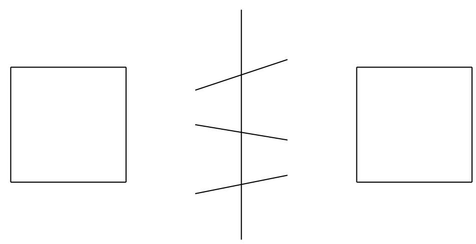

**Figure 2.** Schematic for long-distance routes.

As another empirical aspect of scale-invariance, [16] studied proportions of route-length, within distance-$r$ routes, spent on the $i$'th longest road segment in the route (identifying roads by their highway number designation) and observe that in the U.S. the averages of these ordered proportions are around $(0.40, 0.20, 0.13, 0.08, 0.05)$ as $r$ varies over a range of medium to large distances. Again, in our models (identifying roads as straight segments) scale-invariance implies there is some vector of expected proportions that is precisely independent of $r$.

# 2 Technical setup

In formulating an axiomatic setup there are several alternative choices one could make. In section 2.2 we state concisely the choices we made; section 2.3 discusses alternatives, reasons for choices, and immediate consequences or non-consequences of the setup.

### 2.1 Stochastic geometry background

We quote a fundamental identity from stochastic geometry (see [22] Chapter 8). Let $\mathcal{E}$ be an *edge process* – for our purposes, a union of line segments –

whose distribution is invariant under translation and rotation. Then $\mathcal{E}$ has an *edge-intensity*, a constant $\iota = \text{intensity}(\mathcal{E}) \in [0, \infty]$ such that

$$E(\text{length of } \mathcal{E} \cap A) = \iota \times \text{area}(A), \quad A \subset \mathbb{R}^2. \tag{1}$$

Moreover the positions and angles at which $\mathcal{E}$ intersects the $x$-axis (and hence any other line) are such that

$$\text{mean number intersections per unit length} = 2\pi^{-1} \times \text{intensity}(\mathcal{E}) \tag{2}$$

and the random angle $\Theta \in (0, \pi)$ of a typical intersection has density

$$f_\Theta(\theta) = \tfrac{1}{2} \sin \theta. \tag{3}$$

### 2.2 Definitions

Here we organize the setup via four aspects.

**Some notation.** $\mathbf{0}$ denotes the origin; $\text{disc}(z, r)$ and $\text{circle}(z, r)$ denote the closed disc and the circle centered at $z$.

**Aspect 1. Allowed routes and route-compatability.** Define a *jagged route* between two points $z, z'$ of $\mathbb{R}^2$ to consist of straight line segments between successive points $(z_i, -\infty < i < \infty)$ with $\lim_{i \to -\infty} z_i = z$ and $\lim_{i \to \infty} z_i = z'$, and such that the total length $\sum_{i=-\infty}^{\infty} |z_i - z_{i-1}|$ is finite. A *feasible route* is either a jagged route or the variant with a finite or semi-infinite set of successive line segments; we further require that the route be non-self-intersecting. Write $\mathsf{r}(z, z')$ for a feasible route, which from now on we will just call *route*. We envisage a route $\mathsf{r}(z, z')$ as a one-dimensional subset of $\mathbb{R}^2$, equipped with a label indicating it is the route from $z$ to $z'$. The route $\mathsf{r}(z', z)$ is always the reversal of $\mathsf{r}(z, z')$.

When we have a collection of routes, we require the following *pairwise compatability* property.

If two routes $\mathsf{r}(z_1, z_j)$, $\mathsf{r}(z'_1, z'_2)$ meet at two points then the routes coincide on the subroute between the two meeting points. (4)

**Aspect 2. Subnetworks on locally finite configurations.** Given a locally finite configuration of points $(z_i)$ in the plane, and routes $\mathsf{r}(z_i, z_j)$ satisfying the pairwise compatability property, write $\mathsf{s}$ for the union of all these routes. If $\mathsf{s}$ has the "finite length in bounded regions" property

$$\text{len}(\mathsf{s} \cap \text{disc}(\mathbf{0}, r)) < \infty \text{ for each } r < \infty \tag{5}$$

then call $\mathsf{s}$ a *feasible subnetwork*. Here "len' denotes "length". Formally $\mathsf{s}$ consists of the vertex set $(z_i)$, an edge set which is the union of the edge sets comprising each $\mathsf{r}(z_i, z_j)$, and marks on edges to indicate which routes they are in. Inclusion $\mathsf{s}(1) \subseteq \mathsf{s}(2)$ means that $\mathsf{s}(2)$ can be obtained from $\mathsf{s}(1)$ by adding extra vertices and associated routes.

As outlined in section 2.3 there is a natural $\sigma$-field that makes the set of all feasible subnetworks into a measurable space, so it makes sense below to talk about random feasible subnetworks.

**Aspect 3. Desired distributional properties of subnetworks.** The precise definition of the class of processes we shall study uses "finite-dimensional distributions" (FDDs), as follows. Given a finite set $z_1, \ldots, z_k$ let $\mu_{z_1,\ldots,z_k}$ be the distribution of a random feasible subnetwork $\mathbf{span}(z_1, \ldots, z_k)$ on $z_1, \ldots, z_k$. Suppose a family (indexed by all finite sets) of FDDs satisfies

$$\text{the natural consistency condition} \tag{6}$$

$$\text{invariance under translation and rotation} \tag{7}$$

$$\text{invariance under scaling.} \tag{8}$$

To be precise about (8), recall that the scaling map $\sigma_c : \mathbb{R}^2 \to \mathbb{R}^2$ takes $z$ to $cz$. Then the action of $\sigma_c$ on $\mathbf{span}(z_1, \ldots, z_k)$ gives a random subnetwork whose distribution equals the distribution of $\mathbf{span}(\sigma_c z_1, \ldots, \sigma_c z_k)$.

Appealing to the Kolmogorov extension theorem, we can associate with such a family a process of routes $\mathcal{R}(z_1, z_2)$, for each pair $z_1, z_2$ in $\mathbb{R}^2$, though for a process defined in that way we can only discuss properties determined by FDDs.

As mentioned earlier, much of our study involves *sampled spanning subnetworks*, as follows. For each $0 < \lambda < \infty$ let $\Xi(\lambda)$ be a Poisson point process of intensity $\lambda$ (we sometimes call this *point-intensity* to distinguish from edge-intensity at (1)). Make a process $(\Xi(\lambda),\ 0 < \lambda < \infty)$ by coupling in the natural way (take a space-time Poisson point process and let $\Xi(\lambda)$ be the positions of points arriving during time $[0, \lambda]$). Taking Poisson points independent of the process of routes, we can define $\mathcal{S}(\lambda)$ as the subnetwork of routes $\mathcal{R}(\xi, \xi')$ for pairs $\xi, \xi'$ in $\Xi(\lambda)$. We want the resulting processes $\mathcal{S}(\lambda)$ to have the following properties.

$$\text{for each } \lambda,\ \mathcal{S}(\lambda) \text{ is a random feasible subnetwork on vertex-set } \Xi(\lambda) \tag{9}$$

$$\text{for each } \lambda,\ \mathcal{S}(\lambda) \text{ has translation- and rotation-invariant distribution} \tag{10}$$

$$\mathcal{S}(\lambda_1) \subseteq \mathcal{S}(\lambda_2) \text{ for } \lambda_1 < \lambda_2 \tag{11}$$

$$\text{applying } \sigma_c \text{ to } \mathcal{S}(\lambda) \text{ gives a network distributed as } \mathcal{S}(c^{-2}\lambda). \tag{12}$$

For (12), recall that applying $\sigma_c$ to $\Xi(\lambda)$ gives a point process distributed as $\Xi(c^{-2}\lambda)$.

We omit full measure-theoretic details of the construction of $\mathcal{S}(\lambda)$, and just point out what extra conditions are needed to obtain properties (9 - 12). First, we need to impose the technical condition

$$\text{the map } (z_1, \ldots, z_k) \to \mu_{z_1,\ldots,z_k} \text{ is measurable} \tag{13}$$

to ensure that $\mathcal{S}(\lambda)$ is measurable. Second, part of the "feasible" assertion in (9) is the "finite length in bounded regions" property (5), and this property for $\mathcal{S}(\lambda)$ cannot be a consequence of assumptions on FDDs only, so we need

$$\operatorname{len}(\mathcal{S}(\lambda) \cap \operatorname{disc}(\mathbf{0}, r)) < \infty \ \text{ a.s. for each } r < \infty \tag{14}$$

and this will follow from the stronger assumption (16) below.

**Aspect 4. Final definition of a SIRSN.** To summarize the above: given a process of routes $\mathcal{R}(z_1, z_2)$ with FDDs satisfying (6 - 8, 13), we can define the process of sampled subnetworks $(\mathcal{S}(\lambda), 0 < \lambda < \infty)$ which, if (14) holds, will have properties (9 - 12). Finally, we define a SIRSN as a process (denoted by the routes $\mathcal{R}(z_1, z_2)$ or by the sampled subnetworks $(\mathcal{S}(\lambda), 0 < \lambda < \infty)$) satisfying these assumptions (6 - 8, 13) and also satisfying the extra conditions (15,20) below. These extra conditions merely repeat and formalize the requirements, stated in the introduction, that certain statistics be finite. As noted above, these assumptions imply that (9 - 12) hold.

Write $\mathbf{1} = (1, 0)$ and $D_1 := \operatorname{len} \mathcal{R}(\mathbf{0}, \mathbf{1})$. So $D_1$ represents route-length between points at distance 1 apart. Our definition of *feasible* route implies $1 \le D_1 < \infty$ a.s., and we impose the requirement

$$1 < \mathbb{E} D_1 < \infty. \tag{15}$$

Next, our definition of *feasible subnetwork* implies that $\mathcal{S}(1)$ must have a.s. finite length in a bounded region. We impose the stronger requirement of finite *expected* length. In terms of the edge-intensity (1), we require

$$\ell := \operatorname{intensity}(\mathcal{S}(1)) < \infty. \tag{16}$$

Finally, we define

$$\mathcal{E}(\lambda, r) := \bigcup_{\xi, \xi' \in \Xi(\lambda)} \mathcal{R}(\xi, \xi') \setminus (\operatorname{disc}(\xi, r) \cup \operatorname{disc}(\xi', r)) \tag{17}$$

and edge-intensities

$$p(\lambda, r) \ := \ \text{intensity}(\mathcal{E}(\lambda, r)) \tag{18}$$

$$p(r) \ := \ \lim_{\lambda \to \infty} p(\lambda, r) \tag{19}$$

and impose the requirement

$$p(1) < \infty \tag{20}$$

whose significance is discussed in the next section. Lemma 20 will show that (20) implies (16). If we do not require (20) but instead require (16), call the process a *weak* SIRSN.

### 2.3 Discussion of technical setup

**Aspect 1. Allowed routes and route-compatability.** Because we want routes to have a well-defined lengths, a minimum assumption would be that routes are rectifiable curves. We have assumed "feasible routes" in order to simplify notation. We believe that the theory would be essentially unchanged if instead one allowed rectifiable curves, as in the (quite different) theory mentioned in section 8.7.4.

It turns out (section 5.1) that realizations of our models always have jagged routes. A consequence is that (as in Figure 1) a route $\mathcal{R}(\xi, \xi')$ between two points of $\mathcal{S}(\lambda)$ does not pass through any third point $\xi''$ of $\mathcal{S}(\lambda)$. This prompts the precise definition of *geodesic* below.

The route-compatability property is a property that would hold if routes were defined as minimum-cost paths, for some reasonable notion of "cost". Note that our formal setup does not require routes to be minimum-cost in any explicit sense.

**Aspect 2. Subnetworks on locally finite configurations.** Here are some properties of a fixed feasible subnetwork.

**Lemma 1** *Let* $\mathsf{s}$ *be a feasible subnetwork on a locally finite, infinite configuration* $(z_i)$*.*
*(i) The set* $\{\mathsf{r}(z_i, z_j) \cap \text{disc}(z, r)\}_{i,j}$ *of sub-routes appearing as intersections of some route with a fixed disc* $\text{disc}(z, r)$ *contains only finitely many distinct (non-identical) sub-routes.*
*(ii) For each* $i$ *and each sequence* $(z_j)$ *with* $|z_j| \to \infty$ *there is a subsequence* $z'_k = z_{j(k)}$ *and a semi-infinite path* $\pi$ *from* $z_i$ *in* $\mathsf{s}$ *such that, for each* $r > 0$*,*

$$\mathsf{r}(z_i, z'_k) \cap \text{disc}(z_i, r) = \pi \cap \text{disc}(z_i, r) \textit{ for all large } k.$$

**Proof.** (ii) follows from (i) by a compactness argument. To outline (i), if false then (by the route-compatability property) the subroutes must meet the disc boundary at an infinite number of distinct points, and then (again by the route-compatability property) their extensions must meet the boundary of a slightly larger disc at an infinite number of distinct points, implying infinite length and contradicting the "finite length in bounded regions" property (5) of $\mathsf{s}$. ■

Note that Lemma 1 is implicitly about compactness in a topology on the space of paths within a given subnetwork $\mathsf{s}$. This is quite different from the topology of the space of all subnetworks, mentioned later.

**Terminology: paths, routes and geodesics.** A *path* in $\mathsf{s}$ has its usual network meaning. Typically there will be many paths between $z_i$ and $z_j$, but (as part of the structure of a *feasible subnetwork*) one is distinguished as the *route* $\mathsf{r}(z_i, z_j)$. So a route is a path; and a path may or may not be part of one or more routes. A *singly infinite geodesic* in $\mathsf{s}$ from $z_i$ is an infinite path, starting from $z_i$, such that any finite portion of the path is a subroute of the route $\mathsf{r}(z_i, z_k)$ for some $z_k$. So Lemma 1(ii) says that there always exists at least one singly infinite geodesic from $z_i$. A typical point $\epsilon$ along a route $\mathsf{r}(z_i, z_j)$ will sometimes be called a *path element* to distinguish it from the endpoints.

Now write $\mathfrak{S}$ for the set of all feasible subnetworks $\mathsf{s}$ on all locally finite configurations $\mathbf{x} = (x_j)$. It is natural to want to regard $\mathcal{S}(\lambda)$ as a random element of $\mathfrak{S}$, which requires specifying a $\sigma$-field on $\mathfrak{S}$, and as traditional we can do this by specifying a complete separable metric space structure on $\mathfrak{S}$ and using the Borel $\sigma$-field.

We outline a "natural" topology in an appendix. In this paper the topology plays no explicit role, but one can imagine developments where it does – one can imagine constructions using weak convergence, for instance, and compactness issues would be key to a proof of the existence part of Open Problem 30. However, it might be better to develop such theory within a framework where routes are allowed to be rectifiable curves.

**Aspect 3. Desired distributional properties of subnetworks.** The scale-invariance property (12)

$$\text{applying } \sigma_c \text{ to } \mathcal{S}(\lambda) \text{ gives a network distributed as } \mathcal{S}(c^{-2}\lambda)$$

is what gives SIRSNs a mathematically interesting structure, and almost all our general results in sections 5 and 6 rely on scale-invariance. To indicate how it is used, define $\ell(\lambda)$ analogously to (16):

$$\ell(\lambda) := \text{intensity}(\mathcal{S}(\lambda)) \tag{21}$$

so $\ell(1) = \ell$. Then there is a scaling relation

$$\ell(\lambda) = \lambda^{1/2}\ell, \quad 0 < \lambda < \infty. \tag{22}$$

To derive this relation, consider the scaling map $\sigma_{\lambda^{-1/2}}$ that takes $\mathcal{S}(1)$ to $\mathcal{S}(\lambda)$, and by considering the pre-image $A = [0, \lambda^{1/2}]^2$ of the unit square we see

$$\ell(\lambda) = \lambda^{-1/2} \times \text{area}(A) \times \ell$$

where the $\lambda^{-1/2}$ term is length rescaling.

Similar relations, provable in the same way, will be stated later (28,30, 37) without repeating the proof.

**Aspect 4. Final definition of a SIRSN.** Starting from FDDs, a conceptual and technical issue is how to continue to understand a SIRSN as a process over the whole continuum. As an analogy, for continuous-time stochastic processes one typically seeks some sample path regularity property such as *càdlàg*. So one might seek some notion of "regularity" for a realization, for instance a.e. continuity of routes $\mathcal{R}(z_1, z_2)$ as $z_1$ and $z_2$ vary. A version of such continuity is proved, under extra assumptions, in section 7.2. But as we next explain, in the present context the assumption $p(1) < \infty$ serves as an alternative regularity condition that enables us to study global properties of a SIRSN.

