# Optimal spatial transportation networks where link costs are sublinear in link capacity

## D J Aldous

Department of Statistics, University of California Berkeley, 367 Evans Hall
# 3860, Berkeley, CA 94720, USA
E-mail: aldous@stat.berkeley.edu
URL: www.stat.berkeley.edu/users/aldous

**Abstract.** Consider designing a transportation network on $n$ vertices in the plane, with traffic demand uniform over all source–destination pairs. Suppose the cost of a link of length $\ell$ and capacity $c$ scales as $\ell c^\beta$ for fixed $0 < \beta < 1$. Under appropriate standardization, the cost of the minimum cost *Gilbert network* grows essentially as $n^{\alpha(\beta)}$, where $\alpha(\beta) = 1 - (\beta/2)$ on $0 < \beta \leq \frac{1}{2}$ and $\alpha(\beta) = \frac{1}{2} + (\beta/2)$ on $\frac{1}{2} \leq \beta < 1$. This quantity is an upper bound in the worst case (of vertex positions) and a lower bound under mild regularity assumptions. Essentially the same bounds hold if we constrain the network to be *efficient* in the sense that average route length is only $1 + o(1)$ times the average straight line length. The transition at $\beta = \frac{1}{2}$ corresponds to the dominant cost contribution changing from short links to long links. The upper bounds arise in the following type of hierarchical networks, which are therefore optimal in an order-of-magnitude sense. On the large scale, we use a sparse Poisson line process to provide long-range links. On the medium scale, we use hierarchical routing on the square lattice. On the small scale, we link vertices directly to medium-grid points. We discuss one of the many possible variant models, in which links also have a designed maximum speed $s$ and the cost becomes $\ell c^\beta s^\gamma$.

**Keywords:** random graphs, networks, traffic models, communication, supply and information networks

## Contents

## 1. Introduction

To design a transportation network linking specified points (visualized as cities) in the plane, one might specify a cost functional and a benefit functional on all possible networks, and then consider networks which are optimal in the sense of minimizing cost for a given level of benefit. This paper addresses one particular choice of functionals, but our broader purpose (see section 1.1) is to draw the attention of statistical physicists to this class of problem.

We study a simple model involving the 'economy of scale' idea:

One link of length $\ell$ and capacity $2c$ is less than twice as expensive as two links of length $\ell$ and capacity $c$.

We capture this idea by specifying that the cost of a link of length $\ell$ and capacity $c$ scales as $\ell c^\beta$ for some $0 < \beta < 1$. In the real world, network designers do not know in advance what traffic demand will be. We simplify by assuming that traffic demand is known (and uniform over all source–destination pairs) and routes are controlled, so that the volume $f(e)$ of flow across an edge (link) $e$ can be determined by the designers and the corresponding link-capacity built. (Visualize links as roads and flow volume $f(e)$ as 'number of vehicles per hour'. We are ignoring stochastic fluctuations in traffic.) Thus our cost structure is

$$\text{cost of network} = \sum_e \ell(e) f^{\beta(e)} \tag{1}$$

where $\ell(e) =$ length of link $e$.

To define the model carefully, write $\mathbf{x}^n = \{x_1, x_2, \ldots, x_n\}$ for a configuration of $n$ vertices in the square $[0, n^{1/2}]^2$ of area $n$. So $x_i$ is the position of vertex $i$. Create a connected *network* $G(\mathbf{x}^n)$ by adding links: links are line segments with their natural Euclidean lengths, and links may meet at places not in the given vertex set $\mathbf{x}^n$. To make the distinction clear let us refer to the given $n$ vertices as *cities* and any meeting places (which depend on our choice of network) as *junctions*. Between each source–destination pair $(i, j)$ of cities, a flow of volume $n^{-3/2}$ (this scaling is explained below) is routed through the network. Define $\text{cost}(G(\mathbf{x}^n))$, the cost of the network, via (1). This setting specializes a setting considered by Gilbert [1] and we call the minimum-cost network the *Gilbert network* $\text{Gil}(\mathbf{x}^n)$. See [2] for general properties of, and heuristic algorithms for, Gilbert networks over deterministic points.

