Uniform Multicommodity Flow through the Complete Graph with Random Edge-capacities

David J. Aldous [*] Colin McDiarmid [†] Alex Scott [‡]

Abstract

Give random capacities C to the edges of the complete n-vertex graph. Consider the maximum flow Φ_n that can be simultaneously routed between each source-destination pair. We prove that $\Phi_n \to \phi$ in probability where the limit constant ϕ depends on the distribution of C in a simple way, and that asymptotically one need use only one- and two-step routes. The proof uses a reduction to a random graph problem.
Keywords: multicommodity flow, graph colouring.

1 Introduction

This paper is part of a project studying optimal flows through random networks, where a network has both a graph structure and extra structure such as capacities and costs on edges, and where we are in the "multicommodity flow" setting with simultaneous flows between each source-destination pair. Possible models span a broad spectrum from realistic to mathematically tractable, and at the latter end are models based on the complete graph. Including study of such models within a project is natural both for mathematical completeness and for comparison purposes.

[*]Department of Statistics, University of California, Berkeley, CA 94720-3860, USA; email: aldous@stat.berkeley.edu. Research supported by N.S.F Grant DMS-0704159.

[†]Department of Statistics, 1 South Parks Road, Oxford, OX1 3TG, United Kingdom; email: cmcd@stats.ox.ac.uk.

[‡]Mathematical Institute, University of Oxford, 24-29 St Giles, Oxford OX1 3LB, UK; email: scott@maths.ox.ac.uk.

Consider first the setting of an arbitrary finite connected undirected graph G. Let $\phi > 0$. A flow of volume $\phi/2$ between vertex v and vertex w has net out-flow $= \phi/2$ at v, net out-flow $= -\phi/2$ at w, and zero net out-flow at other vertices. For such a flow write $f_{v,w}(e) \geq 0$ for the absolute value of the flow volume across an undirected edge e. Suppose we have such a flow simultaneously for each ordered pair (v, w) with $w \neq v$; call this collection a *uniform flow* of volume ϕ and write $f(e) := \sum_{(v,w)} f_{v,w}(e)$ for the combined volume of flow across the undirected edge e. Suppose now we are given capacities $C(e)$ for edges e. Then the *maximum uniform flow volume* (MUFV) is defined to be the largest ϕ such that there exists a uniform flow of volume ϕ which satisfies the capacity constraints

$$f(e) \leq C(e) \ \forall e. \tag{1}$$

One modeling paradigm, seeking to combine the spatial inhomogeneity of real networks with mathematical tractability, is to consider some standard family G_n of n-vertex graphs, and to assume the edge-capacities $C(e)$ are random (specifically, are i.i.d. copies of a reference r.v. C). Now the MUFV is a r.v. Φ_n, and one can seek to study its $n \to \infty$ behaviour.

Apparently, and somewhat surprisingly, such questions have not been studied before. There is literature [5, 8, 10, 11] on flows with a single source-destination pair and on flows from the top to the bottom of a square, but these fall into the one-commodity setting of the max-flow min-cut theorem, rather than our multicommodity setting.

In this paper we consider the complete graph; a similar problem on the $m \times m$ square grid was studied by very different methods in [3]. An interesting observation is that in both these models the limit constants for Φ_n depend on the distribution of C (not just on its expectation EC), but for rather different reasons in the two models. An intermediate model is the cube $\{0, 1\}^d$, and here we conjecture that the limit constant does depend only on EC when C is bounded away from zero. See Section 4 for further related work and open problems.