There are several possible real-world measures of "size" of a road segment, quantifying the minor road to major road spectrum – e.g. number of lanes; level in a highway classification system; traffic volume. What about within our model of a SIRSN? Recalling the definition (17) of $\mathcal{E}(\lambda, r)$, the limit

$$\mathcal{E}(\infty, r) := \cup_{\lambda<\infty}\mathcal{E}(\lambda, r)$$

has (because $\cup_{\lambda<\infty}\Xi(\lambda)$ is dense) the interpretation of "the set of path elements $\epsilon$ that are on some route $\mathcal{R}(z_1, z_2)$ with both $z_1$ and $z_2$ at distance $> r$ from $\epsilon$". As shown in section 6, assumption (20) implies that the edge-intensity $p(r)$ of $\mathcal{E}(\infty, r)$ is finite and scales as $p(r) = p(1)/r$. Moreover the random process $\mathcal{E}(\infty, r)$ is independent of the sampling process $(\Xi(\lambda), 0 < \lambda < \infty)$ and is an intrinsic part of the global structure of the SIRSN. So if we intuitively interpret $\mathcal{E}(\infty, r)$ as "the roads of size $\geq r$", then we have a mathematically convenient notion of "size of a road segment" emerging from our setup without explicit design. Intuitively, one could view the limit $\mathcal{E}(\infty, 0+) := \cup_{r>0} \mathcal{E}(\infty, r)$ as the continuum network of interest. But at a technical level it is not clear what are the properties of a realization of $\mathcal{E}(\infty, 0+)$, and we do not study it in this paper.

# 3 The binary hierarchy model

The construction of this model, our basic example of a SIRSN, occupies all of section 3, in several steps.

- A construction on the integer lattice (sections 3.1 - 3.4)
- Extension to the plane (sections 3.5 - 3.6)
- Further randomization to obtain invariance properties (section 3.7).

## 3.1 Routes on the lattice

For an integer $x \neq 0$, write height($x$) for the largest $j \in \mathbb{Z}^+$ such that $2^j$ divides $x$; in other words the unique $j$ such that $x = (2k+1)2^j$ for some $k \in \mathbb{Z}$. Set height(0) $= \infty$. For later use note that in one dimension, any integer interval $[m_1, m_2]$ contains a *unique* integer of maximal height, which we call peak$[m_1, m_2]$. For instance peak$[67, 99] = 96$ and peak$[34, 59] = 48$.

Until section 3.5 we will work on the integer lattice $\mathbb{Z}^2$, with vertices $z = (x, y)$ whose coordinates have heights $\geq 0$. While we are working on the lattice it is convenient to use $L^1$ distance $||z_2 - z_1||_1 := |x_2 - x_1| + |y_2 - y_1|$. Note also that until section 3.4 we work with *deterministic* constructions.

Write $L_x^{(X)}$ and $L_y^{(Y)}$ for the lines through $\{(x, y), y \in \mathbb{Z}\}$ and $\{(x, y), x \in \mathbb{Z}\}$. The height of a line $L_x^{(X)}$is the height of $x$.

Fix a parameter $1/2 < \gamma < 1$. Associate with lines at height $h$ a cost-per-unit length equal to $\gamma^h$. Now each path in the lattice has a cost, being the sum of the edge costs. Visualize a road network in which one can travel along a height-$h$ road at speed $1/\gamma^h$; so the cost equals time taken.

Define the route $\mathsf{r}(z_1, z_2)$ to be a minimum-cost path between $z_1$ and $z_2$. There is a uniqueness issue: for instance, for any minimum-cost path from $(i, i)$ to $(j, j)$ there is an equal cost path obtained by reflection $(x, y) \to (y, x)$. However, the estimates from here through section 3.3 hold when $\mathsf{r}(z_1, z_2)$ is any choice of minimum-cost path. We will deal with uniqueness in section 3.4.

A key point of the construction is that if we scale space by 2 then the scaled structure on the even lattice $(2\mathbb{Z})^2$ agrees with the original substructure on the even lattice, up to a constant multiplicative factor in edge-costs, and so the route between two even points will be the same whether we work in $\mathbb{Z}^2$ or $(2\mathbb{Z})^2$. So this "invariance under scaling by 2" property is built into the model at the start.

The fact that moving along the axes has zero cost may seem worrrying but actually causes no difficulty (we will later apply a random translation, and the original axes do not appear in the final process). Note that the cost associated with the line segment from $(2^h, 2^h)$ to $(2^h, 0)$ is $\gamma^h 2^h$ and the constraint $\gamma > 1/2$ is needed to make this cost increase with $h$. Intuitively, if $\gamma$ is near 1 then the route $\mathsf{r}(z_1, z_2)$ will stay inside or near the rectangle with opposite corners $z_1, z_2$, whereas if $\gamma$ is near $1/2$ then the route may go far away from the rectangle to exploit high-speed roads.

For this model we will show a property stronger than (15); the ratio of route-length to distance is uniformly bounded.

**Proposition 2** *There is a constant $K_\gamma < \infty$ such that*

$$\text{len } \mathsf{r}(z_1, z_2) \leq K_\gamma ||z_2 - z_1||_1, \quad \forall z_1, z_2 \in \mathbb{Z}^2.$$

Some intuition about possible paths in this model is provided by Figure 3 (the reader should imagine the ratios of longer/shorter edge lengths as larger than drawn). We might have a route as shown in the figure, where the two long edges are very fast freeways. But such a route is not possible if the fast freeways are too far from the start and destination points. The latter assertion will follow from Lemma 4.

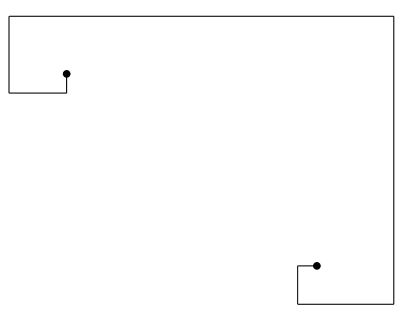

**Figure 3.** Routes like this are possible.

One might expect some explicit algorithmic description of routes $\mathsf{r}(z_1, z_2)$ that one can use to prove the results in sections 3.2 - 3.6, but we have been unable to do so. Instead our proofs rely on finding internal structural properties that routes must have.

### 3.2 Analysis of routes in the deterministic model

Consider the route from $z_1 = (x_1, y_1)$ to $z_2 = (x_2, y_2)$. The $x$-values taken on the route form some interval $I_x \supseteq [\min(x_1, x_2), \max(x_1, x_2)]$, and similarly the $y$-values form some interval $I_y$. Consider the point $z^* = (x^*, y^*) = (\mathrm{peak}(I_x), \mathrm{peak}(I_y))$ and call this point $\mathrm{peak}^{(2)}\mathsf{r}(z_1, z_2)$. The notation reminds us that $\mathrm{peak}^{(2)}\mathsf{r}(z_1, z_2)$ depends on the route $\mathsf{r}(z_1, z_2)$, which may not be unique.

**Lemma 3** *Consider the route* $\mathsf{r}(z_1, z_2)$ *from* $z_1$ *to* $z_2$.
*(i) The route passes through* $z^* = \mathrm{peak}^{(2)}\mathsf{r}(z_1, z_2)$.
*(ii) The route meets the line* $L^{(X)}_{x^*}$ *in either the single point* $z^*$ *or in one line segment containing* $z^*$ *(and similarly for* $L^{(Y)}_{y^*}$*).*
*(iii) Suppose the route passes through a point* $(x^*, y)$ *(for some* $y \neq y^*$*) and through a point* $(x, y^*)$ *(for some* $x \neq x^*$*). Then the route between those points is the two-segment route via* $z^*$.
*(iv) Suppose* $z^* = z_1$. *If* $z_2$ *is in a certain quadrant relative to* $z_1$*, for instance the quadrant* $[x_1, \infty) \times [y_1, \infty)$*, then the route from* $z_1$ *to* $z_2$ *stays in that quadrant.*

**Proof.** We first prove (iii). It is enough to prove that, amongst routes between $(x^*, y)$ and $(x, y^*)$, the two-segment route via $z^*$ is the *unique* minimum-cost route. In order to get from $(x^*, y)$ to the line $L^{(Y)}_{y^*}$ the route must use at least $|y - y^*|$ vertical unit edges; by definition of $x^* = \mathrm{peak}(I_x)$, if these edges are not precisely the line segment from $(x^*, y)$ to $z^*$ then the cost of these edges will be strictly larger; and similarly for horizontal edges. This establishes the uniqueness assertion above, and hence (iii).

For (i), if the hypothesis of (iii) fails then the route must go through $z^*$, whereas if it holds then the conclusion of (iii) implies the route goes through $z^*$. For (ii), if false then the the route passes through some two points $(x^*, y')$ and $(x^*, y'')$ but not the intervening points on that line. But (as in the argument for (iii)) the minimum cost path between those two points is the direct line between them.

Finally, (iv) follows from (ii), because if (iv) fails then the route meets one boundary of the quadrant in more than one segment. ∎

**Lemma 4** *The route* $\mathsf{r}(z_1, z_2)$ *from* $z_1$ *to* $z_2$ *stays within the square of side* $K'_\gamma \; ||z_2 - z_1||_1$ *centered at* $z_1$*, where* $K'_\gamma$ *depends only on* $\gamma$.

**Proof.** Choose the integer $h$ such that

$$2^{h-1} < ||z_2 - z_1||_1 \leq 2^h.$$

As illustrated in Figure 4, there is a square of the form $S = [(i-1)2^h, (i+1)2^h] \times [(j-1)2^h, (j+1)2^h]$ containing both $z_1$ and $z_2$ (note here $i$ and $j$ may be even or odd). We may suppose the route does not stay within $S$ (otherwise the result is trivial). For any point $z$ outside $S$, call the $L^\infty$ distance from $z$ to $S$, that is the number $d$ such that $z$ is on the boundary of the concentric square $S_d = [(i-1)2^h - d, (i+1)2^h + d] \times [(j-1)2^h - d, (j+1)2^h + d]$, the *displacement* of $z$. Now let $d$ be the maximum displacement along the route $\mathsf{r}(z_1, z_2)$, and choose a point $z'$ along the route with displacement $d$. So the route stays within $S_d$.

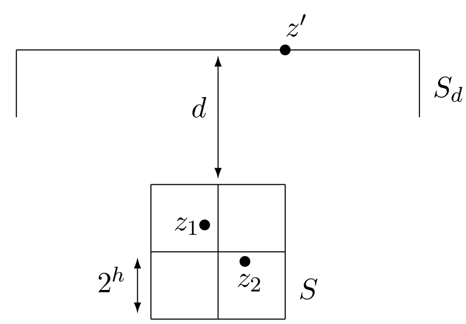


**Figure 4.** Construction for proof of Lemma 4.

We may assume, as in Figure 4, that $z'$ is on the top edge of $S_d$. The route needs to cover the vertical distance $d$ between the top edges of $S$ and $S_d$ *twice* (up and down) while staying within $S_d$, which has side-length $2^{h+1} + 2d$. Now within any integer interval of length $a$ the second-largest height $H$ satisfies $2^H \leq a$. So the cost ($C$, say) of the route outside $S$ is at least the cost associated with this second-largest height, which is given by

$$d\gamma^H \text{ where } 2^H \leq 2^{h+1} + 2d.$$

Setting $d = b2^h$, this inequality implies

$$\log_2 C \geq \log_2 b + h + (h + 1 + \log_2(1+b)) \log_2 \gamma.$$

But for this to be the minimum-cost path, the cost outside $S$ must be less than the cost of going round the boundary of $S$, which is at most $\gamma^h \times 2^{h+2}$. So

$$\log_2 C \leq h \log_2 \gamma + h + 2.$$

This inequalities combine to show

$$\log_2 b + (1 + \log_2(1+b)) \log_2 \gamma \leq 2$$

which, because $\gamma > 1/2$, implies that $b$ is bounded by some constant $b_\gamma$. ■

Lemma 4 makes Proposition 2 look very plausible, but to prove it we need to extend Lemma 3 to develop internal structural properties that routes must have.

Call a sequence of integers $i_1, i_2, \dots, i_m$ a *height-monotone sequence* from $i_1$ to $i_m$ if
(i) $\mathrm{height}(i_1) > \mathrm{height}(i_2) > \dots > \mathrm{height}(i_m) \geq 0$;
(ii) $|i_{j+1} - i_j| < 2^{\mathrm{height}(i_j)}, \quad 1 \leq j < m$.
Suppose, for integers $m_1, m_2, m^*$, we are given a height-monotone sequence $m^* = i_1, i_2, \dots, i_m = m_2$ and a height-monotone sequence $m^* = j_1, j_2, \dots, j_q = m_1$. Then we can form the concatenation
$m_1 = j_q, j_{q-1}, \dots, j_2, m^*, 1_2, \dots, i_m = m_2$. Call a sequence that arises this way an *admissable sequence* from $m_1$ to $m_2$. See Figure 5.

| (height) | | | | | | | |
|---|---|---|---|---|---|---|---|
| 5 | | | | | 96 | | |
| 4 | | | | 80 | | | |
| 3 | 72 | | | | | | |
| 2 | | | | | | | 100 |
| 1 | | 74 | | | | | |
| 0 | | | 75 | | | 99 | |

**Figure 5.** An admissable path from 75 to 99. This path has range $100 - 72 = 28$.

Regard a height-monotone or admissable sequence as a path of steps where a step from $i$ to $j$ has length $|j - i|$. It is clear from (ii) that the length of the path in (i) is at most twice the length of the first step. We deduce the following crude bound.

> (*) The total length of an admissable path is at most 4 times the *range* of the path, where the range is the difference between the maximum and minimum integer points visited by the path.

Proposition 2 follows immediately from Lemma 4, the bound (*) above and the following lemma.

**Lemma 5** *The route from $z_1 = (x_1, y_1)$ to $z_2 = (x_2, y_2)$ consists of alternating horizontal and vertical segments, in which the successive distinct $x$-values of the segment ends (the turning points) form an admissable sequence from $x_1$ to $x_2$, and the successive distinct $y$-values form an admissable sequence from $y_1$ to $y_2$.*

**Proof.** In view of Lemma 3 we can reduce to the case where $z_1 = \mathrm{peak}^{(2)}\mathsf{r}(z_1, z_2)$, and we need to show that the successive distinct $x$-values form a height-monotone sequence, as do the $y$-values. Without loss of generality suppose that $x_1 \le x_2$, that $y_1 \le y_2$ and that the first segment is horizontal. So the route is of the form

$$(x_1, y_1) = (x_{(1)}, y_{(1)}) \to (x_{(2)}, y_{(1)}) \to (x_{(2)}, y_{(2)}) \to (x_{(3)}, y_{(2)}) \to \dots$$

It suffices to show that for each edge of the route, say the edge $(x_{(i)}, y_{(i)}) \to (x_{(i+1)}, y_{(i)})$, and for each point (say $(x^*, y_{(i)})$) on the edge other than the starting point, we have $\mathrm{height}(x^*) < \mathrm{height}(x_{(i)})$. This is true for the first two edges of the route by definition of $z_1$ as $\mathrm{peak}^{(2)}\mathsf{r}(z_1, z_2)$. Suppose it fails first at some point $(x^*, y_{(i)})$. Then the route has proceeded $(x_{(i)}, y_{(i-1)}) \to (x_{(i)}, y_{(i)}) \to (x^*, y_{(i)})$ instead of the alternate path via $(x^*, y_{(i-1)})$. Now inductively $\mathrm{height}(y_{(i)}) < \mathrm{height}(y_{(i-1)})$, so the cost of the horizontal edge is less in the alternate path; so for the route to have smaller cost it must happen that the cost of its vertical edge is smaller than in the alternate path, that is $\mathrm{height}(x_{(i)}) > \mathrm{height}(x^*)$, contradicting the supposed failure. ∎

### 3.3 Further technical estimates

The next lemma will be key to bounding network length, more specifically to showing $\ell < \infty$ later.