Gilbert networks may be optimal from a network *operator* viewpoint, but what about a network *user*? Write $\ell(x_i, x_j)$ for route length and $|x_j - x_i|$ for straight line distance between cities $i$ and $j$. For a typical configuration, the average distance $\text{ave}_{i,j}|x_j - x_i|$ will be of the order of $n^{1/2}$. The kind of 'benefit to users' we have in mind is that the network provides routes almost as short as possible. So we call the sequence of networks $(G(\mathbf{x}^n))$ *modestly efficient* if

$$\text{ave}_{i,j}(\ell(x_i, x_j) - |x_j - x_i|) = \text{o}(n^{1/2}). \tag{2}$$

The name reflects the remarkable fact [3] that there exist *extremely efficient* networks for which this average is $\text{O}(\log n)$ while their length is only $1 + \text{o}(1)$ times the minimum length of any connected network; such results pay no attention to flow volumes or capacities, and so constitute the $\beta = 0$ case of the present model. The problem we address in this paper is:

> given the sequence $(\mathbf{x}^n)$, how small can we make $\text{cost}(G(\mathbf{x}^n))$ subject to the *modestly efficient* constraint (2)?

In the $\beta = 0$ case just mentioned, we can make $\text{cost}(G(\mathbf{x}^n))$ be asymptotically the length of the Steiner tree (minimum length connected network) on $\mathbf{x}^n$, which is well known to be $\text{O}(n)$ in the worst case and in the typical case. Recall that $a_n = \text{O}(b_n)$ means that $a_n/b_n$ is bounded as $n \to \infty$. It is often convenient to write the converse relationship $b_n = \text{O}(a_n)$ as $a_n = \Omega(b_n)$; if both $a_n = \text{O}(b_n)$ and $a_n = \Omega(b_n)$ then we write $a_n = \Theta(b_n)$.

In the case $\beta = 1$ there is no 'economy of scale' and so the minimum-cost network is just the complete graph, that is a direct link between each pair of cities. The associated cost is

$$\sum_i \sum_j n^{-3/2} |x_i - x_j| = n \times \frac{\text{ave}_{i,j}|x_i - x_j|}{n^{1/2}}$$

which is $\text{O}(n)$ in the worst case and in the typical case.

Recall that the Gilbert network $\text{Gil}(\mathbf{x}^n)$ is the minimum-cost network when there is no extra 'modestly efficient' constraint. Theorem 1 shows that imposing the 'modestly efficient' constraint makes little difference in an order-of-magnitude sense: in either case the optimal cost grows roughly as the order of $n^{\alpha(\beta)}$.

**Theorem 1.** *Fix* $0 < \beta < 1$. *Define*

$$\alpha(\beta) = 1 - \frac{\beta}{2}, \qquad 0 < \beta \leq \frac{1}{2}$$
$$= \frac{1 + \beta}{2}, \qquad \frac{1}{2} \leq \beta < 1.$$

*Let* $\mathbf{x}^n$ *be a configuration of* $n$ *cities in the square* $[0, n^{1/2}]^2$.

*(a) Case* $0 < \beta \leq \frac{1}{2}$. *There exist modestly efficient networks for which* $\mathrm{cost}(G(\mathbf{x}^n)) = \mathrm{O}(n^{\alpha(\beta)})$ *(except for* $\beta = 1/2$ *the bound is* $\mathrm{O}(n^{3/4} \log n)$*).* *Under the technical assumption (7) there do not exist connected networks for which* $\mathrm{cost}(G(\mathbf{x}^n)) = \mathrm{o}(n^{\alpha(\beta)})$. *So under (7) we have* $\mathrm{cost}(\mathrm{Gil}(\mathbf{x}^n)) = \Theta(n^{\alpha(\beta)})$ *for* $\beta < \frac{1}{2}$.

*(b) Case* $\frac{1}{2} < \beta < 1$. *Here* $\mathrm{cost}(\mathrm{Gil}(\mathbf{x}^n)) = \mathrm{O}(n^{\alpha(\beta)})$. *Given* $\omega_n \to \infty$ *arbitrarily slowly, there exist modestly efficient networks for which* $\mathrm{cost}(G(\mathbf{x}^n)) = \mathrm{O}(\omega_n n^{\alpha(\beta)})$. *Under the technical assumption (8),* $\mathrm{cost}(\mathrm{Gil}(\mathbf{x}^n)) = \Theta(n^{\alpha(\beta)})$, *but there do not exist modestly efficient networks for which* $\mathrm{cost}(G(\mathbf{x}^n)) = \mathrm{O}(n^{\alpha(\beta)})$.