1.1 Statement of results

Consider the complete n-vertex graph whose edges e have independent random capacities $C(e)$ whose common distribution satisfies

$$0 < EC < \infty. \tag{2}$$

Note that the function

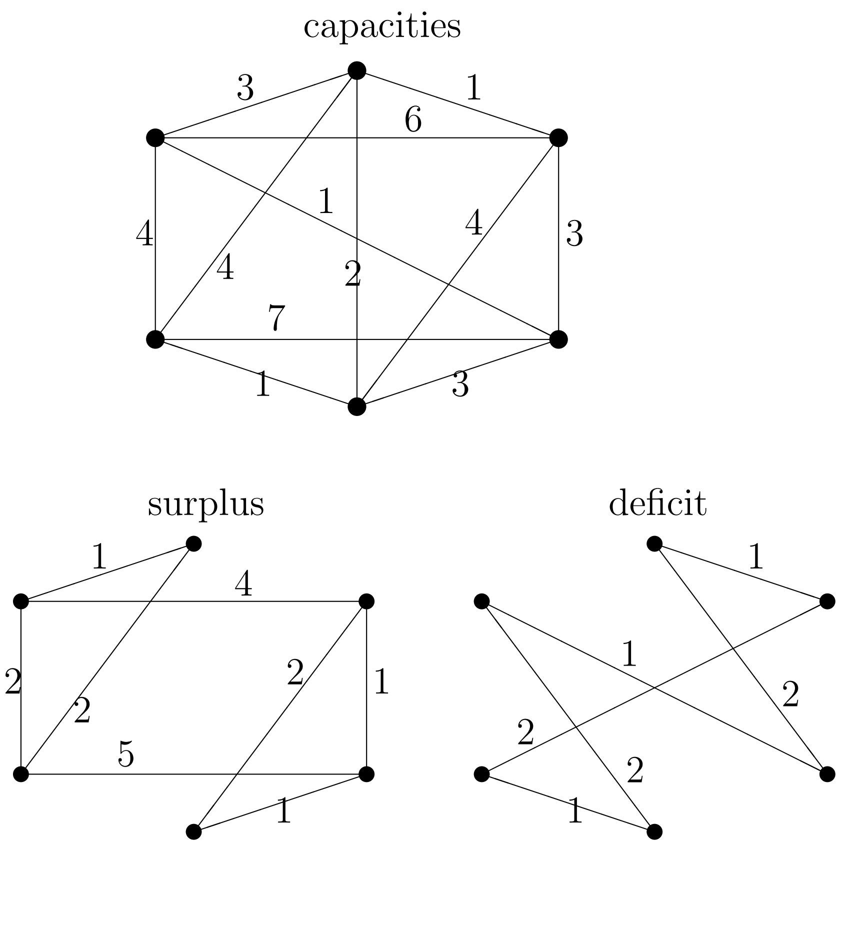
$$\phi \to 2E \max(\phi - C, 0) \ - \ E \max(C - \phi, 0)$$

in continuous and strictly increasing from $-EC$ to ∞ as ϕ increases from 0 to ∞, and so we can define a constant $0 < \phi_* < \infty$ as the unique solution of

$$E \max(C - \phi_*, 0) = 2E \max(\phi_* - C, 0). \tag{3}$$

For example, if C is uniform on $\{0, 1\}$ then $\phi_* = 1/3$, and if C is uniform on $[0, 1]$ then $\phi_* = \sqrt{2} - 1$.

Theorem 1. *Under assumption (2) the MUFV Φ_n satisfies $\Phi_n \to \phi_*$ in probability as $n \to \infty$.*