**Lemma 6** *There exists an integer $b \ge 1$, depending only on $\gamma$, such that for all $h \ge 0$ and all rectangles of the form $[i2^{h+b}, (i+1)2^{h+b}] \times [j2^h, (j+1)2^h]$, the route $\mathsf{r}(z_1, z_2)$ between two points $z_1, z_2 \in \mathbb{Z}^2$ outside (or on the boundary of) the rectangle does not use any horizontal edge strictly inside the rectangle.*

Note there may be routes using a vertical line straight through the rectangle.
**Proof.** Suppose false; then there are two points $z_1, z_2$ on the boundary of the rectangle such that the route between them lies strictly within the rectangle and contains a horizontal edge. Because the speed on an interior edge is less than the speed on a parallel boundary edge, this cannot happen when $z_1$ and $z_2$ are in the same or adjacent boundaries of the rectangle,

because the path around the boundary is faster. Suppose they are on the top and the bottom boundaries. Then the height of the horizontal edge is less than the heights of the starting and ending $y$-values, contradicting Lemma 5. The only remaining case is when $z_1$ and $z_2$ are on the left and right boundaries. Using Lemma 5 again, the route cannot use a vertical edge inside the rectangle, so the only possibility is a single horizontal segment passing through the rectangle. Such a path has cost at least $2^{h+b}\ \gamma^{h-1}$, because the height of the line is at most $h-1$, whereas the path around the boundary has cost at most $2^{h+b}\ \gamma^h + 2^h\ \gamma^{b+h}$. So the potential route is impossible when $2^b + \gamma^b < 2^b\gamma^{-1}$ which holds for sufficiently large $b$. ■

**Corollary 7** *If a route* $\mathsf{r}(z_1, z_2)$ *uses a height-*$h$ *segment through* $z_0$*, then* $\min(||z_1 - z_0||_1, ||z_2 - z_0||_1) \leq 2^h(2^b + 1)$ *for* $b$ *as in Lemma 6.*

**Proof.** Consider a unit-length horizontal (without loss of generality) edge of height $h$ at $z_0$. It is in the interior of some rectangle of the form $[i2^{h+1+b}, (i+1)2^{h+1+b}] \times [j2^{h+1}, (j+1)2^{h+1}]$. By Lemma 6 applied with $h+1$, either $z_1$ or $z_2$ must be within that rectangle. ■

Perhaps surprisingly, we do not make much explicit use of the deterministic function $\mathbf{cost}(z_1, z_2)$ giving the cost of the minimum-cost route in the integer lattice, but will need the following bound.

**Lemma 8** *There exists a constant* $K''_\gamma$ *such that*

$$\mathbf{cost}(z_1, z_2) \leq K''_\gamma ||z_2 - z_1||^\beta$$

*where* $\beta := \log(2\gamma)/\log 2$.

**Proof.** As in Figure 4 in the proof of Lemma 4, there is a square of the form $S = [(i-1)2^h, (i+1)2^h] \times [(j-1)2^h, (j+1)2^h]$ containing both $z_1$ and $z_2$, where $h$ is the integer such that $2^{h-1} < ||z_2 - z_1||_1 \leq 2^h$. As observed there, the cost of going all around the boundary of $S$ is $O(\gamma^h 2^h)$. By considering a path from $z_1$ using the "greedy" rule of always switching to an orthogonal line of greater height, it is easy to check that the cost of this greedy path from $z_1$ to the boundary of $S$ is also $O(\gamma^h 2^h)$. Hence $\mathbf{cost}(z_1, z_2) = O(\gamma^h 2^h)$ and the result follows.

### 3.4 Finessing uniqueness by secondary randomization

As previously observed, minimum-cost paths are not always unique. We conjecture that, at least when $\gamma$ is not algebraic, there is some simple classification of when and how non-uniqueness occurs. But instead of addressing that issue we can finesse it by introducing randomness (which we need later, anyway) at this stage. One possible way to do so would be to use the uniform distribution on minimum-cost paths. Instead we use what we will call *secondary randomization* to choose between non-unique minimum-cost paths. Place i.i.d. Normal$(0,1)$ random variables ("weights") $\zeta_e$ on the edges $e$ of $\mathbb{Z}^2$. Any path has a weight $\sum_{e \text{ in path}} \zeta_e$. Define the route $\mathcal{R}_0(z_1, z_2)$ to be the minimum-weight path in the set of minimum-cost paths from $z_1$ to $z_2$.

### 3.5 Extension to the binary rational lattice

The notion of *height* extends to binary rationals: if $x \in \mathbb{R}$ is a binary rational and $x \neq 0$, write height$(x)$ for the largest $j \in \mathbb{Z}$ such that $2^j$ divides $x$; in other words the unique $j$ such that $x = (2k+1)2^j$ for some $k \in \mathbb{Z}$.

For $-\infty < H < \infty$ let $\mathbb{Z}^2_H$ be the lattice on vertex-set $\{2^H z : \ z \in \mathbb{Z}^2\}$, in other words on the set of points in $\mathbb{R}^2$ whose coordinates have height $\geq H$. So far we have been working on the integer lattice $\mathbb{Z}^2$, but now the results we have proved extend by (binary) scaling to analogous results on the lattices $\mathbb{Z}^2_H$. We will use such scaled results as needed.

Note in particular the following consistency condition as $H$ varies. Take $H_1 < H_2$. Consider the route, in $\mathbb{Z}^2_{H_1}$, between two vertices of $\mathbb{Z}^2_{H_2}$. By Lemma 5 and the definition of admissable, any minimum-cost path stays within the lattice $\mathbb{Z}^2_{H_2}$. So the set of minimum-cost paths is the same whether we work in $\mathbb{Z}^2_{H_1}$ or in $\mathbb{Z}^2_{H_2}$. Note also that each edge $e$ in $\mathbb{Z}^2_H$ corresponds to two edges $e_1, e_2$ of $\mathbb{Z}^2_{H-1}$. So we can couple the edge-weights by making $\zeta_e = \zeta_{e_1} + \zeta_{e_2}$ (only this infinite divisibility property of the Normal is relevant to the construction) and this gives a "consistency of secondary weights" property, which implies that the random route $\mathcal{R}_0(z_1, z_2)$ is also the same whether we work in $\mathbb{Z}^2_{H_1}$ or in $\mathbb{Z}^2_{H_2}$.

So we have now defined random routes $\mathcal{R}_0(z_1, z_2)$ for all unordered pairs of vertices in $\mathbb{Z}^2_{-\infty} := \cup_{H > -\infty} \mathbb{Z}^2_H$. From the "minimality" in the construction it is clear that the routes satisfy the route-compatability properties (iii,iv) from section 2.2.

### 3.6 Extension to the plane

We want to define routes $\mathcal{R}_0(z_1, z_2)$ between general points $z_1, z_2$ of $\mathbb{R}^2$ as $H \to -\infty$ limits of the routes $\mathcal{R}_0(z_1^H, z_2^H)$ between vertices such that

$$z_i^H \in \mathbb{Z}^2_H, \quad z_i^H \to z_i \quad (i = 1, 2) \tag{23}$$

Proposition 9 formalizes this idea. The proof in this section is the most intricate part of the construction, which can thereafter be completed (section 3.7) by "soft" arguments.

As a first issue, what does it mean to say that, under (23),

$$\text{routes } \mathsf{r}(z_1^H, z_2^H) \text{ converge to a route } \mathsf{r}(z_1, z_2)? \tag{24}$$

We define this to mean:

> for each $H_0 > -\infty$, the subroute $\mathsf{r}_{H_0}(z_1^H, z_2^H)$ consisting of path segments of $\mathsf{r}(z_1^H, z_2^H)$ within lines of height $\geq H_0$ is, for sufficiently large negative $H$, a path not depending on $H$ – call this path $\mathsf{r}_{H_0}(z_1, z_2)$.

When this property holds, Lemma 5 implies that $\mathsf{r}_{H_0}(z_1, z_2)$ is a connected path, consistent as $H_0$ decreases, and using Proposition 2 and scaling we see that the closure of $\cup_{H_0 > -\infty} \mathsf{r}_{H_0}(z_1, z_2)$ defines a route $\mathsf{r}(z_1, z_2)$ satisfying the "jagged" condition of section 2.2.

**Proposition 9** *There exists a subset $A \subset \mathbb{R}^2$ of zero area such that, if $z_1$ and $z_2$ are outside $A$, there exists a random jagged route $\mathcal{R}_0(z_1, z_2)$ such that, whenever (23) holds, then (24) holds.*

The proof relies on the fact that, for particular configurations illustrated in Figure 6, routes from a certain neighborhood to distant destinations must pass through a particular point. In fact all that matters is the *existence* of such a configuration, not the particular one we now exhibit. Consider a square $G = [2^h, 2^h + 2]^2$ and points $b = (2^h + 1, 2^h)$ and $d = (2^h + 1, 2^h + 2^{-h})$, illustrated in Figure 6. What is relevant is the heights of the lines involved, indicated in the figure.

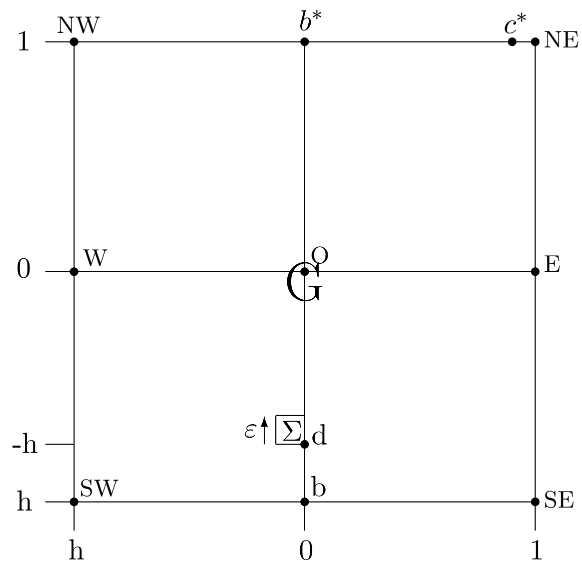


**Figure 6.** The big square $G$ and the small square $\Sigma$. Marginal labels attached to lines are line-heights, not coordinates.

**Lemma 10** *There exist large $h$ and small $\varepsilon$ (depending on $\gamma$) such that, in the configuration shown in Figure 6, every route from inside the small square $\Sigma := d + [-\varepsilon, 0] \times [0, \varepsilon]$ to the boundary of $G$ passes via $b$.*

**Proof.** For each point $c$ on the boundary of $G$ there is a counter-clockwise path $\pi_1(b, c)$ and a clockwise path $\pi_2(b, c)$ along the boundary from $b$ to $c$. These paths have equal cost for the point $c^* = (2^h + 2 - \gamma^{h-1}, 2^h + 2)$, which is near the NE corner point of $G$. We will need the following lemma.

**Lemma 11** *There exists $h$ such that the following hold.*
*(a) The paths $\pi_1(b, c^*)$ and $\pi_2(b, c^*)$ attain the minimum cost over all paths from $b$ to $c^*$, and are the only paths to do so.*
*(b) The only minimum-cost paths from $d$ to $c^*$ are the two paths consisting of the segment $[d, b]$ and the paths $\pi_1(b, c^*)$ or $\pi_2(b, c^*)$.*
*(c) There exists $\eta > 0$ such that any path from $d$ to $c^*$ that avoids the segment $[d, b]$ has cost at least $\eta$ greater that the minimum-cost paths.*

Note a technical point. We are working on $\mathbb{Z}^2_{-\infty} := \cup_{H > -\infty} \mathbb{Z}^2_H$ and $c^*$ may not be in $\mathbb{Z}^2_{-\infty}$. To be precise we should replace $c^*$ in the arguments below by a sequence $c^*_H \to c^*$, but that requires awkward notation we prefer to avoid.

Granted Lemma 11 we deduce Lemma 10 as follows. Consider a point $c$ on the counter-clockwise path from $b$ to $c^*$ (the clockwise case is similar). Then the following must hold, because any counter-example path to $c$ could be extended along the boundary from $c$ to $c^*$ and would give a counter-example to Lemma 11.

(a) The path $\pi_1(b,c)$ is the unique minimum-cost path from $b$ to $c$.

(b) The path consisting of the segment $[d,b]$ and the path $\pi_1(b,c)$ is the unique minimum-cost path from $d$ to $c$.

(c) Any path from $d$ to $c$ that avoids the segment $[d,b]$ has cost at least $\eta$ greater that the minimum-cost path.

Lemma 8 extends by scaling to $\cup_{H>-\infty}\mathbb{Z}^2_H$, and so the function $\mathbf{cost}(\cdot,\cdot)$ extends to a continuous function on $\mathbb{R}^2$. So we can choose $H$ so that the square $\Sigma = d + [-2^{-H}, 0] \times [0, 2^{-H}]$ satisfies $\sup_{s\in\Sigma} \mathbf{cost}(s,d) \leq \eta/3$.

It is easy to check that a minimum-cost path from $s \in \Sigma$ to $d$ does not meet $[d,b]$ except at $d$.

Consider $s \in \Sigma$ and a point $c$ as above. So $\mathbf{cost}(s,c) \leq \mathbf{cost}(d,c) + \eta/3$. Suppose a minimum-cost path from $s$ to $c$ does not meet the segment $[d,b]$. Then the path from $d$ to $c$ via $s$ would have cost $\leq \mathbf{cost}(d,c) + 2\eta/3$ and would not meet $[d,b]$, contradicting (c). So a minimum-cost path from $s$ to $c$ must meet the segment $[d,b]$. Then, by uniqueness in (b) (for the path from $d$ to $c$), it must continue via $b$, establishing Lemma 10. ■

**Proof of Lemma 11.** (I thank Justin Salez for completing the details of this proof.) In outline, we use the "structure of paths" results in Lemmas 3 and 5 to reduce to comparing costs of a finite number of possible routes. We will make use of the following preliminary observations, which are straightforward to check :

(i) the only minimum-cost path from $SE$ to $NW$ is $SE \to SW \to NW$ ;

(ii) the only minimum-cost paths from $SW$ to $NE$ are $SW \to NW \to NE$ and $SW \to SE \to NE$ ;

(iii) $O \to b^* \to NE$ is a minimum-cost path from $O$ to $NE$.

Consider assertion (a). The cost associated with paths $\pi_1(b,c^*)$ and $\pi_2(b,c^*)$ equals $2\gamma^h + 2\gamma$, which (by choosing $h$ large) is less than 2. Now consider some minimum-cost path $\pi$ from $b$ to $c^*$. Since both end-points have their $y-$coordinate at height $\geq 1$, all horizontal segments of $\pi$ must have height $\geq 1$ (Lemma 5). In other words, the length of every vertical segment must be an even integer. If the last vertical segment of $\pi$ were ending strictly between $NW$ and $NE$, then its cost would be at least 2, contradicting optimality. Thus, $\pi$ must pass through $NW$ or $NE$, and observationss (i) or (ii) complete the proof.

Now consider assertions (b) and (c). Let $\pi$ be any path from $d$ to $c^*$, and let $z$ be the point at which $\pi$ first meets the boundary of the rectangle formed by $\{E, W, SW, SE\}$. Let $\pi', \pi''$ denote the subpaths of $\pi$ from $d$ to $z$ and from $z$ to $c^*$, respectively. There are four possible cases :

- $z \in (E, W)$ : since all segments in $\pi'$ have height $\leq 0$, replacing $\pi'$ by $d \to O \to z$ cannot increase the overall cost. In the resulting path, one may further replace the subpath from $O$ to $c^*$ by $O \to b^* \to c^*$ without increasing the cost, by (iii). This shows :

$$\mathrm{cost}(\pi) \geq 2 + \gamma - 2^{-h} - \gamma^h.$$

- $z \in (W, SW)$ : all horizontal segments in $\pi'$ have height $\leq -1$, so $\mathrm{cost}(\pi') \geq \gamma^{-1}$. By (ii), one also has $\mathrm{cost}(\pi'') \geq \mathrm{cost}(z \to NW \to c^*)$. Combining these two facts yields

$$\mathrm{cost}(\pi) \geq 2\gamma + \gamma^{-1}.$$

- $z \in (SE, E)$ : replacing the subpath $\pi''$ by $z \to NE \to c^*$ cannot increase the overall cost, by part (a). In the resulting path, the subpath from $d$ to $E$ costs at least $\gamma^{-1} + \gamma(1 - 2^{-h})$, because the horizontal and vertical heights are $\leq -1$ and $\leq 1$, respectively. Thus,

$$\mathrm{cost}(\pi) \geq 2\gamma + \gamma^{-1}.$$

- $z \in (SW, SE)$ : all segments of $\pi'$ have height $\leq -1$ except those included in $[O, b]$, which have height 0. Thus,

$$\mathrm{cost}(\pi') - \mathrm{cost}(d \to b \to z) \geq (\gamma^{-1} - 1)\mathrm{len}\left([b, d] \setminus \pi'\right).$$

  Moreover, by part (a), $\mathrm{cost}(b \to z) + \mathrm{cost}(\pi'') \geq \mathrm{cost}(\pi_1)$. Thus,

$$\mathrm{cost}(\pi) \geq (\gamma^{-1} - 1)\mathrm{len}\left([b, d] \setminus \pi\right) + \mathrm{cost}(d \to b) + \mathrm{cost}(\pi_1).$$

Let us sum up: in the first three cases, the cost of $\pi$ exceeds that of our two candidates by at least 1, for $h$ sufficiently large. In the fourth case, the excess is at least $(\gamma^{-1} - 1)\mathrm{len}\left([b, d] \setminus \pi\right)$. This proves both (b) and (c), with $\eta = 2^{-h}(\gamma^{-1} - 1)$. ■

**Proof of Proposition 9.** In each basic $2^{h+1} \times 2^{h+1}$ square $G$ of $\mathbb{Z}^2_{h+1}$ there is a copy of the Figure 6 configuration; let $\Sigma_G$ be the corresponding small square. Let $B := \cup_G \Sigma_G$ be the union of those squares, for the fixed $h$ given by Lemma 10. Then for $i \geq 1$ let $B_i := \sigma_{2^{-i}} B$ be rescalings of $B$. Each $B_i$ has the same density, which by Lemma 10 is non-zero, and a straightforward use of the second Borel-Cantelli lemma (with sufficiently well-spaced values of $i$) shows that the set

$$A := \{z \in \mathbb{R}^2 : \ z \text{ in only finitely many } B_i\}$$

has area zero.