Our discussion above of the cases $\beta = 0$ and $\beta = 1$ implies corresponding results in these cases with $\alpha(0) = 1$ and $\alpha(1) = 1$.

The transition at $\beta = \frac{1}{2}$ corresponds to the dominant cost contribution changing from short links to long links, as we will explain in section 2.5. The technical regularity assumptions that we need to impose to obtain lower bounds reflect this transition: for $\beta < 1/2$ we need to assume that nearest-neighbor distances are not atypically small, whereas for $\beta > 1/2$ we assume a large-scale equidistribution of the city configuration. (We defer statements of these assumptions until the place they are actually used in the proof, to avoid interrupting the conceptual discussion here.) We show (section 2) that the upper bounds arise in the following type of hierarchical networks, which are therefore optimal in an order-of-magnitude sense. On the large scale, use a sparse Poisson line process to provide long-range links. On the medium scale, use hierarchical routing on the square lattice. On the small scale, link cities directly to medium-grid points. It is perhaps counter-intuitive that one can use the same network for the whole range of $\beta$; the point is that only the medium–small scale structure really matters for $\beta < 1/2$ and only the large-scale structure really matters for $\beta > 1/2$. Our arguments implicitly imply some weak properties of the *exactly* optimal networks. Understanding in detail the structure of the Gilbert network (or the asymptotically optimal modestly efficient network) over random points in the critical case $\beta = 1/2$ is a challenging problem, interesting because one expects the network to have some *scale-free* structure, in the (correct) sense of invariance under spatial and flow-volume rescaling.

One can imagine many variant models in which extra structure is incorporated. In section 4 we briefly discuss the case where links have designed speed $s$ and where the cost of a link becomes $\ell c^\beta s^\gamma$; in this case an analog of theorem 1 remains true.

## 1.1. Optimal spatial network design methodology

This paper contributes to a general program concerning networks linking points in the plane:

> for mathematically simple cost/benefit functionals, study the properties (geometry, cost and benefit values) of optimal networks as the number $n$ of points tends to infinity.

Network design problems arise in many applied fields, but serious real-world modeling leads to more complicated functionals tuned to specific applications than we have in mind. As complementary work, [3] gives a detailed treatment of the *extremely efficient* networks mentioned above that minimize average route length subject to total network length; and [4] analyzes a model (for, e.g., passenger air travel or package delivery) where there is a substantial cost to transfer from one link to another. In the latter model, theory predicts that hub-and-spoke networks (as seen in the real world) are near-optimal and that, constraining the average number of transfers to be say 2, the length of the shortest possible network scales as $n^{13/10}$.

The methodological feature we want to emphasize concerns models for the position of $n$ cities (assumed for simplicity in a square of area $n$). In each problem we have studied one gets the same order of magnitude for optimal network cost for worst-case positions as one gets for arbitrary positions (under mild assumptions) and in particular the same as for random positions or for regular (e.g. lattice) positions.

The bulk of the statistical physics literature on spatial networks (surveyed in [5]) analyzes networks built according to some specific probability model which combines ingredients such as

(a) geometric random graphs (link probability depends on inter-vertex distance);
(b) proportional attachment probabilities for arriving vertices;
(c) prescribed power law distribution of lattice vertex degrees;
(d) networks based on recursive partitioning of space.

This theoretical literature makes passing reference to optimality, but we have not seen analytic results demonstrating optimality over *all* possible networks in the spatial context (see [6] for non-spatial results and [7, 8] for assumptions under which optimal networks are trees). For interesting empirical work see [9].

Our scaling conventions (a square of area $n$; flow-volume $n^{-3/2}$ between each source–destination pair) may seem arbitrary, but are chosen to fit the following standardizations:

 (i) cities have density 1 per unit area;
(ii) flow volume across unit area is of the order of 1.
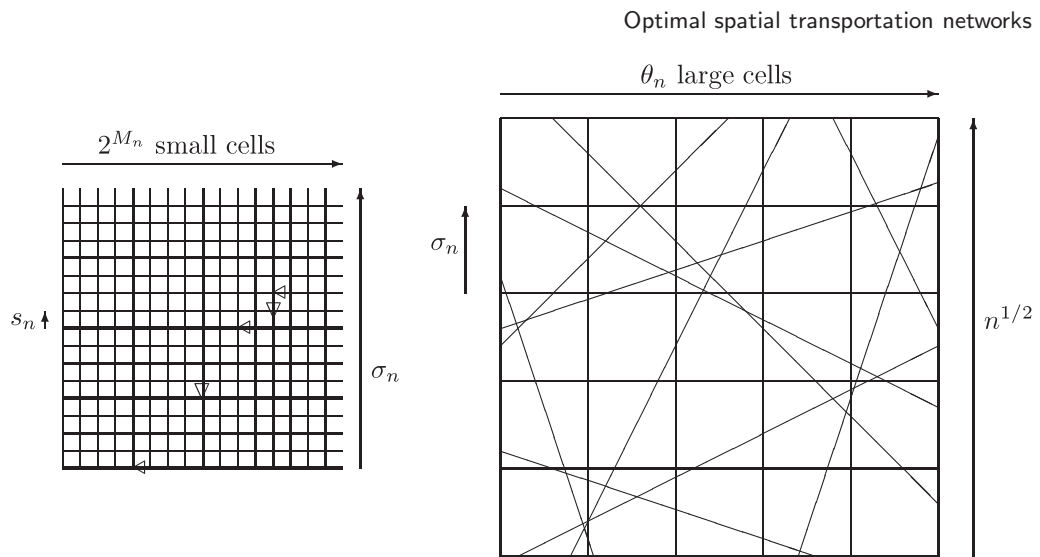
## 2. The construction

A network satisfying the requirements of theorem 1 will be constructed in section 2.3 using mathematical ingredients described in sections 2.1 and 2.2. Figure 1 illustrates the construction.

### 2.1. Hierarchical routing on the square lattice

Fix $M$ and consider the square grid on vertices $\{0, 1, 2, \ldots, 2^M - 1\}^2$. Declare lines (and their edges) to be of some type $0, 1, 2, \ldots, M$ according to the rule:

the horizontal lines $\{(x, y) : y = (2j - 1)2^m\}$, $\qquad j = 1, 2, \ldots$ are type $m$

the boundary line $\{(x, 0)\}$ is type $M$;

**Figure 1.** Ingredients of the construction. Left: the hierarchical routing lattice, with higher type edges indicated by thicker lines, and a typical route shown. Right: the large-scale grid and the Poisson line process.

and similarly for vertical lines. For each vertex $(x, y)$, define a route from $(x, y)$ to $(0, 0)$ using only downward and leftward edges as follows. First choose the edge at $(x, y)$ of higher type (breaking ties arbitrarily). Then repeat the rule

> Follow the current edge until it crosses an edge of strictly higher type, then transfer to that edge

until reaching $(0, 0)$. See figure 1, left side.

It is elementary to verify

**Lemma 2.** *For each $0 \leq m \leq M$, the number of type-m edges traversed by the route is at most $2^{m+1}$.*

### 2.2. The Poisson line process

A line in the plane may be parameterized by the point $z$ on the line which is closest to the origin (so the line segment from the origin to $z$ is orthogonal to the line); then write $z$ in radial coordinates as $(r, \theta)$. Recall [10] the notion of a *Poisson line process* (PLP) of intensity $\eta > 0$, which makes precise the notion of 'completely random' lines in the plane. Parameterizing lines by their closest points $(r, \theta)$, this PLP has intensity $\eta$ with respect to the Lebesgue measure on parameter space $(0, \infty) \times (0, 2\pi)$. The PLP distribution is invariant under Euclidean transformations, and for a fixed set $A$

$$\mathbb{E}(\text{length of line segments intersecting} A) = \pi \eta \times \text{area}(A). \tag{3}$$

(We write $\mathbb{E}$ for expectation and $\mathbb{P}$ for probability). The next result shows how the PLP is useful in constructing spatial networks. See figure 1, right side.

**Lemma 3.** *Let $n^{1/2}/\sigma_n$ be an integer. Construct a network as the superposition of the rectangular grid with cell side length $\sigma_n$ and the Poisson line process of intensity $\eta$, intersected with the square $[0, n^{1/2}]^2$. Let $v_i, v_j$ be vertices of the grid. Then*

$$\mathbb{E}(route\text{-}length\ v_i\ to\ v_j) \leq |v_i - v_j| + C_2\frac{1}{\eta}\log(\eta\sqrt{2n})$$

*for an absolute constant $C_2$.*

Lemma 3 is proved in [3], lemma 11, and we will not repeat the argument here. (In essence, one analyzes the natural routing algorithm: move to a nearby line of the PLP, move along that line in the direction closer to the direction of the destination city, and when encountering another line of the PLP, switch to that line if its direction is closer to the destination city direction). Using the PLP gives us *random* networks, but a typical realization will have costs and lengths of the same order as the expectations in our formulas.