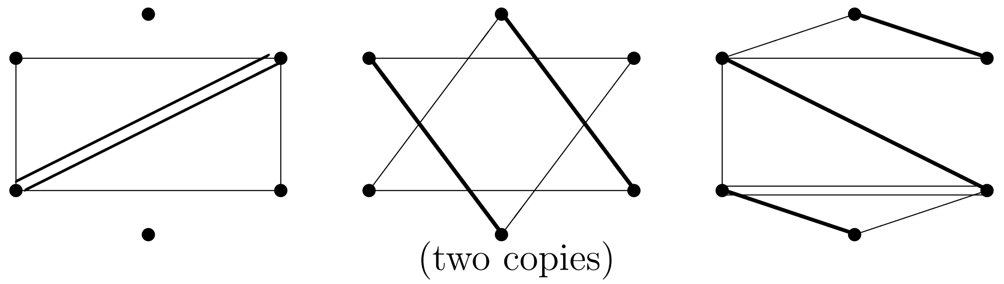


Figure 1: a worked example

In order to gain an intuitive understanding of the problem, let us consider a small example. Figure 1 shows (top) a particular instance of integer edge-capacities with $n = 6$: three edges have zero capacity and are not shown. How might we try to route a flow of volume 2 between each vertex pair? First try to route along the direct edge e; if capacity $C(e) \geq 2$ then we can do so, leaving surplus capacity $C(e) - 2$, and the center left diagram shows the graph of surplus capacity edges. If instead $C(e) < 2$ then the edge has "deficit" $2 - C(e)$ and the center right diagram shows the graph of deficit capacity edges. Regard each of those graphs as a multigraph whose edges have surplus/deficit 1. To complete the routing it is enough to associate each deficit edge with a pair of surplus edges forming a triangle (with each surplus edge being used at most once), because then the flow required between endpoints of the deficit edge can be routed through the two surplus edges. The deficit multigraph corresponding to Figure 1 has 9 edges. The bottom row diagrams illustrate one way of producing the required 9 triangles (the deficit edges are drawn thicker).

How does this construction idea lead intuitively to the formula (3) and Theorem 1? Suppose we wish to route a uniform flow of volume ϕ. As in the example, first route as much flow as possible across the direct edge, that is route volume $\min(\phi, C(e))$ across an edge e. This leaves a deficit volume $\max(\phi - C(e), 0)$ for edge e. Now the mean surplus capacity per edge is $E \max(C - \phi, 0)$. We try to route the unsatisfied demand via 2-step paths with surplus capacity; for this to work it is plausibly necessary that

$$E \max(C - \phi, 0) \geq 2E \max(\phi - C, 0).$$

Conversely, this should be sufficient because the set of edges with surplus capacity forms a dense random graph which should be sufficiently well-connected to permit construction of the desired 2-step paths.

The "necessary" part is indeed easy to formalize (Lemma 2), and we do this in Section 2. We prove the converse (Lemma 5) in Section 3.1; the proof uses a reduction to a result (Proposition 3) on random coloured graphs which we prove in Section 3.2.

Finally, we mention a possible connection between our setting and the more elaborate setting of *dynamic routing* (of e.g. phone calls) on a complete network. Section 2.1 of [7] analyzes the throughput of a model in which it is assumed that calls use either a one-link or two-link route, having previously commented

> We shall occasionally mention the possibility that a call might be connected along a path of more than two links, but ... this possibility is rarely of interest and we shall exclude it from our formal development.

Our results suggest the possibility of proving that asymptotically one cannot improve throughput by using such longer paths.

2 The upper bound

The upper bound in Theorem 1 is provided by

Lemma 2. *Fix* $\phi > \phi_*$. *Then* $\lim_n P(\Phi_n \geq \phi) = 0$.

Proof. Fix a realization of the edge-capacities. Suppose a uniform flow of volume ρ exists. For an edge (v, w)

$$\begin{aligned} \sum_e (f_{v,w}(e) + f_{w,v}(e)) &\geq \rho \text{ if } C(vw) \geq \rho \\ &\geq C(vw) + 2(\rho - C(vw)) \text{ if } C(vw) \leq \rho \end{aligned}$$

because in the latter case volume of at least $\rho - C(vw)$ must use at least a 2-step route. Combining the two cases,

$$\sum_e (f_{v,w}(e) + f_{w,v}(e)) \geq \min(\rho, C(vw)) + 2\max(\rho - C(vw), 0).$$

Summing over edges $e' = (v, w)$ and using the capacity constraint (1),

$$\sum_e C(e) \geq \sum_{e'} \left(\min(\rho, C(e')) + 2\max(\rho - C(e'), 0)\right).$$

Dividing by $\binom{n}{2}$ and recalling we supposed that the uniform flow exists, we have shown

$$Q_n := \frac{1}{\binom{n}{2}} \sum_{e'} C(e') - \frac{1}{\binom{n}{2}} \sum_{e'} \left(\min(\rho, C(e')) + 2\max(\rho - C(e'), 0)\right) \geq 0 \text{ on } \{\Phi_n \geq \phi\}.$$