Now consider $z_1 \in A^c$. Then there exists a sequence $i_j = i_j(z_1) \to \infty$ such that $z_1 \in B_{i_j}$ and the associated $b_{i_j}(z_1) \to z_1$. Consider $z_2 \neq z_1$ and $(z_1^H, z_2^H) \to (z_1, z_2)$ as in (23). For $j$ larger than some $j_0(z_1, z_2)$, Lemma 10 implies that for all sufficiently large $H$ the route $\mathcal{R}_0(z_1^H, z_2^H)$ passes through $b_{i_j}(z_1)$. But the routes between the $b_{i_j}(z_1), j \geq 1$ are specified by the construction on $\cup_{H>-\infty}\mathbb{Z}_H^2$. It follows that, when $z_1$ and $z_2$ are both in $A^c$ we have convergence in the sense of (24) to a route $\mathcal{R}_0(z_1, z_2)$. ■

We digress to give the technical estimate that will show $\ell < \infty$ in this model.

**Lemma 12** *For the routes $\mathcal{R}_0$ in Proposition 9, take the union over points $\xi, \xi'$ of a rate-1 Poisson point process $\Xi(1)$ of the routes $\mathcal{R}_0(\xi, \xi')$, and let $\mathcal{S}^*$ be the intersection of that union with the interior of a unit square $U = [i, i+1] \times [j, j+1]$. Then the expected length of $\mathcal{S}^*$ is at most $2^{b+2}$, for $b$ as in Lemma 6.*

**Proof.** Lemma 6 was stated for $h \geq 0$ and vertices in $\mathbb{Z}^2$, but by scaling it holds for $h < 0$ and vertices in $\mathbb{R}^2$. Consider $h < 0$. Within $U$ there are $2^{-h-1}$ horizontal unit-length line segments at height $h$, and these can be split into $2^{-2h-1}$ segments of length $2^h$. Consider such a line segment, $\zeta$ say. It is in the interior of some rectangle of the form $[i2^{h+1+b}, (i+1)2^{h+1+b}] \times [j2^{h+1}, (j+1)2^{h+1}]$. By Lemma 6 applied with $h+1$, the only possible way that the segment $\zeta$ can be in a route $\mathcal{R}_0(\xi, \xi')$ is if $\xi$ or $\xi'$ is within the rectangle. (And the same holds for any piece of $\zeta$, by considering a sub-rectangle). The chance the Poisson process contains such a point is at most the area of the rectangle, which is $2^{2h+2+b}$.

So the contribution to mean length from a particular segment $\zeta$ is at most $2^h \times 2^{2h+2+b}$, and then the contribution from height-$h$ horizontal lines is at most $2^h \times 2^{2h+2+b} \times 2^{-2h-1} = 2^{h+1+b}$. Summing over $h \leq -1$ and adding the same contribution from vertical lines gives the bound $2^{2+b}$. ■

### 3.7 Completing the construction by forcing invariance

Proposition 9 gives paths $\mathcal{R}_0(z_1, z_2)$ when $z_1, z_2 \in A^c$. The process $\mathcal{R}_0$ cannot be translation- or rotation-invariant (in distribution), because the axes play a special role (infinite speed); though by construction the process is invariant under $\sigma_2$ (scaling space by a factor 2). But there is a standard way of trying to make translation-invariant random processes out of deterministic processes, by taking weak limits of random translations of the original process. In our setting this can be done fairly explicitly as follows. For $u \in \mathbb{R}^2$ let $T_u$ be the translation map $T_u(z) = u + z,\ z \in \mathbb{R}^2$ on points, and let $T_u$ act on routes in the natural way. Take $U_n$ uniform on the square $[0, 2^n]^2$, and couple the random variables $(U_n, n \geq 1)$ by setting $U_n = U_{n+1} \bmod 2^n$ coordinatewise. Define

$$\mathcal{R}^{(n)}(z_1, z_2) = T_{-U_n}(\mathcal{R}_0(z_1 + U_n, z_2 + U_n)). \tag{25}$$

In words, translate points by $U_n$, use $\mathcal{R}_0$ to define a route between the translated points, and then translate back to obtain a route between the original points.

Now the only way that $\mathcal{R}^{(n+1)}(z_1, z_2)$ could be different from $\mathcal{R}^{(n)}(z_1, z_2)$ is if the route $\mathcal{R}_0(z_1 + U_n, z_2 + U_n)$ intersects the boundary of the square $[0, 2^n]^2$, which, using Lemma 4, has chance $O(2^{-n})$. So we can define a random network $\mathcal{R}_{\text{t-i}}$ via the a.s. limits

$$\mathcal{R}_{\text{t-i}}\ (z_1, z_2) = \mathcal{R}^{(n)}(z_1, z_2) \text{ for all sufficiently large } n. \tag{26}$$

This process is translation-invariant, because for fixed $z \in \mathbb{R}^2$ the variation distance between the distributions of $U_n$ and $U_n + z \bmod 2^n$ tends to zero.

For $0 < c < \infty$ write $\sigma_c$ for the scaling map $z \to cz$ on $\mathbb{R}^2$, and recall that $\mathcal{R}_0$ is invariant under $\sigma_2$. Now for $\mathcal{R}^{(n)}$ at (25),

$$\begin{aligned}
\sigma_2 \mathcal{R}^{(n)}(z_1, z_2) &= \sigma_2 T_{-U_n} \mathcal{R}_*(z_1 + U_n, z_2 + U_n) \\
&= T_{-2U_n} \sigma_2 \mathcal{R}_*(z_1 + U_n, z_2 + U_n) \\
&\stackrel{d}{=} T_{-2U_n} \mathcal{R}_*(2z_1 + 2U_n, 2z_2 + 2U_n) \text{ by invariance of } \mathcal{R}_0 \text{ under } \sigma_2 \\
&\stackrel{d}{=} T_{-U_{n+1}} \mathcal{R}_*(2z_1 + U_{n+1}, 2z_2 + U_{n+1}) \text{ because } U_{n+1} \stackrel{d}{=} 2U_n \\
&= \mathcal{R}^{(n+1)}(2z_1, 2z_2).
\end{aligned}$$

Hence the distribution of the limit $\mathcal{R}_{\text{t-i}}$ is invariant under $\sigma_2$.

Of course our construction so far is not rotationally invariant, but applying a uniform random rotation to $\mathcal{R}_{\text{t-i}}$ gives a network $\mathcal{R}_{\text{r-i}}$ whose distribution is invariant under rotation, as well as preserving distributional invariance under translation and under $\sigma_2$. Finally, we get a process $\mathcal{R}$ with scale-invariant distribution by random rescaling via the scale-free distribution:

$$\mathcal{R}(z_1, z_2) = \sigma_{1/C}\mathcal{R}_{\text{r-i}}\ (Cz_1, Cz_2), \quad \mathbb{P}(C \in dc) = \tfrac{1}{c\ \log 2},\ 1 < c < 2. \quad (27)$$

This completes the construction of the *binary hierarchy model* $\mathcal{R}$. To check it satisfies the formal setup of a SIRSN in section 2.2, the only remaining issue is to check that the parameters $\mathbb{E}D_1, \ell$ and $p(1)$ are finite. For the former, Proposition 2 implies the corresponding bound in terms of Euclidean distance

$$\text{len } \mathcal{R}_0(z_1, z_2) \leq 2^{1/2} K_\gamma ||z_2 - z_1||_2$$

and this bound is unaffected by the transformations taking $\mathcal{R}_0$ to $\mathcal{R}$. So $\mathbb{E}D_1 \leq 2^{1/2}K_\gamma$. For $\ell$, in the notation of Lemma 12, the edge-intensity of $\cup_{\xi,\xi' \in \Xi(1)}\mathcal{R}_0(\xi, \xi')$ is at most $2^{b+2} + 2$, the "+2" terms arising from the edges of $\mathbb{Z}^2$. This edge-intensity is unaffected by the transformations taking $\mathcal{R}_0$ to $\mathcal{R}_{\text{r-i}}$ . Scaling by $C$ in (27) multiplies edge-intensity by $C$, so finally $\ell \leq (2^{b+2} + 2)\mathbb{E}C$. To bound $p(1)$, set $r(h) = 2^h(2^b + 1)$. Corollary 7 implies that, for routes $\mathcal{R}_0$, if an edge element is in a route between some two points at distance $\geq r(h)$ from the element, then the edge has height $\geq h$. The edge-intensity of edges with height $\geq h$ equals $2^{1-h}$. These quantities are unaffected by translation and rotation; and the scaling by $\sigma_C$ can at most increase the edge-intensity by 4. So the edge-intensity in $\mathcal{R}$ of $\mathcal{E}(\lambda, r(h))$ is $p(\lambda, r(h)) \leq 4 \cdot 2^{1-h}$ Choosing $h$ such that $r(h) < 1$ we deduce $p(1) < \infty$.

### 3.8 Remarks on section 3.

The "combinatorial" arguments in sections 3.1 - 3.4 are obviously specific to this model. But the property implicit in Lemma 10 (that there exist configurations in which all long routes from a small neighborhood exit the unit disc at the same point) is closely related to desirable structural properties of SIRSNs discussed in section 7.

Lemma 20 later shows that in general $\ell \leq 2p(1)$, so our argument above that $\ell < \infty$ could be omitted, though it is pleasant to have a self-contained construction.

## 4 Other possible constructions

The model in section 3 has some very special features, in particular that in any realization we see a (scaled and rotated) square lattice of roads. Below we outline two other constructions which, we conjecture, produce SIRSNs, the technical dificulty being to prove a.s. uniqueness of routes defined as minimum-cost paths.

### 4.1 The Poisson line process model

For each $m = 1, 2, 3, \ldots$ take a rate-1 Poisson line process, and attach Uniform$(m-1, m)$ marks to the lines; the union of all these is a Poisson line process with "mark measure" being Lebesgue measure on $(0, \infty)$. By a one-to-one mapping of marks one can transform to the mark measure with density $x^{-\gamma}$ on $0 < x < \infty$, where we take the parameter $2 < \gamma < \infty$. So in any small disc, there is some finite largest mark amongst lines intersecting the disc.

Picturing the lines as freeways and the marks as speeds, for any pair of points $z_1, z_2$ on the lines there is some finite minimum time $\mathrm{t}(z_1, z_2)$ over all routes from $z_1$ to $z_2$, and analogous to Lemma 8 one can show (Wilf Kendall: personal communication) that this function extends to a random continuous function $\mathrm{t}(z_1, z_2)$ on the plane. The technical difficulty is to show that for given $(z_1, z_2)$ there is an a.s. unique route attaining that time; if that were proved, establishing the remaining properties required of an SIRSN would be straightforward. In particular, scale-invariance would follow from the form $x^{-\gamma}$ of the mark density.

### 4.2 A dynamic proximity graph model

This potential construction of a SIRSN is based on a space-time Poisson point process $(\Xi(\lambda), 0 < \lambda < \infty)$. Note that to study such a SIRSN one would use an independent Poisson point process to define $\mathcal{S}(\lambda)$. Note also that the corresponding "static" model, called the *Gabriel* network, is a member of the family of *proximity graphs* described in [15, 5]; any family member could be used in the construction below.

Here's the construction rule.

> When a point $\xi$ arrives at time $\lambda$, consider in turn each existing point $\xi' \in \Xi(\lambda-)$, and create an edge $(\xi, \xi')$ if the disc with diameter $(\xi, \xi')$ contains no other point of $\Xi(\lambda-)$.

Write $\mathbb{G}(\lambda)$ for the time-$\lambda$ network on points $\Xi(\lambda)$. Note the automatic scale-invariance property

> the action of $\sigma_c$ on $\mathbb{G}(\lambda)$ gives a network distributed as $\mathbb{G}(c^{-2}\lambda)$.

Now fix a parameter $0 \leq \gamma < \gamma_*$ for some sufficiently small $\gamma_* > 0$ and view an edge created at time $\lambda$ as a road with speed $\lambda^{-\gamma}$. Defining routes in $\mathbb{G}(\lambda)$ as minimum-time paths, it seems intuitively plausible, as in the Poisson line process model, that that we can extend the minimum-time function on $\cup_\lambda \Xi(\lambda)$ to a continuous function $\mathrm{t}(z_1, z_2)$ and then prove there is an a.s. unique route attaining that time. Again, if that were proved, establishing the remaining properties required of an SIRSN would be straightforward. In particular, scale-invariance would follow from the fact that the construction rule is scale-invariant.

# 5 Properties of weak SIRSNs

In this section we study properties that hold for any weak SIRSN, that is when we do not require (20) but instead require (16). These are essentially properties of the sampled subnetworks $\mathcal{S}(\lambda)$ for fixed $\lambda$ – we cannot get $\lambda \to \infty$ results.

## 5.1 No straight edges at typical points

If a point $\xi$ of $\Xi(\lambda)$ is the start of some straight line segment of length $\geq r$ in $\mathcal{S}(\lambda)$ then consider the subroutes of length exactly $r$ from $\xi$. The edge process of such subroutes has some edge-intensity $\iota(\lambda, r)$. In independent copies of $\Xi(1)$ these edge-processes cannot have any positive-length overlap. So by regarding $\Xi(n)$ as the union of $n$ independent copies of $\Xi(1)$ we have $\iota(n, r) = n\iota(1, r)$. But by the general scaling property (12)

$$\iota(\lambda, r) = \lambda^{1/2}\iota(1, r\lambda^{1/2}). \tag{28}$$

Since $\iota(1, r) \leq \ell < \infty$ these two different scaling relations imply $\iota(1, r) = 0$ for all $r > 0$.

This proves (a) below; note the consequence (b), implied by the definition of *feasible path* in the section 2.2 setup.

**Proposition 13** *$S(\lambda)$ has the following properties a.s.*
*(a) $S(\lambda)$ contains no line segment $[\xi, z]$ of positive length, for any $\xi \in \Xi(\lambda)$.*
*(b) The route $\mathcal{R}(\xi_1, \xi_2)$ between two points of $\Xi(\lambda)$ does not pass through any third point $\xi_3$ of $\Xi(\lambda)$.*

## 5.2 Singly and doubly infinite geodesics

Recall from section 2.3 that a *singly infinite geodesic* from a point $\xi_0$ in $\mathcal{S}(\lambda)$ is an infinite path, starting from $\xi_0$, such that any finite portion of the path is a subroute of some route $\mathcal{R}(\xi_0, \xi)$. Lemma 1 showed

$$\text{There is a.s. at least one singly infinite geodesic from each point of } \mathcal{S}(\lambda). \tag{29}$$

A *doubly infinite geodesic* in $\mathcal{S}(\lambda)$ is a path $\pi$ which is an increasing union of segments $\pi_k$, where each $\pi_k$ is a segment of some route $\mathcal{R}(\xi_k, \xi'_k)$ between two points of $\Xi(\lambda)$, and both endpoints of $\pi_k$ go to infinity.

Previous work on very different (e.g. percolation-type [19]) networks suggests there may be a general principle:

> In natural models of random networks on $\mathbb{R}^2$ or $\mathbb{Z}^2$, doubly infinite geodesics do not exist.

Proposition 14 proves this for weak SIRSNs based on a simple scaling argument. Note however this argument depends implicitly upon our assumption $\ell < \infty$ which seems rather special to our setting.

Recall the setup of (17, 18). $\mathcal{E}(\lambda, r) \subset \mathcal{S}(\lambda)$ is the set of points $z$ in edges of $\mathcal{S}(\lambda)$ such that $z$ is in the route $\mathcal{R}(\xi, \xi')$ for some $\xi, \xi'$ of $\Xi(\lambda)$ such that $\min(|z - \xi|, |z - \xi'|) \geq r$. And $p(\lambda, r)$ is the edge-intensity of $\mathcal{E}(\lambda, r)$. By scaling,

$$p(\lambda, r) = \lambda^{1/2} p(1, r\lambda^{1/2}) \tag{30}$$

**Proposition 14** *$p(\lambda, r) \to 0$ as $r \to \infty$. In particular, $\mathcal{S}(\lambda)$ has a.s. no doubly infinite geodesics.*

**Proof.** For fixed $\lambda$ the edge-processes $\mathcal{E}(\lambda, r)$ can only decrease as $r$ increases, and the limit $\mathcal{E}(\lambda, \infty) := \cap_r \mathcal{E}(\lambda, r)$ is by definition the set of path elements in doubly infinite geodesics. This limit has edge-intensity $p(\lambda, \infty) = \lim_{r \to \infty} p(\lambda, r) \geq 0$. So it is enough to prove $p(\lambda, \infty) = 0$. Suppose not. Then by the scaling relation (30)

$$p(\lambda, \infty) = \lambda^{1/2} p(1, \infty), \quad 0 < \lambda < \infty.$$

We claim that in fact

$$\mathcal{E}(\lambda, \infty) = \mathcal{E}(1, \infty) \text{ a.s. for } \lambda < 1,$$

which (because we know $p(1, \infty) < \infty$) implies $p(1, \infty) = 0$ and completes the proof.