### 2.3. Construction of the networks

We now describe how the ingredients above (hierarchical routing on the square lattice, the PLP) are used in a network construction. Recall $\mathbf{x}^n$ denotes the given configuration of $n$ cities. Take integers $\theta_n \uparrow \infty$ slowly and define

$$\sigma_n = n^{1/2}/\theta_n.$$

Let $M_n$ be the integer such that

$$\sigma_n/2 < 2^{M_n} \leq \sigma_n.$$

Define

$$s_n = \sigma_n/2^{M_n}, \qquad (\text{so} 1 \leq s_n < 2).$$

Construct a network $G(\mathbf{x}^n)$ as follows.

(i) Take the *large-scale* network in lemma 3, with $\eta_n = \theta_n n^{-1/2}$. This network contains *large cells* of side length $\sigma_n$.

(ii) Inside each large cell put a copy of the hierarchical routing lattice of section 2.1, with $M = M_n$, and scaled so that the basic *small cell* of this lattice has side length $s_n$.

(iii) Link each city $x \in \mathbf{x}^n$ via a straight edge to the bottom left corner vertex $v(x)$ of its small cell.

Figure 1 illustrates (i) and (ii). There is a natural way to define a route from $x_i$ to $x_j$ in this network. From $x_i$ take the link to $v(x_i)$, then follow the section 2.1 routing scheme to the lower left corner $V(x_i)$ of the large cell; navigate from $V(x_i)$ to $V(x_j)$ via the shortest route in the lemma 3 graph.

Note that, in addition to the given $n$ cities, this network has several different kinds of junctions: the vertices of the grid and places where lines of the PLP cross each other or cross the grid lines or cross the short stage (iii) links. In our model there is no cost associated with creating a junction or with routes using junctions; the costs involve only link lengths and route lengths. So the exact number of junctions is unimportant.

### 2.4. Analysis of the networks

Clearly

$$\ell(x_i, x_j) \le \ell(V(x_i), V(x_j)) + 2^{3/2}\sigma_n$$

and so by lemma 3

$$\mathbb{E}\ell(x_i, x_j) \le |x_i - x_j| + 2^{3/2}\sigma_n + C_2 \frac{1}{\eta_n} \log(\eta_n \sqrt{2n}).$$

From the definitions of $\sigma_n, \eta_n$ we see

$$\mathbb{E}(\ell(x_i, x_j) - |x_i - x_j|) = \mathrm{o}(n^{1/2})$$

establishing the modestly efficient property.

To analyze costs, we treat stages (i)–(iii) separately and check that each stage cost is less than the bounds stated in theorem 1.

*Stage (iii).* There are $n$ links of the form $(x, v(x))$, each carrying flow-volume $2(1 - (1/n))n^{-1/2}$, and each having length at most $s_n\sqrt{2}$, and so

$$\text{the total cost of stage (iii) links is } \mathrm{O}(n^{1-\beta/2}). \tag{4}$$

*Stage (ii).* Now let $\mathcal{E}_m$ be the set of type-$m$ edges. The number of such edges is $\#\mathcal{E}_m = \mathrm{O}(n2^{-m})$. Recall that Hölder's inequality shows that for any edge set $\mathcal{E}$

$$\sum_{e\in\mathcal{E}} f^\beta(e) \le (\#\mathcal{E})^{1-\beta} \left( \sum_{e\in\mathcal{E}} f(e) \right)^\beta.$$

Now

$$\sum_{e\in\mathcal{E}_m} f(e) = 2n^{-1/2} \sum_{x\in\mathbf{x}^n} \#\{ \text{ type-}m \text{ edges in route } v(x) \text{ to } V(x)\} \le 2^{m+2}n^{1/2}$$

using lemma 2. Thus

$$\sum_{e\in\mathcal{E}_m} f^\beta(e) = \mathrm{O}\left( (n2^{-m})^{1-\beta} \, (2^m n^{1/2})^\beta \right) = \mathrm{O}\left( n^{1-(\beta/2)} 2^{m(2\beta-1)} \right). \tag{5}$$