But as $n \to \infty$ the quantity Q_n converges in probability to

$$q := EC - (E\min(\rho, C) + 2E\max(\rho - C, 0)) = E\max(C - \rho, 0) - 2E\max(\rho - C, 0).$$

If $\rho > \phi_*$ then $q < 0$ and hence we must have $\lim_n P(\Phi_n \geq \phi) = 0$. □

3 The reduction argument

3.1 The reduction

We will use a reduction to the following "random graph" result. To motivate this reduction, consider the case where the edge-capacity C takes only values $\{0, 1, 2\}$ and where we seek to route a uniform flow of volume 1. Then traffic across capacity-0 edges (colored scarlet, say) needs to be routed through two capacity-2 edges (colored blue, say). Colors are mnemonics for *smaller* and *bigger* capacity.

Proposition 3. *Fix $0 < p_s < p_b/2$ with $p_s + p_b \leq 1$. Randomly colour the edges of the complete n-vertex graph as blue (probability p_b) or scarlet (probability p_s) or neither (probability $1 - p_b - p_s$). Then whp there exists a collection of edge-disjoint triangles, each triangle having one scarlet edge and two blue edges, such that every scarlet edge is in some triangle.*

We defer the proof of this proposition to the next subsection, and show here how to deduce the lower bound in Theorem 1, stated as Lemma 5. Note that the condition $\rho < \phi_*$ is equivalent to

$$r := \frac{E\max(\rho - C, 0)}{E\max(C - \rho, 0)} < \frac{1}{2}. \tag{4}$$

We first prove a version of the result for integer capacities and demands, and then use this to deal with the general case in Lemma 5.

Lemma 4. *Suppose C is integer-valued and bounded, and suppose ρ is an integer satisfying* (4)*. Then, with high probability, we can construct flows of volume ρ between every pair of vertices such that the capacity constraint* (1) *holds.*

Proof. Let M be an upper bound for C. We construct M separate flow problems $\mathcal{P}_1, \ldots, \mathcal{P}_M$ of the following type. Each problem $\mathcal{P}_i$ will be encoded by an n-vertex graph with scarlet and blue edges. A scarlet edge vw indicates a demand of 1 and capacity of 0 between v and w, while blue edges have capacity 1 and demand 0. The absence of an edge indicates that demand and capacity are both 0.

We construct the problems as follows. For each edge vw with $C(vw) \geq \rho$ we choose (uniformly at random) a subset $I \subset \{1, \ldots, M\}$ of size $C(vw) - \rho$ and insert a blue edge between v and w in $\mathcal{P}_i$ for each $i \in I$. For each edge vw with $C(vw) < \rho$ we choose (uniformly at random) a subset $I \subset \{1, \ldots, M\}$ of size $\rho - C(vw)$ and insert a scarlet edge between v and w in $\mathcal{P}_i$ for each $i \in I$. Note that the instances $\mathcal{P}_i$ and $\mathcal{P}_j$ may be dependent, but the edges inside any instance $\mathcal{P}_i$ are present (and coloured) independently.

By (4) the hypothesis of Proposition 3 is satisfied, so with high probability, we can find for each instance $\mathcal{P}_i$ a collection of edge-disjoint triangles with two scarlet edges and one blue edge covering all scarlet edges. For each scarlet edge in one of the triangles, route unit flow between its end-vertices by using the two blue edges in the triangle. Taking the sum of these flows over all M instances, we establish the lemma. □

Lemma 5. *Assume (2) and let $\phi < \phi_*$. Then* $\lim_n P(\Phi_n \geq \phi) = 1$.