To prove the claim, note that for any finite-length segment $\pi_0$ of a doubly infinite geodesic in $\mathcal{S}(1)$, there are an infinite number of distinct pairs $\xi_j, \xi'_j$ of $\Xi(1)$ such that $\mathcal{R}(\xi_j, \xi'_j)$ contains $\pi_0$, and for each pair there is chance $\lambda^2$

that both points are in $\Xi(\lambda)$. These events are independent (because $\Xi(\lambda)$ is obtained from $\Xi(1)$ by independent sampling) so a.s. an infinite number of pairs $\xi_j, \xi'_j$ are in $\Xi(\lambda)$, implying that $\pi_0$ is in a doubly infinite geodesic of $\mathcal{S}(\lambda)$. ■

**Remark.** The limit used here is different from the limit $p(r) := \lim_{\lambda\to\infty} p(\lambda, r)$ featuring in assumption (20).

### 5.3 Marginal interpretation of $\ell$

Recall $\ell$ is defined as the edge-intensity of $\mathcal{S}(1)$, which is the subnetwork on a rate-1 Poisson point process $\Xi(1)$. Now augment the network $\mathcal{S}(1)$ by including the point at the origin and the routes from the origin to each $\xi \in \Xi(1)$. The newly added edges have some random total length $L$.

**Proposition 15** $\mathbb{E}L = \ell/2$.

**Proof.** Recall (22) the scaling relation $\ell(\lambda) = \lambda^{1/2}\ell$, where $\ell(\lambda)$ is the edge-intensity of the subnetwork $\mathcal{S}(\lambda)$ of a Poisson process of point-intensity $\lambda$ . Differentiating with respect to $\lambda$,

$$\ell'(1) = \tfrac{1}{2}\ell.$$

So we need to show $\mathbb{E}L = \ell'(1)$.

Consider the space-time Poisson point process $(\Xi(\lambda), 0 < \lambda < \infty)$ from section 2.2. Each arriving point creates some additional network length, say $\tilde{L}(\xi)$, and for a point arriving at time $\lambda$, write $\tilde{\ell}(\lambda)$ for the mean additional network length. Now

$$\ell(\lambda_0) = \mathbb{E} \sum_{\xi\in\Xi(\lambda_0)\cap[0,1]^2} \tilde{L}(\xi) = \int_0^{\lambda_0} \tilde{\ell}(\lambda)\ d\lambda$$

and so $\ell'(1) = \tilde{\ell}(1)$.

### 5.4 A lower bound on network length

Write $\Delta$ for the parameter $\mathbb{E}D_1$ of a SIRSN. Write $\ell_*(\Delta)$ for the minimum possible value of $\ell$ in a SIRSN with a given value of $\Delta$.

**Proposition 16** $\ell_*(\Delta) = \Omega((\Delta - 1)^{-1/2})$ *as* $\Delta \downarrow 1$.

The proof is based on a bound (Proposition 17) involving the geometry of deterministic paths, somewhat similar to bounds used in [8] section 4. Figure 7 illustrates the argument to be used.

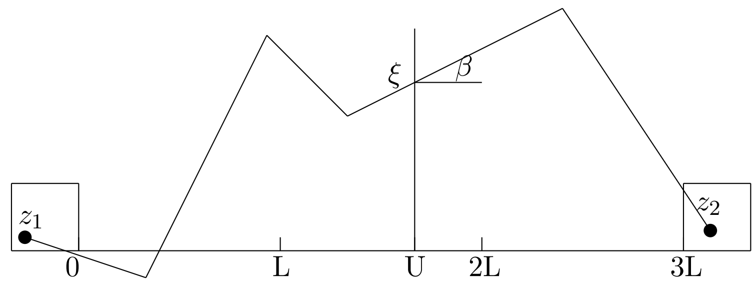


**Figure 7.**

**Proposition 17** *Let $\alpha, L, D$ and $\theta_0$ be positive reals satisfying $\theta_0 < \pi/2$ and*

$$2\sqrt{(D-1)^2 + (\tfrac{3}{2}L+1)^2} \;=\; (1+2\alpha)(3L+2) \tag{31}$$

$$L\left(\frac{1}{\cos\theta_0} - 1\right) \;=\; 4\alpha\sqrt{(3L+2)^2+1}. \tag{32}$$

*Let $\mathcal{R}$ be a route from some point $z_1$ in the unit square $[-1,0]\times[0,1]$ to some point $z_2$ in the unit square $[3L, 3L+1]\times[0,1]$, and suppose*

$$\operatorname{len}(\mathcal{R}) \le (1+2\alpha)|z_2 - z_1|. \tag{33}$$

*Take $U$ uniform random on $[L, 2L]$. The route $\mathcal{R}$ first crosses the vertical line $\{(U,y), -\infty < y < \infty\}$ at some random point $(U, \xi(U))$ and at some angle $\beta(U) \in (-\frac{\pi}{2}, \frac{\pi}{2})$ relative to horizontal. Then*
*(i) $|\xi(U)| \le D$.*
*(ii) $\mathbb{P}(|\beta(U)| \le \theta_0) \ge \frac{1}{2}$.*

**Proof.** The maximum possible value of $\xi(U)$ arises in the case where $z_1 = (-1,1)$, $z_2 = (3L+1, 1)$, $U = \frac{3}{2}L$, the route consists of straight lines from $z_1$ to $(U,\xi(U))$ to $z_2$, and the route-length attains equality in (33). In this case the value of $\xi(U)$ is the quantity $D$ satisfying (31), establishing (i).

Writing $\beta(u)$ for the angle (relative to horizontal) of the route at $x$-coordinate $u$, then the length ($\Lambda$, say) of the route between $x$-coordinates $L$ and $2L$ equals $\int_L^{2L} \frac{1}{\cos\beta(u)}\, du$. This implies

$$\Lambda - L \ge (\tfrac{1}{\cos\theta_0} - 1) \times L\mathbb{P}(\beta(U) \ge \theta_0).$$

But by considering excess length (relative to a horizontal route), (33) implies

$$\Lambda - L \le 2\alpha |z_2 - z_1| \le 2\alpha \sqrt{(3L+2)^2 + 1}.$$

Combining these inequalities gives a lower bound on $\mathbb{P}(\beta(U) \ge \theta_0)$ which equals 1/2 when $\theta_0$ satisfies (32), establishing (ii). ∎

**Proof of Proposition 16** Consider a SIRSN with parameters $\ell$ and $\Delta$ and with induced subnetwork $\mathcal{S}$ on a Poisson point process $\Xi$. Set $\alpha = \Delta - 1$. Suppose we can choose $L, D, \theta_0$ to satisfy, along with the given $\alpha$, the equalities (31,32) – note this leaves us one degree of freedom.

With probability $(1-e^{-1})^2$ there are points $z_1$ and $z_2$ of the Poisson process in the unit squares $[-1,0] \times [0,1]$ and $[3L, 3L+1] \times [0,1]$. By Markov's inequality and the definition of $\Delta$, with probability at least 1/2 the route $\mathcal{R}(z_1, z_2)$ has length at most $(1+2\alpha)|z_2 - z_1|$. Applying Proposition 17 we deduce that, with probability $\ge (1-e^{-1})^2/4$, the network $\mathcal{S}$ contains an edge that crosses the random vertical line $\{(U,y) : \ -\infty < y < \infty\}$ at some point $(U, \xi(U))$ with $-D \le \xi(U) \le D$ and crosses at some angle $\beta(U) \in (-\theta_0, \theta_0)$ relative to horizontal.

If we translate vertically by $2D$, to consider routes between the unit squares $[-1,0] \times [2D, 2D+1]$ and $[3L, 3L+1] \times [2D, 2D+1]$, then the potential crossing points (using the same r.v. $U$) for the translated and untranslated cases are distinct. Now by considering translates by all multiples of $2D$, and noting that the distribution of crossings of the random vertical line $\{(U,y) : \ -\infty < y < \infty\}$ is the same as for the $y$-axis, we have shown

> the mean intensity of crossings of the network $\mathcal{S}$ over the $y$-axis at angles $\in (-\theta_0, \theta_0)$ relative to horizontal is at least $\frac{(1-e^{-1})^2}{8D}$.

The stochastic geometry identities (2, 3) relates this mean intensity to the parameter $\ell$ via

$$\text{this mean intensity} = \tfrac{\ell}{\pi} \int_{-\theta_0}^{\theta_0} \cos\theta \ d\theta \le \tfrac{2\ell\theta_0}{\pi}.$$

Combining with the previous inequality we find

$$\ell \ge \frac{1}{21 D \theta_0}.$$

Now set $L = \alpha^{-1/2}$ and consider the solutions of (31,32) in the limit as $\alpha \downarrow 0$: we find that solutions exist with

$$\theta_0 \sim \sqrt{24\alpha}; \quad D \to 10$$

which establishes Proposition 16.

### 5.5 The minimum value of $\ell$ and the Steiner tree constant

Take $k$ uniform random points $Z_1, \ldots, Z_k$ in a square of area $k$ and consider the length $L_{ST}(k)$ of the Steiner tree (the minimum-length connected network) on $Z_1, \ldots, Z_k$. Well-known subadditivity arguments [21, 25] imply that $\mathbb{E}L_{ST}(k) \sim c_{\mathrm{ST}} k$ for some constant $0 < c_{\mathrm{ST}} < \infty$. One can define $c_{\mathrm{ST}}$ equivalently (see [4] for results of this kind) as the infimum of $c$ such that there exists a translation-invariant connected random network over $\Xi(1)$ with edge-intensity $c$. From the latter description it is obvious that in any SIRSN we have $\ell \geq c_{\mathrm{ST}}$. So the overall infimum

$$\ell_* := \text{ infimum of } \ell \text{ over all SIRSNs} \tag{34}$$

satisfies $\ell_* \geq c_{\mathrm{ST}}$, and below we outline an argument that the inequality is strict. First we derive some simple lower bounds on $c_{\mathrm{ST}}$ and $\ell_*$.

(i) Write $b(\xi)$ for the distance from $\xi$ to its closest neighbor in $\Xi(1)$. The discs of center $\xi$ and radius $b(\xi)/2$ are disjoint as $\xi$ varies and must contain network length at least $b(\xi)/2$, so

$$c_{\mathrm{ST}} \geq \tfrac{1}{2}\mathbb{E}b(\xi) = \tfrac{1}{4}.$$

(ii) In a network of edge-intensity $c$, (2) shows the mean number of edges crossing circle$(0, r)$ equals $2\pi r \times 2\pi^{-1} c = 4rc$. If there is a point of $\Xi(1)$ inside disc$(0, r)$ then there must be some such crossing edge, so

$$1 - \exp(-\pi r^2) \leq 4rc.$$

So

$$c_{\mathrm{ST}} \geq \sup_r \frac{1 - \exp(-\pi r^2)}{4r} \approx 0.283.$$

(iii) We can get a better bound on $\ell_*$ by using Proposition 15 as follows. Using the intensity calculation above, in a network of edge-intensity $\ell$ the probability that no edge crosses circle$(0, r)$ is at least $1 - 4r\ell$. When a new point arives at $\xi$ in the $\mathcal{S}(\lambda)$ process at time $\lambda = 1$, if no existing edges cross circle$(\xi, r)$ then the added network length $L$ is at least $r$. So

$$\mathbb{E}L \geq \sup_r r(1 - 4r\ell) = \tfrac{1}{16\ell}.$$

But Proposition 15 says $\ell = 2\mathbb{E}L$ and so we have shown

$$\ell_* \geq \sqrt{1/8} \approx 0.353. \tag{35}$$

One could no doubt obtain small improvements by similar arguments.

Here is an outline argument that $\ell_* > c_{\mathrm{ST}}$.
(i) In the Steiner tree on the Posson point process $\Xi(1)$, vertices of degree $> 1$ have non-zero density, and their edges meet at some varying angles, whereas at the Steiner points (non-vertex junctions) edges must meet at 120 degree angles.
(ii) If there were a SIRSN with $\ell \approx c_{\mathrm{ST}}$, then $\mathcal{S}(1)$ would have essentially the properties (i). But then in $\mathcal{S}(1/2)$, obtained by deleting half the vertices of $\Xi(1)$ to get $\Xi(1/2)$, some of the deleted vertices would remain as junction points. The "varying angles" property implies the edge-intensity $\ell(1/2)$ of $\mathcal{S}(1/2)$ is strictly larger than that of the Steiner tree on $\Xi(1/2)$, contradicting the scale-invariance property that the edge-intensities of $\mathcal{S}(1)$ and the Steiner tree on $\Xi(1)$ are essentially equal.

# 6 General SIRSNs and their properties

In this section we study properties of $\mathcal{S}(\lambda)$ in the $\lambda \to \infty$ limit, for a general SIRSN. Roughly speaking, this is studying "the whole SIRSN" instead of sampled subnetworks, and such results depend on assumption (20).

Recall again the setup from (17) - (20). So $p(\lambda, r)$ is the edge-intensity of $\mathcal{E}(\lambda, r)$, which is the process of points $z$ in edges of $\mathcal{S}(\lambda)$ such that $z$ is in the route $\mathcal{R}(\xi, \xi')$ for some $\xi, \xi'$ in $\Xi(\lambda)$ such that $\min(|z-\xi|, |z-\xi'|) \geq r$. Recall also from (30) the scaling relation $p(\lambda, r) = \lambda^{1/2} p(1, r\lambda^{1/2})$. Defining

$$p(r) := \lim_{\lambda \to \infty} p(\lambda, r) < \infty \tag{36}$$

the assumption (20) that $p(1) < \infty$ and scaling imply

$$p(r) = p(1) \times r^{-1}, \quad 0 < r < \infty. \tag{37}$$

## 6.1 A connectivity bound

Assumption (20) has a direct implication for the qualitative structure of a SIRSN: all the routes linking two regions, once they get away from a neighborhood of the regions, use only a finite number of different paths. We first give a version of this result in terms of discs.

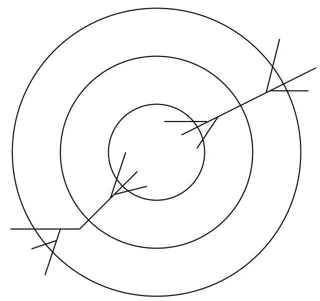

**Figure 8.** Schematic for routes from inside disc$(0,1/2)$ to outside disc$(0,3/2)$ crossing the unit circle.

**Proposition 18** *Take $0 < r < 1$ and let $N(\lambda, r)$ be the number of distinct points on the unit circle at which some route $R(\xi, \xi')$ between some $\xi \in \Xi(\lambda) \cap \mathrm{disc}(0, 1-r)$ and some $\xi' \in \Xi(\lambda) \cap (\mathbb{R}^2 \setminus \mathrm{disc}(0, 1+r))$ crosses the unit circle. Then*

$$\mathbb{E} \lim_{\lambda \to \infty} N(\lambda, r) \le 4p(1)\ r^{-1}.$$

**Proof.** Any crossing point is in $\mathcal{E}(\lambda, r)$ and so by identity (2)

$$\mathbb{E}N(\lambda, r) \le 2\pi \times 2\pi^{-1} p(\lambda, r) < \infty$$

and the result follows from (37). ■

The following general version can be proved similarly.