Writing $\mathcal{E}_{\mathrm{med}}$ for all edges in the copies of the hierarchical routing lattice, we find after summing over $0 \le m \le M$

$$\sum_{e\in\mathcal{E}_{\mathrm{med}}} f^\beta(e) = \mathrm{O}(n^{1-(\beta/2)}), \qquad 0 < \beta < \tfrac{1}{2}$$

$$= \mathrm{O}(n^{3/4}\log n), \qquad \beta = \tfrac{1}{2}$$

$$= \mathrm{O}(n^{1-(\beta/2)}2^{M(2\beta-1)}) = \mathrm{O}(n^{(1/2)+(\beta/2)}), \qquad \tfrac{1}{2} < \beta < 1$$

using $2^M < n^{1/2}$. Because edge lengths here are $s_n < 2$, these are bounds for the costs associated with stage (ii).

*Stage (i).* Write $\mathcal{E}_{\mathrm{large}}$ for the set of links of the large-scale network, that is the large-scale grid and the PLP lines. Flow along the route from $V(x_i)$ to $V(x_j)$ contributes $n^{-3/2}\ell(V(x_i), V(x_j))$ to the 'flow × distance' measure, and so

$$\int_{\mathcal{E}_{\mathrm{large}}} f(e)\,\mathrm{d}e = n^{-3/2} \sum_i \sum_j \ell(V(x_i), V(x_j))$$

where the left side denotes integrating along all links of the large-scale network. By the already-established modestly efficient property,

$$\sum_i \sum_j \ell(V(x_i), V(x_j)) = (1 + o(1)) \sum_i \sum_j |x_i - x_j| = \mathrm{O}(n^{5/2})$$

and so

$$\int_{\mathcal{E}_{\text{large}}} f(e)\, \mathrm{d}e = \mathrm{O}(n).$$

The total length $L_n$ of $\mathcal{E}_{\text{large}}$ is the sum of $\mathrm{O}(n^{1/2}\theta_n)$ (= contribution from large-scale grid) and $\mathrm{O}(\eta_n n)$ (= contribution from the PLP, using (3)), and so $L_n = \mathrm{O}(n^{1/2}\theta_n)$. The integral form of Hölder's inequality now shows that the cost associated with $\mathcal{E}_{\text{large}}$ is

$$\int_{\mathcal{E}_{\text{large}}} f^\beta(e)\, \mathrm{d}e \leq L_n^{1-\beta} \times \left( \int_{\mathcal{E}_{\text{large}}} f(e)\, \mathrm{d}e \right)^\beta = \mathrm{O}\left( \theta_n^{1-\beta} n^{(1+\beta)/2} \right). \tag{6}$$

Examining the cost of each stage, we check that the modestly efficient network we have constructed has its cost bounded as stated in theorem 1. Moreover, if we eliminate the 'modestly efficient' constraint then we can eliminate stage (iii) of the construction (take $\theta_n = 1$) and get the stated $\mathrm{O}(n^{\alpha(\beta)})$ upper bound.

### 2.5. The transition at $\beta = 1/2$

To summarize, the costs associated with the constructed networks arising from short-, medium- and large-scale links are bounded by expressions (4), (5) and (6), respectively. By examining the exponents of $n$ we see that the transition at $\beta = \frac{1}{2}$ corresponds to the dominant cost contribution changing from short links to long links. The arguments we give below for the lower bound show this is a genuine effect (no alternate networks can do essentially better), not an artifact of the particular networks constructed above.

## 3. The lower bound

In the settings of [3, 4] the lower bounds require some effort to prove, but in the present setting the proofs are short.

### 3.1. The case $0 < \beta \leq 1/2$

Consider first the case $0 < \beta \leq 1/2$. Impose the condition: there exists some small $\delta > 0$ such that

for at least $\delta n$ of the cities of $\mathbf{x}^n$, the distance to the nearest neighbor is at least $\delta$. (7)

Consider a city $x \in \mathbf{x}^n$ satisfying this condition and consider the link segments of an arbitrary connected network within distance $\delta/2$ from $x$. Because flow of volume $2n^{-1/2}$ must enter or leave $x$, the cost associated with these link segments (which by concavity of $f \to f^\beta$ is minimized when there is a single link segment) is at least $\delta/2 \times (2n^{-1/2})^\beta$. Summing over all (there are at least $\delta n$) such cities $x$, noting the link segments are distinct as $x$ varies, the network cost is at least $\delta n \times \delta/2 \times (2n^{-1/2})^\beta = \Omega(n^{1-(\beta/2)})$.