Proof. Let $p_0 = P(C > 0)$ and choose $c_0 > 0$ such that $P(C \geq c_0) \geq p_0/2$. Define

$$C_k = \max\{\min(2^{-k}\lfloor C2^k - 1\rfloor, k), 0\}$$

for k sufficiently large that $2^{-k} < c_0$. So $0 \leq C_k \leq \max\{C - 2^{-k}, 0\}$. Define ρ_k as the largest multiple of 2^{-k} for which

$$E \max(C_k - \rho_k, 0) > 2E \max(\rho_k - C_k, 0).$$

It is easy to check that $\rho_k \uparrow \phi_*$ as $k \to \infty$. Thus it is sufficient to show that, for each fixed large k,

$$P(\text{uniform flow of volume } \rho_k \text{ exists}) \to 1 \text{ as } n \to \infty. \tag{5}$$

But by applying Lemma 4 to the bounded integer-valued quantities $2^k C_k$ and $2^k \rho_k$, then rescaling by a factor 2^{-k}, we find that (with high probability) we can construct flows of volume ρ_k between every pair such that the total flow volume $f(e)$ satisfies the capacity constraints $f(e) \leq C_k(e)\ \forall e$. □

3.2 Proof of Proposition 3

Given nonnegative reals $p_1, \ldots, p_k$ with $\sum_{i=1}^k p_i \leq 1$, we write $\mathcal{G}(n; p_1, \ldots, p_k)$ for the probability space of edge-coloured graphs on n vertices, obtained as follows: for each pair of vertices independently we have an edge of colour i with probability p_i, and no edge with probability $1 - \sum_{i=1}^k p_i$. Proposition 3 follows immediately from the following result (give scarlet edges colour 1, and blue edges colour 2 or colour 3 with probability 1/2 each).

Lemma 6. *Fix $\delta > 0$, and suppose that $p_1, p_2, p_3 \geq 0$ with sum at most one satisfy $p_1 + \delta \leq \min\{p_2, p_3\}$. Then for $G \in \mathcal{G}(n; p_1, p_2, p_3)$ there is whp a collection $\mathcal{T}$ of edge-disjoint triangles such that every triangle in $\mathcal{T}$ contains one edge of each colour and every edge of colour 1 is contained in some triangle in $\mathcal{T}$.*

Proof. The proof will go in two steps: we begin by setting aside a subset of the edges of colours 2 and 3, and use the remainder to cover most of the edges of colour 1. We then use the edges we have set aside to cover the remaining edges of colour 1.

Let $G \in \mathcal{G}(n; p_1, p_2, p_3)$. We define edge-disjoint subgraphs G_1 and G_2 of G with $V(G_1) = V(G_2) = V(G)$ as follows: G_1 contains all edges of colour 1; each edge of colour $i > 1$ is placed in G_1 with probability p_1/p_i and in G_2 with probability δ/p_i (and is discarded with the remaining probability $(p_i - p_1 - \delta)/p_i \geq 0$). Then G_1 has distribution $\mathcal{G}(n; p_1, p_1, p_1)$ and G_2 has distribution $\mathcal{G}(n; 0, \delta, \delta)$.

We begin with G_1, and try to cover edges of colour 1 with multicoloured triangles. We shall first partition $V = V(G_1)$ into a number of sets of size $\Theta(\sqrt{n})$: we ignore the edges inside vertex sets, and break up the remainder into edge-disjoint tripartite graphs that we handle separately.

Recall that a *Steiner triple system* on a set U is a collection $\mathcal{S}$ of triples in U such that every pair of elements of U is contained in exactly one triple from $\mathcal{S}$. Steiner triple systems exist whenever $|U| \equiv 1$ or $3 \bmod 6$. So let $t \sim \sqrt{n}$ be of form $6k+1$ and let $V_1, \ldots, V_t$ be a partition of V into t sets of size $\lfloor n/t \rfloor$ or $\lceil n/t \rceil$. Let $\mathcal{S}$ be a Steiner triple system on $[t] = \{1, \ldots, t\}$, so $\mathcal{S}$

contains $\binom{t}{2}/3$ triples. For each element $S = \{i, j, k\}$ of $\mathcal{S}$, we consider the corresponding tripartite subgraph G_S of G_1 with vertex classes (V_i, V_j, V_k) that contains all edges from G_1 joining these vertex classes. The graphs $\{G_S : S \in \mathcal{S}\}$ are independent (disjoint) tripartite random graphs, and every edge of G_1 is contained either in some vertex class V_i or in exactly one of the G_S.