**Proposition 19** *Let $\varepsilon > 0$ and let $K_1, K_2$ be compact sets whose $\varepsilon$-neighborhoods $K_1^\varepsilon, K_2^\varepsilon$ are disjoint. For $z_1 \in K_1, z_2 \in K_2$ let $\mathcal{R}_\varepsilon(z_1, z_2)$ be the subroute of $\mathcal{R}(z_1, z_2)$ crossing from the boundary of $K_1^\varepsilon$ to the boundary of $K_2^\varepsilon$. Let $N(\lambda, \varepsilon, K_1, K_2)$ be the number of distinct paths amongst the set $\{\mathcal{R}_\varepsilon(\xi_1, \xi_2) : \xi_i \in \Xi(\lambda) \cap K_i\}$. Then*

$$\mathbb{E} \lim_{\lambda \to \infty} N(\lambda, \varepsilon, K_1, K_2) < \infty.$$

### 6.2 A bound on normalized length

**Lemma 20** *$\ell \le 2p(1)$.*

**Proof.** Define

$$\mathcal{R}_\delta(\xi, \xi') = \mathcal{R}(\xi, \xi') \cap (\mathrm{disc}(\xi, \delta) \cup \mathrm{disc}(\xi', \delta))$$

in words, the part of the route that is *within* distance $\delta$ from one or both endpoints. Then define

$$\widehat{\mathcal{S}}_\delta(\lambda) = \cup_{\xi,\xi' \in \Xi(\lambda)} \mathcal{R}_\delta(\xi, \xi').$$

Note that clearly

$$S(\lambda) \setminus \widehat{\mathcal{S}}_1(\lambda) \subseteq \mathcal{E}(\lambda, 1). \tag{38}$$

By considering $\lambda = 1$,

$$\ell \leq p(1,1) + \iota(\widehat{\mathcal{S}}_1(1))$$

where $\iota(\cdot)$ denotes edge-intensity. Now write

$$\iota(\widehat{\mathcal{S}}_1(1)) = \sum_{k \geq 1} \iota(\widehat{\mathcal{S}}_{2^{1-k}}(1) \setminus \widehat{\mathcal{S}}_{2^{-k}}(1)).$$

For fixed $k \geq 1$, scaling by $2^k$ gives

$$\begin{aligned} \iota(\widehat{\mathcal{S}}_{2^{1-k}}(1) \setminus \widehat{\mathcal{S}}_{2^{-k}}(1)) &= 2^{-k} \iota(\widehat{\mathcal{S}}_2(2^{-2k}) \setminus \widehat{\mathcal{S}}_1(2^{-2k})) \\ &\leq 2^{-k} p(2^{-2k}, 1) \text{ by (38).} \end{aligned}$$

So

$$\ell \leq \sum_{k \geq 0} 2^{-k} p(2^{-2k}, 1) \leq \sum_{k \geq 0} 2^{-k} p(1).$$

■

### 6.3 The network $\mathcal{E}(\infty, r)$ of major roads

Intuitively, the point of assumption (20) and the scaling relation (37) is that we can define a proces $\mathcal{E}(\infty, r) := \cup_{\lambda < \infty} \mathcal{E}(\lambda, r)$ which must have edge-intensity $p(r) = p(1)/r$, and that in results like Proposition 18 we can replace $\lim_{\lambda \to \infty} N(\lambda, r)$ by $N(\infty, r)$. We don't want to give details of a completely rigorous treatment, but let us just suppose we can set up $\mathcal{E}(\infty, r)$ as a random element of some suitable measurable space, as we did for $\mathcal{S}(\lambda)$ in section 2.3.

The conceptual point is that $\mathcal{S}(\lambda)$ and $\mathcal{E}(\lambda, r)$ depend on the external randomization, that is on the fact that we were studying a SIRSN via the random points $\Xi(\lambda)$, but as outlined below $\mathcal{E}(\infty, r)$ doesn't depend on such external randomization. Intuitively this is simply because $\cup_\lambda \Xi(\lambda)$ is dense in $\mathbb{R}^2$; we outline a measure-theoretic argument below.

**Proposition 21** *The FDDs* $(\mathbf{span}(z_1, \ldots, z_k))$ *of a SIRSN can be extended to a joint distribution, of these FDDs jointly with a random process* $\mathcal{E}^*(\infty, r)$*, such that, for any space-time PPP* $(\Xi(\lambda), 0 < \lambda < \infty)$ *independent of the FDDs, we have* $\mathcal{E}^*(\infty, r) := \cup_{\lambda<\infty}\mathcal{E}(\lambda, r)$ *a.s.*

**Outline proof.** For a suitable formalization of "random subset of $\mathbb{R}^2$" we have the implication

> if $\mathcal{A}_1$ and $\mathcal{A}_2$ are i.i.d. random subsets, and if $\mathcal{A}_1 \cup \mathcal{A}_2 \subseteq_{a.s.} \mathcal{A}' \stackrel{d}{=} \mathcal{A}_1$, then $\mathcal{A}_1 = A$ a.s. for some non-random subset $A$

and then the corresponding "conditional" implication

> if $Z$ is a random element of some space, if $\mathcal{A}_1$ and $\mathcal{A}_2$ are random subsets conditionally i.i.d. given $Z$, and if $\mathcal{A}_1 \cup \mathcal{A}_2 \subseteq_{a.s.} \mathcal{A}'$ where $(Z, \mathcal{A}') \stackrel{d}{=} (Z, \mathcal{A}_1)$, then $\mathcal{A}_1 = \mathcal{A}$ a.s. for some $Z$-measurable random subset $A$.

So take two independent space-time PPPs $\Xi^1(\lambda), \Xi^2(\lambda)$ and use a measure-preserving bijection $[0, \infty) \cup [0, \infty) \to [0, \infty)$ to define another space-time PPP $\Xi'(\lambda)$ in terms of $\Xi^1$ and $\Xi^2$. The associated networks satisfy

$$\mathcal{E}^1(\infty, r) \cup \mathcal{E}^2(\infty, r) = \mathcal{E}'(\infty, r) \stackrel{d}{=} \mathcal{E}^1(\infty, r)$$

and this holds jointly with the FDDs of the SIRSN. Since $\mathcal{E}^1(\infty, r)$ and $\mathcal{E}^2(\infty, r)$ are conditionally i.i.d. given the SIRSN. Proposition 21 follows from the general "conditional implication" above.

### 6.4 Transit nodes and shortest path algorithms

Here we make a connection with the "shortest path algorithms" literature mentioned in section 1.4.

Fix $h$ and take the square grid of lines with inter-line spacing equal to $h$. Define $\mathcal{T}_h$ to be the set of points of intersection of $\mathcal{E}(\infty, h)$ with that grid.

**Lemma 22** *(i)* $\mathcal{T}_h$ *has point-intensity* $4\pi^{-1}h^{-2}p(1)$.
*(ii) For each* $z \in \mathbb{R}^2$ *there is a subset* $T_z$ *of* $\mathcal{T}_h$*, of mean size* $24\pi^{-1}p(1)$*, and with* $|z' - z| \le 2^{3/2}h$ *for each* $z' \in T_z$*, such that for each pair* $z_1, z_2$ *with* $|z_2 - z_1| > 3h$ *the route* $\mathcal{R}(z_1, z_2)$ *passes through some point of* $T_{z_1}$ *and some point of* $T_{z_2}$.

**Proof.** The grid has edge-intensity $2h^{-1}$, so from (2) the point-intensity of $\mathcal{T}_h$ is $2\pi^{-1} \times p(h) \times 2h^{-1}$, and (i) follows from scaling (37).

For any starting point $z$ consider the closest grid intersection $(ih, jh)$. Then $z$ is in some square with corner $(ih, jh)$, say the square $[(i-1)h, ih] \times [jh, (j+1)h]$. Let $T_z$ be the set of points of intersection of $\mathcal{E}(\infty, h)$ with the concentric square $S_z = [(i-2)h, (i+1)h] \times [(j-1)h, (j+2)h]$. This square has boundary length $12h$ and so the mean size of $T_z$ equals $2\pi^{-1} \times p(h) \times 12h = 24\pi^{-1}p(1)$. By construction

$$\tfrac{3}{2}h < |z' - z| \leq 2^{3/2}h \text{ for each } z' \text{ on the boundary of} S_z$$

and in particular for each $z' \in T_z$. If $|z_2 - z_1| > 3h$ then the squares $S_{z_1}$ and $S_{z_2}$ do not overlap, and the points $z_1'$ and $z_2'$ at which the route crosses their boundaries are in $\mathcal{T}_h$. ■

**Informal algorithmic implications** One cannot rigorously relate our "continuum" setup to discrete algorithms, but in talks we present the following informal calculation. For the real-world road network in a country we have empirical statistics

- $A$: area of country
- $\eta$: average number of road segments per unit area
- $M = \eta A$: total number of road segments in country
- $p(r)$: "length per unit area" of the subnetwork consisting of segments on routes with start/destination each at distance $> r$ from the segment.

For a real-world network there is an inconsistency between scale-invariance and having a finite number $\eta$ of road segments per unit area, but let us imagine approximate scale-invariance over scales of say 2 - 100 miles, and modify a scale-invariant model by deleting road segments of very short length. In what follows it is helpful to imagine the unit of length to be (say) 20 miles.

Fix $r$. Lemma 22 (with $h = r$) suggests that in the real-world network we can find transit nodes such that there are $O(p(1))$ transit nodes within distance $O(r)$ of a typical point. If so then we can analyze the algorithmic procedure outlined in section 1.4. The local search involves a region of radius $r$ and hence with $O(\eta r^2)$ edges. Regarding the time-cost of a single Dijkstra search as $c_1 \times$ (number of edges), the time-cost of finding the route to each local transit node is $O\left(c_1\ (\eta r^2)p(1)\right)$. Transit nodes have point-intensity

$O(p(1)/r^2)$, so the total number is $O(Ap(1)/r^2)$. Regard the space-cost of storing a $k \times k$ matrix of inter-transit-node routes as $c_2 k^2$; so this space-cost is $O\left(c_2\, (p(1)A/r^2)^2\right)$. Summing the two costs and optimizing over $r$, the optimal cost is $O(c_1^{2/3} c_2^{1/3} \eta^{2/3} A^{2/3} p^{4/3}(1)) = O(c_1^{2/3} c_2^{1/3} p^{4/3}(1) M^{2/3})$ and this $O(M^{2/3})$ scaling represents the improvement over the $O(M)$ scaling for Dijkstra. The corresponding optimal number of transit nodes is $O((c_1/c_2)^{1/3} p^{2/3}(1) M^{1/3})$. The latter has a more interpretable formulation. If the only alternative algorithms were a Dijkstra search of cost $c_1 \times$ (number of edges) or table look-up of cost $c_2 \times$ (number of edges)$^2$ , then there would be some critical number of edges at which one should switch between them, and this is just the solution $m_{\text{crit}}$ of $c_1 m_{\text{crit}} = c_2 m_{\text{crit}}^2$. So the optimal number of transit nodes is $O(m_{\text{crit}}^{1/3} p^{2/3}(1) M^{1/3})$.

### 6.5 Number of singly infinite geodesics

Write $\mathcal{S}^*(\lambda)$ for the spanning subnetwork obtained from $\mathcal{S}(\lambda)$ by adding a city at the origin $\mathbf{0}$. This process inherits the scaling-invariance property (12) of $\mathcal{S}(\lambda)$. We know from (29) that at least one singly infinite geodesic from $\mathbf{0}$ exists. The set of all singly infinite geodesics in $\mathcal{S}^*(\lambda)$ from $\mathbf{0}$ forms *a priori* a tree, because two geodesics that branch cannot re-join, by route compatability property (iv) from section 2.2. So consider

$$q(\lambda, r) := \mathbb{E}(\text{number of distinct points at which some singly infinite geodesic in } \mathcal{S}^*(\lambda) \text{ from } \mathbf{0} \text{ first crosses the circle of radius } r).$$

What we know in general is

$$1 \le q(\lambda, r) \le \infty; \quad r \to q(\lambda, r) \text{ is increasing}; \quad \lambda \to q(\lambda, r) \text{ is increasing}$$

and the scaling property gives

$$q(\lambda, r) = q(1, r\lambda^{1/2}). \tag{39}$$

So the $\lambda \to \infty$ limit $q(\infty, r) := \lim_{\lambda\to\infty} q(\lambda, r)$ exists (maybe infinite), and the scaling property implies

$$q(\infty, r) = q(\infty, 1) \in [1, \infty], \quad 0 < r < \infty.$$

So consider the property

$$q(\infty, 1) < \infty. \tag{40}$$

By applying Proposition 18 with $r \approx 1$ we see

$$q(\infty, 1) \le 4p(1). \tag{41}$$

So (36) implies (40). So we have shown the following.

**Corollary 23** *As $\lambda \to \infty$ the number of singly infinite geodesics in $\mathcal{S}^*(\lambda)$ from $\mathbf{0}$ increases to a finite limit number (perhaps a random number with finite mean) $G$. Moreover, if $G > 1$ then these geodesics branch at $\mathbf{0}$.*

# 7 Unique singly-infinite geodesics and continuity

For a SIRSN, let us call the property $G = 1$ a.s. (in the notation of Corollary 23 above) the *unique singly-infinite geodesics* property. It is conceivable that this property always holds – we record this later in Open Problem 32. Uniqueness of geodesics is closely related to continuity of routes $\mathcal{R}(z_1, z_2)$ as $(z_1, z_2)$ vary, as will be seen in section 7.2.

## 7.1 Equivalent properties

Here we show that several properties, the simplest being (42), are equivalent to the unique singly-infinite geodesics property. We will give definitions and proofs as we proceed, and then summarize as Proposition 24.

Consider two independent uniform random points $U_1, U_2$ in disc$(\mathbf{0}, 1)$. By the route-compatability property, the intersection of $\mathcal{R}(\mathbf{0}, U_1)$ and $\mathcal{R}(\mathbf{0}, U_2)$ is a sub-route from $\mathbf{0}$ to some *branchpoint* $B_{1,2}$, where either $B_{1,2} \neq \mathbf{0}$ or the intersection consists of the single point $\mathbf{0}$ (in which case, set $B_{1,2} = \mathbf{0}$). So we can define a property

$$\mathbb{P}(B_{1,2} = \mathbf{0}) = 0. \tag{42}$$

**Unique singly-infinite geodesics imply (42).** Suppose (42) fails. Then there exists $\varepsilon > 0$ such that, for independent random points $U_1^1, U_2^1$ in disc$(\mathbf{0}, 1) \setminus$ disc$(\mathbf{0}, \varepsilon)$, their branchpoint $B_{1,2}^1$ satisfies $\mathbb{P}(B_{1,2}^1 = \mathbf{0}) \geq \varepsilon$. Scaling by $\varepsilon^{-m}, m \geq 1$ and using scale-invariance, for independent random points $U_1^m, U_2^m$ in disc$(\mathbf{0}, \varepsilon^{-m}) \setminus$ disc$(\mathbf{0}, \varepsilon^{1-m})$, their branchpoint $B_{1,2}^m$ satisfies $\mathbb{P}(B_{1,2}^m = \mathbf{0}) \geq \varepsilon$. It follows that, with probability $\geq \varepsilon - o(1)$ as $m \to \infty$, there exists points $\xi_1^m, \xi_2^m$ of $\Xi(1) \cap ($disc$(\mathbf{0}, \varepsilon^{-m}) \setminus$ disc$(\mathbf{0}, \varepsilon^{1-m}))$ such that

$$\text{routes } \mathcal{R}(\mathbf{0}, \xi_1^m) \text{ and } \mathcal{R}(\mathbf{0}, \xi_2^m) \text{ branch at } \mathbf{0}.$$

So on an event of probability $\geq \varepsilon$ this property holds for infinitely many $m$. Then on that event we have $G > 1$, by compactness within the spanning subnetwork $\mathcal{S}^*(1)$ (Lemma 1). ∎

Next consider the spanning subnetwork $\mathcal{S}^*(\lambda)$ on points $\Xi(\lambda) \cup \{\mathbf{0}\}$. The intersection of all routes $\mathcal{R}(\mathbf{0}, \xi)$, $\xi \in \Xi(\lambda)$ is a sub-route from $\mathbf{0}$ to some branchpoint $B(\lambda)$. So we can define a property

$$\mathbb{P}(B(1) = \mathbf{0}) = 0. \tag{43}$$

Clearly (43) implies (42); we need to argue the converse.