## 3.2. The case $1/2 < \beta < 1$

In the case $1/2 < \beta < 1$ we impose the classical equidistribution property for the configuration $\mathbf{x}^n = (x_i^n, 1 \le i \le n)$ rescaled back to the unit square:

$$\text{the empirical distribution of } \{n^{-1/2} x_i^n, 1 \le i \le n\} \text{ converges}$$
$$\text{in distribution to the uniform distribution on } [0,1]^2. \tag{8}$$

Our standardization conventions imply that the total volume of flow through the network is $\Theta(n^{1/2})$ and so assertion (a) below is obvious.

**Lemma 4.** *(a) In the Gilbert network* $\text{Gil}(\mathbf{x}^n)$*, the maximum edge flow is bounded as*

$$\max_e f(e) = \mathrm{O}(n^{1/2}).$$

*(b) For any modestly efficient network* $(G(\mathbf{x}^n))$ *on configurations satisfying the equidistribution condition (8), the maximum edge flow is bounded as*

$$\max_e f(e) = \mathrm{o}(n^{1/2}).$$

Granted this result, use the fact

$$\sum_e \ell(e) f(e) \ge n^{-3/2} \sum_i \sum_j |x_i - x_j| = \Theta(n) \text{ by equidistribution}$$

and the general inequality

$$\text{cost}(G(\mathbf{x}^n)) = \sum_e \ell(e) f^\beta(e) \ge \frac{\sum_e \ell(e) f(e)}{(\max_e f(e))^{1-\beta}}$$

to deduce that $\text{cost}(G(\mathbf{x}^n))$ grows strictly faster than $n/n^{(1-\beta)/2} = n^{\alpha(\beta)}$ for any modestly efficient network, and no slower than order $n^{\alpha(\beta)}$ for the Gilbert network.

**Proof of lemma 4(b).** We first quote an easy fact from geometry.

**Lemma 5.** *Let* $Z_1, Z_2$ *be two independent uniform random points in the unit square* $[0,1]^2$*. There exists a constant* $C$ *such that for all* $x \in [0,1]^2$ *and all* $\delta > 0$

$$\mathbb{P}(|Z_1 - x| + |Z_2 - x| \le |Z_1 - Z_2| + \delta) \le C\delta^{1/2}.$$

Now fix $\delta > 0$. Write $X_1, X_2$ for two uniform random picks from the set $\mathbf{x}^n$ of cities. The modestly efficient assumption implies

$$\mathbb{P}(\ell(X_1, X_2) \ge |X_1 - X_2| + \delta n^{1/2}) \to 0 \qquad \text{as } n \to \infty.$$

Lemma 5 and the equidistribution assumption (8) imply

$$\mathbb{P}(|X_1 - x| + |X_2 - x| \le |X_1 - X_2| + \delta \ \text{ for all } x) \le C\delta^{1/2} + \mathrm{o}(1).$$

In order for the route from $X_1$ to $X_2$ to pass through point $x$, one of the two inequalities above must hold, and so

$$\sup_x \mathbb{P}(\text{route } X_1 \text{ to } X_2 \text{ passes through } x) \le C\delta^{1/2} + \mathrm{o}(1).$$

But $\delta$ is arbitrary, so this probability is $\mathrm{o}(1)$, and the flow volume is exactly $n^{1/2}$ times this probability.

## 4. Associating speeds with links

The main feature of our model—that the cost of building a link is sublinear in link capacity—is just one of many realistic features one might want to incorporate into a model. By focusing on route lengths, we have implicitly assumed that users travel at constant speed. A notable feature of real road or rail networks is that different links permit different speeds. In this section we state and briefly discuss a variant model in which links can be designed to permit different speeds.

Suppose a link with length $\ell$, nominal capacity $c_0$ and nominal speed $s_0$ costs $\ell c_0^\beta s_0^\gamma$, for fixed $0 < \gamma < \infty$. On such a link, traffic moves with speed $s_0$ provided the flow volume $f$ is at most $c_0$; for larger flow volumes, congestion causes the speed to drop, reaching speed 0 (jammed) at volume $\sigma c_0$ for a constant $\sigma$. So $\sigma c_0$ is the maximum capacity. Precisely,

$$\text{speed at flow-volume } f = s_0 G(f/c_0)$$

where $G(u) = 1$ for $0 \le u \le 1$ and $G(u)$ decreases from 1 to 0 as $u$ increases from 1 to $\sigma$. Otherwise the model is the same as before: we are given a configuration of $n$ cities in the square of area $n$ and we are required to route flow of volume $n^{-3/2}$ between each source–destination pair.