Consider a fixed $S = \{\alpha_1, \alpha_2, \alpha_3\} \in \mathcal{S}$. The corresponding subgraph G_S has vertex classes $V_{\alpha_1}, V_{\alpha_2}$ and V_{α_3}. We decompose G_S into subgraphs G_S^1, G_S^2 and G_S^3 where, for each $i = 1, 2, 3$, G_S^i has all edges from G_S with colour j that join $V_{\alpha_{i+j}}$ and $V_{\alpha_{i+j+1}}$ for $j = 1, 2, 3$ (all subscripts taken modulo 3). Thus each G_S^i has edges of different colours between different pairs of vertex classes, and in particular all triangles in G_S^i have edges of three different colours. Furthermore, if we ignore colours, then G_S^i is a random tripartite graph with vertex classes $V_{\alpha_1}, V_{\alpha_2}, V_{\alpha_3}$ and edge probability p_1.

Now fix i, and consider G_S^i. Given vertices x, y in different classes, the number X of x-y paths of length 2 in G_S^i has distribution $B(r, p_1^2)$, where $r \sim n/t \sim \sqrt{n}$ is the size of the third vertex class. By Chernoff's inequality, the probability that $|X - rp_1^2| \geq n^{\frac{1}{2}}/\ln n$ is at most $\exp(-\Omega(n^{\frac{1}{2}}/\ln^2 n))$. It follows that, with failure probability $O(\exp(-n^{1/3}))$, every edge of G_S^i is contained in $(1 + o(1))p_1^2 n/t$ triangles. Thus, giving each triangle in G_S^i weight $(1 - o(1))t/np_1^2$, we obtain a fractional triangle-packing of size $(1 - o(1))e(G_S^i)/3$. We now use Theorem 1 of Haxell and Rödl [9], which implies that the maximum size of a triangle-packing in a graph on m vertices differs from the maximum size of a fractional triangle-packing by $o(m^2)$ (we could also use arguments of Frankl and Rödl [6]). Applying this result with $m \sim 3n/t \sim 3\sqrt{n}$, we see that, with failure probability as above, there is a (proper) triangle-packing in G_S^i of size $(1 + o(1))e(G_S^i)/3 - o(n)$, which thus covers all but $o(n)$ edges from G_S^i; and by symmetry we may require that at each vertex of G_S^i the expected number of uncovered edges is $o(n/t) = o(\sqrt{n})$.

We therefore have, with probability $1 - o(1)$, for every S and i, a triangle-packing such that at each vertex v of G_S^i the expected number of uncovered edges is $o(\sqrt{n})$, and this happens uniformly over all S, i and v. Let $\mathcal{T}_1$ be the union of these triangle-packings, and let H be the subgraph of G_1 that remains after removing all the triangles in $\mathcal{T}_1$.

We claim that, with probability $1 - o(1)$, H has maximum degree $o(n)$. Consider a vertex v, say $v \in V_i$. The edges inside V_i only contribute $O(\sqrt{n})$ to the degree of v, so it is enough to consider edges between vertex classes. Since

i belongs to $(t-1)/2 \sim \sqrt{n}/2$ triples from $\mathcal{S}$, v belongs to $(1+o(1))\sqrt{n}/2$ subgraphs G_S. These subgraphs and their packings are independent: each subgraph contains at most $2\lceil n/t \rceil \sim 2\sqrt{n}$ edges that remain incident with v in H, and an expected $o(\sqrt{n})$ such edges. Thus $Ed_H(v) = o(n)$ and $d_H(v) - Ed_H(v)$ is the sum of $O(\sqrt{n})$ independent random variables, each with absolute value at most $O(\sqrt{n})$. It follows by (for instance) the Azuma-Hoeffding inequality that with probability $1 - o(1/n)$, we have $\Delta(H) = o(n)$. Let us choose n large enough that we can assume $\Delta(H) < \delta^2 n/10$.