**(42) implies (43).** Suppose (42). For each $r < \infty$ the intersection of routes $\mathcal{R}(\mathbf{0}, \xi)$, $\xi \in \Xi(1) \cap \mathrm{disc}(\mathbf{0}, r)$ is a subroute $\pi(1, r)$ from $\mathbf{0}$ to some branchpoint $B(1, r)$, and by (42), scaling and the finiteness of $\Xi(1) \cap \mathrm{disc}(\mathbf{0}, r)$ we have

$$\mathbb{P}(B(1, r) = \mathbf{0}) = 0, \text{ each } r < \infty. \tag{44}$$

As $r$ increases the subroute $\pi(1, r)$ can only shrink, and the quantity in (43) is the limit $B(1) = \lim_{r \to \infty} B(1, r)$. To prove (43) it suffices, by (44), to prove

$$B(1, r) \text{ is constant for all large } r, \text{ a.s.} \tag{45}$$

We may suppose (otherwise the result is obvious) that for some $r_0 \geq 4$ the subroute $\pi(1, r_0)$ stays within $\mathrm{disc}(\mathbf{0}, 1)$. As $r$ increases, the only way that $B(1, r)$ can change at $r$ is if there is a point $\xi \in \Xi(1) \cap \mathrm{circle}(\mathbf{0}, r)$ for which the route $\mathcal{R}(\mathbf{0}, \xi)$ diverges from the existing subroute $\pi(1, r-)$ before the existing branchpoint $B(1, r-)$. If this happens, at $r_1$ say, then consider the subroute $\theta(r_1) = \mathcal{R}(\mathbf{0}, \xi) \cap (\mathrm{disc}(\mathbf{0}, 4) \setminus \mathrm{disc}(\mathbf{0}, 1))$ which has length at least 3. Now suppose $B(1, r)$ again changes at some larger value $r_2$. Then the corresponding subroute $\theta(r_2)$ must be disjoint from $\theta(r_1)$, by the route-compatability property. Now the "finite length in bounded regions" property (5) implies that $B(1, r)$ can change at only finitely many large values of $r$, establishing (45). ■

Now make a slight re-definition of $B(\lambda)$, by considering only points $\xi$ outside the unit disc. That is, the intersection of all routes $\mathcal{R}(\mathbf{0}, \xi)$, $\xi \in \Xi(\lambda) \setminus \mathrm{disc}(\mathbf{0}, 1)$ is a sub-route $\tilde{\pi}(\lambda)$ from $\mathbf{0}$ to some branchpoint $B_1(\lambda)$. Using scale-invariance it is easy to check that (43) is equivalent to

$$\mathbb{P}(B_1(\lambda) = \mathbf{0}) = 0 \text{ for each } \lambda < \infty. \tag{46}$$

As $\lambda$ increases, the sub-routes $\tilde{\pi}(\lambda)$ can only shrink, and the intersection of these subroutes over all $\lambda < \infty$ is again a subroute from $\mathbf{0}$ to some point $B_1(\infty)$. So we can define a property

$$\mathbb{P}(B_1(\infty) = \mathbf{0}) = 0. \tag{47}$$

Clearly (47) implies (46); we need to argue the converse.

**(46) implies (47).** Suppose (46). To prove (47) we essentially repeat the argument above, but use assumption (20) instead of (5). It is enough to show that, as $\lambda$ increases, $B_1(\lambda)$ can change at only finitely many large values of $\lambda$. And we may suppose that for large $\lambda$ the subroute $\tilde{\pi}(\lambda)$ stays within $\mathrm{disc}(\mathbf{0}, 1/4)$. If $B_1(\lambda)$ changes at $\lambda_1$ then there is a point $\xi$ appearing at "time" $\lambda_1$ for which $\mathcal{R}(\mathbf{0}, \xi)$ diverges from the existing subroute $\tilde{\pi}(\lambda_1-)$ and so must cross $\mathrm{circle}(\mathbf{0}, 5/8)$ at some point $z(\lambda_1) \in \mathcal{E}(\lambda_1, 3/8) \subset \mathcal{E}(\infty, 3/8)$. By route-compatability the points $z(\lambda_i)$ corresponding to different values $\lambda_i$ where $B_1(\lambda)$ changes must be distinct, and then (20) implies $\mathcal{E}(\infty, 3/8) \cap \mathrm{circle}(\mathbf{0}, 5/8)$ is an a.s. finite set of points. ■

Clearly (47) implies unique singly-infinite geodesics, by the final assertion of Corollary 23. We have now shown a cycle of equivalences. Finally, by scaling (47) is equivalent to the following property, where the notation is chosen to be consistent with notation in the next section. Define $Q(\lambda, 0, B)$ to be the probability that the routes $\mathcal{R}(\mathbf{0}, \xi')$ to all points $\xi' \in \mathcal{S}(\lambda) \cap (\mathrm{disc}(\mathbf{0}, B))^c$ do **not** all first exit $\mathrm{disc}(\mathbf{0}, 1)$ at the same point. Then (47) is equivalent to

$$\lim_{B \uparrow \infty} \lim_{\lambda \to \infty} Q(\lambda, 0, B) = 0. \tag{48}$$

To summarize:

**Proposition 24** *Properties (42), (43), (44), (47) and (48) are each equivalent to the unique singly-infinite geodesics property.*

### 7.2 Continuity properties

In the previous section we studied properties of long routes from a single point. We now consider long routes from nearby points, and in this context it seems harder to understand whether different properties are equivalent. Suppose, for this discussion, the unique singly-infinite geodesics property holds. Then the geodesics from $\mathbf{0}$ and from $\mathbf{1} = (1, 0) \in \mathbb{R}^2$ are either disjoint or coalesce; we do not know (Open Problem 32) whether the property

$$\text{the geodesics from } \mathbf{0} \text{ and from } \mathbf{1} \text{ coalesce a.s} \tag{49}$$

always holds or is stronger. There are several equivalent ways of saying (49) – see the end of this section – but what's relevant now is that it is equivalent to the property that, for each $\lambda$, the geodesics from each point of $\mathcal{S}^*(\lambda) \cap \mathrm{disc}(\mathbf{0}, 1)$ coincide outside a disc of random radius $R(\lambda) < \infty$ a.s.. So we can then ask whether the property

$$R(\infty) := \lim_{\lambda \to \infty} R(\lambda) < \infty \text{ a.s.}$$

is implied by property (49) or is stronger. We restate this latter property as (50) below.

For $0 < \varepsilon < 1 < B$ define $Q(\lambda, \varepsilon, B)$ to be the probability that the routes $\mathcal{R}(\xi, \xi')$ between points $\xi \in \mathcal{S}(\lambda) \cap \mathrm{disc}(\mathbf{0}, \varepsilon)$ and points $\xi' \in \mathcal{S}(\lambda) \cap (\mathrm{disc}(\mathbf{0}, B))^c$ do **not** all first exit $\mathrm{disc}(\mathbf{0}, 1)$ at the same point. Note $Q(\lambda, \varepsilon, B)$ is monotone increasing at $\lambda$ increases, and decreasing as $B$ increases or $\varepsilon$ decreases. So we can define

$$Q(\infty, \varepsilon, B) := \lim_{\lambda \to \infty} Q(\lambda, \varepsilon, B)$$

and then define a property of a SIRSN

$$\lim_{\varepsilon \downarrow 0, B \uparrow \infty} Q(\infty, \varepsilon, B) = 0 \tag{50}$$

where the limit value is unaffected by the order of the double limit. In words, (50) says that (with high probability) every route from a small neighborhood of the origin to any distant point will first cross the unit circle at the same place. Property (50) implies (49) and implies form (48) of the unique singly-infinite geodesics property, which is the same assertion for routes from the origin only.

The kinds of properties described above relate to questions about continuity of the routes $\mathcal{R}(z_1, z_2)$ as $z_1, z_2$ vary, and we will give one such relation as Lemma 25 below.

Consider $0 < \eta < \delta < 1/2$ and for points $\xi \in \mathcal{S}(\lambda) \cap \mathrm{disc}(\mathbf{0}, \eta)$ and $\xi' \in \mathcal{S}(\lambda) \cap \mathrm{disc}(\mathbf{1}, \eta)$ with route $\mathcal{R}(\xi, \xi')$ let $\mathcal{R}_\delta(\xi, \xi')$ be the sub-route between the first exit from $\mathrm{disc}(\mathbf{0}, \delta)$ and the last entrance into $\mathrm{disc}(\mathbf{1}, \delta)$. Let $\Psi(\lambda, \eta, \delta)$ be the probability that the sub-routes $\mathcal{R}_\delta(\xi, \xi')$ for all $\xi \in \mathcal{S}(\lambda) \cap \mathrm{disc}(\mathbf{0}, \eta)$ and all $\xi' \in \mathcal{S}(\lambda) \cap \mathrm{disc}(\mathbf{1}, \eta)$ are **not** all an identical sub-route. As above, by monotonicity we can define

$$\Psi(\infty, \eta, \delta) := \lim_{\lambda \to \infty} \Psi(\lambda, \eta, \delta)$$

and then define a property of a SIRSN

$$\lim_{\eta \downarrow 0} \Psi(\infty, \eta, \delta) = 0 \quad \forall \delta. \tag{51}$$

In words, (51) says that (with high probability) all routes from a very small neighborhood of the origin to a very small neighborhood of $\mathbf{1}$ coincide outside of larger small neighborhoods.

**Lemma 25** *Property (50) implies property (51).*

**Proof.** Choose $a$ such that $a\eta < 1 < a\delta$. Take the definition of $\Psi$, scale by $a$, and use scale-invariance to obtain the following.

> The probability that the sub-routes $\mathcal{R}_{a\delta}(\xi, \xi')$ for all $\xi \in \mathcal{S}(a^{-2}\lambda) \cap \mathrm{disc}(\mathbf{0}, a\eta)$ and all $\xi' \in \mathcal{S}(a^{-2}\lambda) \cap \mathrm{disc}((a,0), a\eta)$ are **not** all an identical sub-route equals $\Psi(\lambda, \eta, \delta)$.

When this occurs there are two non-identical sub-routes between circle$(\mathbf{0}, a\delta)$ and circle$((a,0), a\delta)$, which imply two non-identical sub-routes between circle$(\mathbf{0}, 1)$ and circle$((a,0), 1)$. For this to happen, either the defining event for $Q(a^{-2}\lambda, a\eta, a/2)$, or the analogous event with reference to $(a,0)$ instead of $\mathbf{0}$, must occur; otherwise all routes in question pass through the same points on circle$(\mathbf{0}, 1)$ and circle$((a,0), 1)$, contradicting the route-compatability properties of section 2.2. So

$$\Psi(\lambda, \eta, \delta) \leq 2Q(a^{-2}\lambda, a\eta, a/2).$$

Letting $\lambda \to \infty$

$$\Psi(\infty, \eta, \delta) \leq 2Q(\infty, a\eta, a/2).$$

Choosing $a = \eta^{-1/2}$ establishes the lemma. ■

**Remark.** Lemma 25 is almost enough to prove that, under condition (50), we have the continuity property

$$\text{if } (z_1^n, z_2^n) \to (z_1, z_2) \text{ then } \mathcal{R}(z_1^n, z_2^n) \to \mathcal{R}(z_1, z_2) \text{ a.s.} \tag{52}$$

where convergence of paths is in the sense of section 2.3. To deduce (52) one would need also to show that the lengths of $\mathcal{R}(z_1^n, z_2^n) \cap (\mathrm{disc}(z_1, \varepsilon_n) \cup \mathrm{disc}(z_2, \varepsilon_n))$ tend to 0 a.s. for all $\varepsilon_n \to 0$. This is loosely related to Open Problem 34.

**Another property equivalent to (49).** Because geodesics either coalesce or are disjoint, for any countable set of initial points there is some set of "geodesic ends", where each such "end" corresponds to a tree of coalescing geodesics from originating "leaves". By a small modification of the proof of Corollary 23, the mean number of such ends from the points $\Xi(\lambda) \cap \mathrm{disc}(\mathbf{0}, 1)$ is at most $4p(1)$, so we can let $\lambda \to \infty$ and deduce that the number $G^* \geq 1$ of ends from initial points $\Xi(\infty) \cap \mathrm{disc}(\mathbf{0}, 1)$ satisfies $\mathbb{E}G^* \leq 4p(1)$. Then by scale-invariance, for each $0 < r < \infty$ the number of ends from initial points $\Xi(\infty) \cap \mathrm{disc}(\mathbf{0}, r)$ equals $G^*$. So the property

$$G^* = 1 \text{ a.s.}$$

is clearly equivalent to property (49) (plus the unique singly-infinite geodesics property). Note that if $G^* > 1$ then there are a finite number of different "geodesic trees" each of whose leaf-sets is dense in $\mathbb{R}^2$ – behavior hard to visualize.

### 7.3 The binary hierarchy model

**Proposition 26** *The binary hierarchy model has property (50) .*

**Proof.** Consider the last stages of construction of the model in section 3.7. Rotation and scaling do not affect the property of interest, so it will suffice to prove the property in the model $\mathcal{R}_{\text{t-i}}$ . Consider the argument from "proof of Proposition 9" in section 3.6 but with large rescalings of $B$ instead of small rescalings. Combining this argument with the construction of $\mathcal{R}_{\text{t-i}}$ at the start of section 3.7 one can show (details omitted) that the set

$$A' := \{z \in \mathbb{R}^2 : \ z \text{ in only finitely many } B'_i\}$$

has area zero; here $B'_i := \sigma_{2^i} B$ is the "large" rescaling of the union $B := \cup_G \Sigma_G$ of the translates $\Sigma_G$ of the small subsquare $\Sigma$ of the basic $2^{h+1} \times 2^{h+1}$ square $G$ in the Figure 6 configuration. By translation-invariance, this implies that a.s. $\mathbf{0} \notin A'$. For such a realization there is a random infinite sequence $i(j)$ with $\mathbf{0} \in \sigma_{2^{i(j)}} B$, and any singly-infinite geodesic from $\mathbf{0}$ must pass through the corresponding infinite sequence $b_{i(j)}$ of points determined by Figure 6. This establishes the unique singly-infinite geodesic property. Moreover $\mathbf{0}$ lies in some translated square $\Sigma_{i(j)}$ of side $\varepsilon 2^{i(j)}$ and for any other point in that square its geodesic must coalesce with the geodesic from $\mathbf{0}$ at or before $b_{i(j)}$. It is easy to check that the squares $\Sigma_{i(j)}$ eventually cover any fixed disc, and this establishes property (50). ■

## 8 Open problems and final discussion

### 8.1 Other specific models?

A major challenge is finding other explicit examples of SIRSN models. Let us pose the vague problems

**Open Problem 27** *Give a construction of a SIRSN which is "mathematically natural" in some sense, e.g. in the sense that there is an explicit formula for the distribution of subnetworks* $\mathbf{span}(z_1, \ldots, z_k)$*.*

**Open Problem 28** *Give a construction of a SIRSN which is "visually realistic" in the sense of not looking very different from a real-world road network.*

## 8.2 Quantitative bounds on statistics

In designing a finite road network there is an obvious tradeoff between total length and the network's effectiveness in providing short routes, so in our context there is a tradeoff between $\ell$ and $\Delta := \mathbb{E}D_1$. More generally

**Open Problem 29** *What can we say about the set of possible values, over all SIRSNs, of the triple* $(\Delta = \mathbb{E}D_1, \ell, p(1))$ *of statistics of a SIRSN?*

This is a sensible question because each statistic is dimensionless, that is not dependent on choice of unit of length – a non-dimensionless statistic would take all values in $(0, \infty)$ by scaling.

We have given three results relating to this problem. Proposition 16 gave a crude lower bound on the function $\ell_*(\Delta)$ defined as the infimum value of $\ell$ over all SIRSNs with the given value of $\Delta$.

**Open Problem 30** *(i) Give quantitative estimates of the function* $\ell_*(\Delta)$*, improving Proposition 16.*
*(ii) Do "optimal" networks attaining the infimum exist, and (if so) can we say something about the structure of the associated optimal networks?*

One might make the (vague) conjecture that for some value of $\Delta$ the optimal network exploits 4-fold symmetry in some way analogous to our section 3 model, and that for some other value it exploits 6-fold symmetry.

In section 5.5 we showed (35) that the overall minimum normalized length $\ell_* := \inf_\Delta \ell_*(\Delta)$ satisfies $\ell_* \geq \sqrt{1/8}$. The third result was Lemma 20, showng $\ell \leq 2p(1)$.

## 8.3 Traffic intensity

As mentioned in section 2.3, the conceptual point of $\mathcal{E}(\infty, r)$ is to capture the idea of the major road - minor road spectrum, and the particular definition of $\mathcal{E}(\infty, r)$ is mathematically convenient because of the scaling property (37) of the edge-intensity $p(r)$. But from a real-world perspective it seems more natural to use some notion of traffic intensity. Given any measure $\psi$ on source-destination pairs $(z_1, z_2)$, then length measure $\mathrm{Leb}_1$ along the routes $\mathcal{R}(z_1, z_2)$ in a SIRSN induces a "traffic intensity" measure $\widetilde{\psi}$ on $\cup_r \mathcal{E}(\infty, r)$. The natural measures $\psi$ to consider are specified by

(i) $z_1$ has Lebesgue measure $\mathrm{Leb}_2$ on $\mathbb{R}^2$
(ii) given $z_1$, the measure on $z := z_2 - z_1$ has density $|z|^{-\beta}$.
The action of $\sigma_c$ on $\psi$ gives the measure specified by
(i) $z_1$ has measure $c^{-2}\mathrm{Leb}_2$ on $\mathbb{R}^2$
(ii) $z := z_2 - z_1$ has density $c^{\beta-2}|z|^{-\beta}$
(iii) the measure along paths is $c^{-1}\mathrm{Leb}_1$
and this is the measure $\widetilde{\psi}$ scaled by $c^{\beta-5}$.

To make a rigorous treatment, the issue is to show that $\widetilde{\psi}$ is a *locally finite* measure on $\mathcal{E}(\infty,1)$. Heuristically one needs $\beta > 2$ so that the contribution from large $|z_2 - z_1|$ is finite, and $\beta < 4$ so that the contribution from small $|z_2 - z_1|$ is finite.