For any network and feasible routing, define the average speed as

$$\overline{\text{speed}} = \frac{\text{ave}_{i,j}|x_i - x_j|}{\text{ave}_{i,j}t(x_i, x_j)}$$

where $t(x_i, x_j)$ is the time taken to travel from $x_i$ to $x_j$. For this model, we ask

What is the minimum cost for a network on a given configuration $\mathbf{x}^n$ of cities that allows $\overline{\text{speed}} = s$?

The answer is that, under the regularity assumptions of theorem 1 (which are needed only for lower bounds), and ignoring $\mathrm{O}(\log n)$ terms

minimum cost grows as order $s^\gamma n^{\alpha^*(\beta,\gamma)}$, where

$$
\begin{aligned}
\alpha^*(\beta,\gamma) &= 1 - \frac{\beta}{2} - \frac{\gamma}{2}, & 0 < 2\beta + \gamma \le 1 \\
&= \frac{1+\beta}{2}, & 1 \le 2\beta + \gamma.
\end{aligned}
\tag{9}
$$

Let us briefly indicate how the previous analysis is adapted to this setting. Because costs scale with design speed $s_0$ as $s_0^\gamma$, it is enough to consider the case $\overline{\text{speed}} = 1$ and show that minimum cost grows as order $n^{\alpha^*(\beta,\gamma)}$. To construct a network, use the networks constructed previously and assign design speeds as follows. For links of the large-scale network, which will be used by routes for a distance of order $n^{1/2}$, we design speed of order 1. For type $m$ edges in the hierarchical routing lattice, which will be used by routes for a distance of order $2^m$, we design speed of order $2^m n^{-1/2} \log n$. For the local links of the form $(x, v(x))$, which will be used by routes for distance $\mathrm{O}(1)$, we design speed of order $n^{-1/2}$. This ensures the typical times $t(x_i, x_j)$ are of order $n^{1/2}$ as required. To calculate the cost, we simply combine the previous estimates (4), (5) and (6) of costs of providing

flow volumes of different links with the costs of the design speeds stipulated above; the total cost is of the order of

$$n^{1-(\beta/2)} \times n^{-\gamma/2} + \sum_{m=0}^{M} n^{1-(\beta/2)} 2^{m(2\beta-1)} \times (2^m n^{-1/2} \log n)^\gamma + n^{(1+\beta)/2} \times 1$$

and this works out to be of the form $n^{\alpha^*(\beta,\gamma)}$ stated.

## Acknowledgment

## References

[1] Gilbert E N, *Minimum cost communication networks*, 1967 *Bell Syst. Tech. J.* **46** 2209–27
[2] Thomas D A and Weng J F, *Minimum cost flow-dependent communication networks*, 2006 *Networks* **48** 39–46
[3] Aldous D J and Kendall W S, *Short-length routes in low-cost networks via Poisson line patterns*, 2007 *Adv. Appl. Probab.* at press [math.PR/0701140]
[4] Aldous D J, *Asymptotics and optimality for hub and spoke models in spatial transportation networks*, 2007 *Math. Proc. Camb. Phil. Soc.* at press [cond-mat/0702502]
[5] Hayashi Y and Matsukubo J, *A review of recent studies of geographical scale-free networks*, 2006 *IPSJ Trans.* **47** 776 [physics/0512011]
[6] Donetti L, Neri F and Muñoz M A, *Optimal network topologies: expanders, cages, Ramanujan graphs, entangled networks and all that*, 2006 *J. Stat. Mech.* P08007
[7] Bejan A, 2000 *Shape and Structure, from Engineering to Nature* (Cambridge: Cambridge University Press)
[8] Barthélemy M and Flammini A, *Optimal traffic networks*, 2006 *J. Stat. Mech.* L07002
[9] Gastner M T and Newman M E J, *Shape and efficiency in spatial distribution networks*, 2006 *J. Stat. Mech.* P01015
[10] Stoyan D, Kendall W S and Mecke J, *Stochastic geometry and its applications*, 1995 *Applied Probability and Statistics* (*Wiley Series in Probability and Mathematical Statistics*) 2nd edn (Chichester: Wiley)