Finally, we use the edges from G_2 to cover the remaining edges of colour 1 in H. Let us orient the edges of G_2 at random. Then (easily, by Chernoff's inequality) with probability $1 - o(1)$, for every ordered pair (x, y) of vertices there are at least $\delta^2 n/5$ oriented paths xzy in G_2 such that xz has colour 2 and zy has colour 3. We now choose triangles greedily: for each edge xy of colour 1 (taking edges in arbitrary order and with arbitrary orientation), we pick an oriented path xzy in G_2 (with xz of colour 2 and zy of colour 3). Since we have previously used at most $\Delta(H) - 1$ edges out of x and at most $\Delta(H) - 1$ edges into y, and there are at least $2\Delta(H)$ directed paths to choose from, there is at least one path edge-disjoint from all previous choices. This enables us to cover all colour 1 edges of H; adding the resulting triangles to $\mathcal{T}_1$ gives our desired collection $\mathcal{T}$ of triangles. □

4 Related work and open problems

As mentioned in the Introduction, the problem studied here for the complete graph could be posed for any family of n-vertex graphs, and the cube graph $\{0, 1\}^d$ seems particularly interesting. For a large class of related problems, suppose there exist feasible flows of given volume but that our objective is to minimize (over such flows) some "cost" of the flow. Such problems have been intensively studied as finite algorithmic problems [1] but only sporadically studied in our "probabilistic model of network and $n \to \infty$" setting. For instance, consider the complete n-vertex graph with independent exponentially distributed edge-lengths, and for each pair of distinct vertices send unit flow along the shortest path between them, without capacity constraints. A typical route uses about $\ln n$ short edges (and is much shorter than the one direct edge). The explicit limit distribution of suitably scaled flows across different edges is obtained in [4]. But it seems a challenging problem to understand what happens when edge capacities are imposed, in which case we can no longer use shortest-path routing but instead seek to minimize overall mean route length: how does optimal mean route length vary with capacity constraint?

Another variant is to replace the "hard constraint" of edge-capacities by a "soft constraint" of congestion costs – the cost to the system of a volume of flow across an edge grows super-linearly with volume. This variant was studied in the lattice setting in [3] and in a locally tree-like directed network in [2]; but analytic understanding of the behavior of any more realistic model of e.g. road networks seems far out of reach, as does understanding of the optimal *design* of such networks.

Acknowledgement We would like to thank a referee for detailed comments.

References

[1] R.K. Ahuja, T.L. Magnanti and J.B. Orlin, *Network Flows: Theory, Algorithms and Applications*, Prentice Hall, 1993

[2] D.J. Aldous, Cost-volume relationship for flows through a disordered network, *Math. Oper. Res.* **33** (2008), 769–786

[3] D. J. Aldous, Optimal flow through the disordered lattice, *Ann. Probab.* **35** (2007), 397–438.

[4] D.J. Aldous and S. Bhamidi, Edge flows in the complete random-lengths network, `http://arxiv.org/abs/0708.0555`, (2007)

[5] J. T. Chayes and L. Chayes, Bulk transport properties and exponent inequalities for random resistor and flow networks, *Comm. Math. Phys.* **105** (1986), 133–152

[6] P. Frankl and V. Rödl, Near perfect coverings in graphs and hypergraphs, *European J. Combinatorics* **6** (1985), 317–326

[7] R.J. Gibbens and F.P. Kelly, Dynamic routing in fully connected networks, *IMA J. Math. Control Inform.* **7** (1990), 77–111

[8] G. Grimmett and H. Kesten, First-passage percolation, network flows and electrical resistances, *Z. Wahrsch. Verw. Gebiete* **66** (1984), 335–366

[9] P.E. Haxell and V. Rödl, Integer and fractional packings in dense graphs, *Combinatorica* **21** (2001), 13–38

[10] H. Kesten, Surfaces with minimal random weights and maximal flows: a higher-dimensional version of first-passage percolation, *Illinois J. Math.* **31** (1987), 99–166

[11] Y. Zhang, Critical behavior for maximal flows on the cubic lattice, *J. Statist. Phys.* **98** (2000), 799–811