**Open Problem 31** *Show that, perhaps under regularity assumptions on the SIRSN, for $2 < \beta < 4$ the construction above gives a locally finite measure $\widetilde{\psi}$ on $\mathcal{E}(\infty,1)$ and hence on $\cup_r\mathcal{E}(\infty,r)$.*

## 8.4 Technical questions raised by results

### 8.4.1 Implications between different properties of a SIRSN

We have given various results of the form "one property of a SIRSN implies another" for which we conjecture the converse is false. In particular, we expect there are counter-examples to most of the following, though of course this requires constructing other examples of SIRSNs.

**Open Problem 32** *Prove, or give a counter-example to:*
*(i) (16) implies (20)*
*(ii) the unique singly-infinite geodesics property implies (49)*
*(iii) (49) implies (50)*
*(iv) (51) implies (50).*

### 8.4.2 Understanding the structure of $\mathcal{E}(\infty,1)$.

We envisage $\mathcal{E}(\infty,1)$ as looking somewhat like a real-world network of major roads, but it is not clear what aspects of real networks appear automatically in our SIRSN model. For instance, a priori $\mathcal{E}(\infty,1)$ need not be connected (it might contain a short segment in the middle of a route between two points at distance $2+\varepsilon$ apart) but it must contain an unbounded connected component (most of a singly-infinite geodesic).

**Open Problem 33** *Does $\mathcal{E}(\infty,1)$ have a.s. only a single unbounded connected component?*

### 8.4.3 Questions about lengths

Even though we started the whole topic of SIRSNs by considering route-lengths, they have played a rather small role in our results, and many questions about route-lengths could be asked.

**Open Problem 34** *Under what extra assumptions (if any) is it true that, for* $U_1, U_2, \ldots$ *independent uniform on* $\mathrm{disc}(\mathbf{0}, 1)$*,*

$$\mathbb{E} \sup_{i \geq 1} \mathrm{len}[\mathcal{R}(\mathbf{0}, U_i)] < \infty?$$

The following (intuitively obvious) claim seems curiously hard to prove; the difficulty lies in showing that the spanning subnetwork does not have (necessarily with low probability) huge length a long way away from the square.

**Open Problem 35** *Take* $k$ *uniform random points* $Z_1, \ldots, Z_k$ *in a square of area* $k$ *and consider the length* $\mathrm{len}[\mathbf{span}(Z_1, \ldots, Z_k)]$ *of the spanning subnetwork random network* $\mathbf{span}(Z_1, \ldots, Z_k)$*. Prove*

$$\mathbb{E}\ \mathrm{len}[\mathbf{span}(Z_1, \ldots, Z_k)] \sim \ell k \text{ as } k \to \infty.$$

## 8.5 Alternative starting points for a setup

We started the whole modeling process by assuming we are given routes between points, but one can imagine two different starting points. The first involves starting with a network of major roads and then adding successively more minor roads, so eventually the road network is dense in the plane. In other words, base a model on some *explicit* construction as $r$ decreases of some process $(\mathcal{E}(r),\ \infty > r > 0)$ of "roads of size $\geq r$" (in our setup this is achieved implicitly by the networks $\mathcal{E}(\infty, r)$). Of course this corresponds to what we see when zooming in on an online map of the real-world road network; the maps are designed to show only the relatively major roads within the window, and hence to show progressively more minor roads as one zooms in. In talks we show such zooms along with the online "zooming in" demonstration [24] of *Brownian scaling* to illustrate the concept of scale-invariance.

The second, mathematically abstract, approach is to start with a random metric $d(z, z')$ on the plane, and define routes as geodesics.

But a technical difficulty with both of these approaches is that there seems no simple way to guarantee *unique* routes between a.a. pairs of points in the plane – in general one needs to add an *assumption* of uniqueness. The explicit models constructed in section 3 and outlined in section 4 do use the "random metric" idea, but the hard part of the construction is proving the uniqueness of routes, even in these simplest models we can imagine. It is perhaps remarkable that our approach, taking routes as given with only the route-compatability property but with no explicit requirement that routes be minimum-cost in some sense, does lead to some non-obvious results.

### 8.6 Empirical evidence of scale-invariance?

For real-world road networks, can scale-invariance be even roughly true over some range of distance? We mentioned one explicit piece of evidence (ordered segment lengths) in section 1.5; there is also evidence that mean route length is indeed roughly proportional to distance, though this is also consistent with other (non scale-invariant) models [6].

An interesting project would be to study the spanning subnetworks on (say) 4 real-world addresses, whose positions form roughly a square, randomly positioned, and find the empirical frequencies with which the various topologically different networks appear. Scale-invariance predicts these frequencies should not vary with the side-length of square; is this true?

### 8.7 Other related literature

#### 8.7.1 Hop count in spatial networks

There has been study of spatial networks with respect to the trade-off between total network length and average *graph distance* (hop count), instead of route-length. See [23] for a recent literature survey and empirical analysis.

#### 8.7.2 Continuum random trees in the plane

Existence of continuum limits of discrete models of random trees has been conjectured, and studied non-rigorously in statistical physics, for a long time, and since 2000 spectacular progress has been made on rigorous proofs. For three models of random trees (uniform random spanning tree on $\mathbb{Z}^2$, minimal spanning tree on $\mathbb{Z}^2$ (with random edge lengths), and the Euclidean minimal spanning tree on Poisson points), [2] established a rigorous "tightness" result and gave sample properties of subsequential limits. A subsequent deep result [18] established the existence of a continuum limit in the first model. In these limits the paths have Hausdorff dimension greater than 1 so $D_1 = \infty$ a.s.. There should be a simple proof of the following, because our definition of SIRSN requires $\mathbb{E}D_1 = \infty$.

**Open Problem 36** *In a SIRSN, the subnetwork $\mathcal{S}(1)$ cannot be a tree (with Steiner points).*

### 8.7.3 Geodesics in first-passage percolation

Geodesics in particular models of first-passage percolation have been studied in [19]. It is unclear whether there is any substantial connection between the behavior of geodesis in that setting and in our setting.

### 8.7.4 A Monge-Kantorovitch approach

A completely different approach to continuum networks, starting from Monge-Kantorovitch optimal transport theory, is developed in the monograph by Buttazzo et al. [11]. Their model assumes
(i) some continuous distribution of sources and sinks
(ii) an *a priori* arbitrary set $\Sigma$ representing location of roads
(iii) two different costs-per-unit-length for travel inside [resp. outside] $\Sigma$.
An optimal network in one that minimizes total transportation cost for a given cost functional on $\Sigma$. It is shown that, under regularity conditions, the optimal network is covered by a finite number of Lipschitz curves of uniformly bounded length, although it may have even uncountably many connected components. But this theory does not seem to address statistics analogous to our $\Delta$ and $\ell$ in any quantitative way.

### 8.7.5 The method of exchangeable substructures

The general methodology of studying complicated random structures by studying induced substructures on random points has many applications [7]. In particular, the *Brownian continuum random tree* [3] provides an analogy for what we would like to see (Open Problem 27) in some "mathematically natural" SIRSN – see e.g. the formula (13) therein for the distribution of the induced subtree on random points – though that is in the "mean-field" setting without any $d$-dimensional geometry.

### 8.7.6 Urban road networks.

There is scattered literature on models for *urban* road networks, mostly with a rather different focus, though [17] has some conceptual similarities with our work.

### 8.7.7 Dynamic random graphs.

Conceptually, what we are doing with routes $\mathcal{R}(z_1, z_2)$ and subnetworks $\mathcal{S}(\lambda)$ is *exploring* a given network. This is conceptually distinct from using sequential *constructions* of a network, a topic often called *dynamic random graphs* [13], even though the particular "dynamic Gabriel" model outlined in section 8.1 does fit the "dynamic" category.

**Acknowldgements** My thanks to Justin Salez for details of the proof of Lemma 11, to Wilfrid Kendall for ongoing collaboration, and to Cliff Stein for references to the algorithmic literature.

## References

[1] I. Abraham, A.Fiat, A.V. Goldberg, and R. F. Werneck. Highway dimension, shortest paths, and provably efficient algorithms. In *Proc. ACM-SIAM Symposium on Discrete Algorithms (SODA10)*. SIAM, 2010.

[2] M. Aizenman, A. Burchard, C. M. Newman, and D. B. Wilson. Scaling limits for minimal and random spanning trees in two dimensions. *Random Structures Algorithms*, 15(3-4):319–367, 1999.

[3] D. J. Aldous. The continuum random tree. II. An overview. In *Stochastic analysis (Durham, 1990)*, volume 167 of *London Math. Soc. Lecture Note Ser.*, pages 23–70. Cambridge Univ. Press, Cambridge, 1991.

[4] D. J. Aldous and M. Krikun. Percolating paths through random points. *ALEA Lat. Am. J. Probab. Math. Stat.*, 1:89–109, 2006.

[5] D. J. Aldous and J. Shun. Connected spatial networks over random points and a route-length statistic. *Statist. Sci.*, 25(3):275–288, 2010.

[6] D.J. Aldous. The shape theorem for route-lengths in connected spatial networks on random points. arXiv:0911.5301v1, 2009.

[7] D.J. Aldous. More uses of exchangeability: Representations of complex random structures. In N. H. Bingham and C. M. Goldie, editors, *Probability and Mathematical Genetics: Papers in Honour of Sir John Kingman*. Cambridge University Press, 2010.

[8] D.J. Aldous and W.S. Kendall. Short-length routes in low-cost networks *via* Poisson line patterns. *Advances Appl. Probab.*, 40:1–21, 2008.

[9] H. Bast, S. Funke, P. Sanders, and D. Schultes. Fast routing in road networks with transit nodes. *Science*, 316:566, 2007.

[10] I. Benjamini. Random planar metrics. In *Proceedings of the International Congress of Mathematicians. Volume IV*, pages 2177–2187, New Delhi, 2010. Hindustan Book Agency.

[11] G. Buttazzo, A. Pratelli, S. Solimini, and E. Stepanov. *Optimal Urban Networks via Mass Transportation*, volume 1961 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009.

[12] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Vol. II*. Probability and its Applications (New York). Springer, New York, second edition, 2008. General theory and structure.

[13] R. Durrett. *Random Graph Dynamics*. Cambridge University Press, Cambridge, 2007.

[14] R. J. Gutman. Reach-based routing: A new approach to shortest path algorithms optimized for road networks. In *Proceedings of the Sixth Workshop on Algorithm Engineering and Experiments and the First Workshop on Analytic Algorithmics and Combinatorics, New Orleans, LA, USA, January 10, 2004*, pages 100–111. SIAM, 2004.

[15] J.W. Jaromczyk and G.T. Toussaint. Relative neighborhood graphs and their relatives. *Proceedings of the IEEE*, 80:1502–1517, 1992.

[16] V. Kalapala, V. Sanwalani, A. Clauset, and C. Moore. Scale invariance in road networks. *Phys. Rev. E*, 73(2):026130, Feb 2006.

[17] S. Lämmer, B. Gehlsen, and D. Helbing. Scaling laws in the spatial structure of urban road networks. *Physica A*, 363:89–95, 2006.

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

[19] C. Licea and C. M. Newman. Geodesics in two-dimensional first-passage percolation. *Ann. Probab.*, 24(1):399–410, 1996.

[20] M. Penrose. *Random Geometric Graphs*. Oxford Univ. Press, 2003.

[21] J.M. Steele. *Probability Theory and Combinatorial Optimization*. Number 69 in CBMS-NSF Regional Conference Series in Applied Math. SIAM, 1997.

[22] D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1987. With a foreword by D. G. Kendall.

[23] J. Wang and G. Provan. Topological analysis of specific spatial complex networks. *Adv. Complex Syst.*, 12(1):45–71, 2009.

[24] Wikipedia. A demonstration of Brownian scaling, 2010. http://en.wikipedia.org/wiki/File:Wiener_process_animated.gif [Online; accessed 14-May-2010].

[25] J.E. Yukich. *Probability Theory of Classical Euclidean Optimization Problems*. Number 1675 in Lecture Notes in Math. Springer, 1998.

# A Appendix: A topology on the space of feasible subnetworks.

We first define convergence of routes. Recall a feasible route $\mathsf{r}(z, z')$ involves line segments between points $(z_i)$ which we will call *turn points* of the route. Given $\varepsilon < |z' - z|/2$ there is (starting from $z$) a last turn point $z_{(\varepsilon)}$ before the route $\mathsf{r}(z, z')$ first exits $\mathrm{disc}(z, \varepsilon)$ and there is (starting from $z'$) a last turn point $z'_{(\varepsilon)}$ before the reverse route $\mathsf{r}(z', z)$ first exits $\mathrm{disc}(z', \varepsilon)$. Define

$$\mathsf{r}(z(n), z'(n)) \to \mathsf{r}(z, z')$$

to mean
(i) $z(n) \to z,\ z'(n) \to z' \neq z$.
(ii) For each $\varepsilon < |z' - z|/2$ such that $\mathrm{circle}(z, \varepsilon)$ and $\mathrm{circle}(z', \varepsilon)$ do not contain any turn point of $\mathsf{r}(z, z')$, writing the turn points of the subroutes $\mathsf{r}(z_{(\varepsilon)}(n), z'_{(\varepsilon)}(n))$ as $(y_0(n), y_1(n), \ldots, y_k(n))$, we have

$$(y_0(n), y_1(n), \ldots, y_k(n)) \to (y_0, y_1, \ldots, y_k)$$

the limit being the turn points of the subroute $\mathsf{r}(z_{(\varepsilon)}, z'_{(\varepsilon)})$, where $k$ is finite from the definition of feasible route.
(iii) The total lengths $L_{(\varepsilon)}(n)$ of $\mathsf{r}(z(n), z'(n)) \cap (\mathrm{disc}(z, \varepsilon) \cup \mathrm{disc}(z', \varepsilon))$ satisfy

$$\lim_{\varepsilon \to 0} \limsup_n L_{(\varepsilon)}(n) = 0.$$

Despite its inelegant formulation, this seems the "natural" notion of convergence.

Now we specify, in a way analogous to (i-iii) above, what it means for a sequence $\mathsf{s}(n)$ of feasible subnetworks on locally finite sets $\mathbf{z}(n) = \{z^i(n)\}$ to converge to a limit subnetwork $\mathsf{s}$ on $\mathbf{z}$.

(i) We need $\mathbf{z}(n)$ to converge to $\mathbf{z}$ in the usual sense of convergence of simple point processes [12]. This is equivalent to saying that if we take any $R$ such that circle$(\mathbf{0}, R)$ contains no point of $\mathbf{z}$, then we can label the points of $\mathbf{z}(n) \cap$ disc$(\mathbf{0}, R)$ as $(z^1(n), \ldots, z^K(n))$ in such a way that $(z^1(n), \ldots, z^K(n)) \to (z^1, \ldots, z^K)$ , the limit (here and in analogous assertions below) being the points of $\mathbf{z} \cap \text{disc}(\mathbf{0}, R)$.

(ii) Take $R$ and $(z^1, \ldots, z^K)$ as above and take $\varepsilon < \frac{1}{2} \min_{1 \leq i < j \leq K} |z^i - z^j|$ such that $\cup_{1 \leq i \leq K} \text{circle}(z^i, \varepsilon)$ does not contain any turn point within $\mathsf{s}$. Then we can label the turn points of $\cup_{1 \leq i \leq K} \mathsf{r}(z^i_{(\varepsilon)}(n), z^j_{(\varepsilon)}(n))$ as $(y^u(n), 1 \leq u \leq L)$ in such a way that

$(y^u(n), 1 \leq u \leq L) \to (y^u, 1 \leq u \leq L)$

$(y^u(n), y^v(n))$ is an edge-segment of route $\mathsf{r}(z^i(n), z^j(n))$ iff $(y^u, y^v)$ is an edge-segment of route $\mathsf{r}(z^i, z^j)$.

(iii) For each $1 \leq i < j \leq K$ the routes $\mathsf{r}(z^i(n), z^j(n))$ satisfy (iii) above.

(iv) Lemma 1 (i) implies that given $R$, the following quantity (referring to the subnetwork $\mathsf{s}$) is finite:

$$R^* := \min\{r : \bigcup_{z_i, z_j \in \text{disc}(\mathbf{0}, r)} \mathsf{r}(z_i, z_j) \cap \text{disc}(\mathbf{0}, R) = \mathsf{s} \cap \text{disc}(\mathbf{0}, R) \ \}$$

(that is, each edge of $\mathsf{s}$ within disc$(\mathbf{0}, R)$ is part of some route between endpoints in disc$(\mathbf{0}, R^*)$). We require

$$\limsup_n R^*(n) < \infty \text{ for each } R < \infty.$$

We have described sequential convergence within the space of feasible subnetworks. It is routine to show this is convergence in some complete separable metric space, but we won't pursue such theory here.