

# The VPN Conjecture Is True

NAVIN GOYAL, Microsoft Research  
 NEIL OLVER, MIT  
 F. BRUCE SHEPHERD, McGill University

We consider the following network design problem. We are given an undirected graph  $G = (V, E)$  with edge costs  $c(e)$  and a set of terminal nodes  $W \subseteq V$ . A *hose* demand matrix is any symmetric matrix  $D$ , indexed by the terminals, such that for each  $i \in W$ ,  $\sum_{j \neq i} D_{ij} \leq 1$ . We must compute the minimum-cost edge capacities that are able to support the oblivious routing of every hose matrix in the network. An *oblivious routing* template, in this context, is a simple path  $P_{ij}$  for each pair  $i, j \in W$ . Given such a template, if we are to route a demand matrix  $D$ , then for each  $i, j$ , we send  $D_{ij}$  units of flow along each  $P_{ij}$ . Fingerhut et al. [1997] and Gupta et al. [2001] obtained a 2-approximation for this problem, using a solution template in the form of a tree. It has been widely asked and subsequently conjectured [Italiano et al. 2006] that this solution actually results in the optimal capacity for the single-path VPN design problem; this has become known as the *VPN Conjecture*. The conjecture has previously been proven for some restricted classes of graphs [Fingerhut et al. 1997; Fiorini et al. 2007; Grandoni et al. 2008; Hurkens et al. 2007]. Our main theorem establishes that this conjecture is true in general graphs. This also has the implication that the single-path VPN problem is solvable in polynomial time.

A natural fractional version of the conjecture had also been proposed [Hurkens et al. 2007]. In this version, the routing may split flow between many paths, in specified proportions. We demonstrate that this multipath version of the conjecture is in fact false. The multipath and single path versions of the VPN problem are essentially direct analogues of the randomized and nonrandomized versions of oblivious routing schemes for minimizing congestion for permutation routing [Borodin and Hopcroft 1982; Valiant 1982].

Categories and Subject Descriptors: C.2.1 [Computer-Communication Networks]: Network Architecture and Design; F.2.2 [Analysis of Algorithms and Problem Complexity]: Nonnumerical Algorithms and Problems; G.2.2 [Discrete Mathematics]: Graph Theory

General Terms: Theory, Algorithms

Additional Key Words and Phrases: Network design, robust optimization, oblivious routing

## ACM Reference Format:

Goyal, N., Olver, N., and Shepherd, F. B. 2013. The VPN conjecture is true. *J. ACM* 60, 3, Article 17 (June 2013), 17 pages.  
 DOI : <http://dx.doi.org/10.1145/2487241.2487243>

## 1. INTRODUCTION

Robustness has become one of several paradigms used to address data uncertainty in optimization. While the input data may not be known perfectly, it assumes that

---

N. Olver was supported by a MELS Quebec Merit Scholarship for Foreign Students and a Schulich Fellowship. F. B. Shepherd was supported by a NSERC Discovery Grant.

Authors' addresses: N. Goyal, Microsoft Research India, "Vigyan", #9, Lavelle Road, Bangalore 560 001, India; email: [navingo@microsoft.com](mailto:navingo@microsoft.com); N. Olver, Massachusetts Institute of Technology, Department of Mathematics, 77 Massachusetts Avenue, Cambridge, MA 02139-4307; email: [olver@math.mit.edu](mailto:olver@math.mit.edu); F. B. Shepherd, Burnside Hall, Room 1113, McGill University, 805 Sherbrooke West, Montreal, Quebec H3A 2K6, Canada; email: [bruce.shepherd@mcgill.ca](mailto:bruce.shepherd@mcgill.ca).

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or [permissions@acm.org](mailto:permissions@acm.org).

© 2013 ACM 0004-5411/2013/06-ART17 \$15.00  
 DOI : <http://dx.doi.org/10.1145/2487241.2487243>

each input comes from some known bounded region (called the *universe* or *uncertainty set*) of legal inputs. Robust optimization then seeks an optimal solution that is valid for all possible inputs in the universe; robustness is thus a worst-case approach for coping with uncertainty. The area has seen significant development in the 1990's due especially to work by Ben-Tal and Nemirovski [1998, 1999] and El Ghaoui et al. [1997, 1998] (see also Ben-Tal et al. [2009] for more references).

Robust optimization has also gained popularity in discrete optimization [Bertsimas and Sim 2003; Kouvelis and Yu 1997], and one particularly active area of research arises in the context of network design. This work is largely motivated by the fact that traffic demands in internet protocol (IP) networks are often hard to determine and/or are rapidly changing. In the basic model for robust network design the input consists of a graph (network topology) and a set of terminals that must be connected. One must determine how much capacity to allocate to each edge of the network so that all traffic patterns can be supported. The uncertainty in the problem is that the network may have to support a wide variety of traffic patterns. We can represent a specific traffic pattern with a *demand matrix*; this is a matrix  $D$ , indexed by the terminals, with entry  $D_{ij}$  specifying the demand from  $i$  to  $j$ . It is assumed that we are given a universe  $\mathcal{U}$  of possible demand matrices, specified as a polyhedron (or more generally, as a convex body). The problem is to design a "minimum cost" network such that each demand matrix in  $\mathcal{U}$  can be routed in the resulting network. There are several possible cost metrics. For instance, the *congestion measure* would seek to minimize the maximum capacity. This represents a broad class of network optimization problems that contains as special cases problems that have been widely studied. This includes problems in the theory of parallel computation [Borodin and Hopcroft 1982; Valiant and Brebner 1981], multi-commodity flows [Leighton and Rao 1999] and online algorithms [Räcke 2002, 2008]. In the present paper, we consider the *total cost measure*; we assume that each edge has a per-unit cost of reserving capacity, and the aim is to minimize the sum of capacity costs.

We have omitted a crucial issue in our discussion so far: the way in which demands are routed between terminals. In particular, we need to specify whether and how rapidly the routing may be updated in response to changes in the traffic matrix. In *dynamic* routing, one has total freedom; the routing may be fractional, and may be completely updated whenever incoming traffic demands change. This version was introduced in Ben-Ameur and Kerivin [2003] and can be seen to be a convex optimization problem. However, it turns out to be coNP-hard [Chekuri et al. 2007] based on a reduction from (uniform) sparsest cut, which also shows that it is hard to approximate if sparsest cut is.

A second well-known (and more practical) model is that of *oblivious routing*. This approach, inspired by routing in packet networks, asks for a *template* that defines ahead of time how any future demands will be routed. For each terminal pair  $i, j$ , the template specifies a path  $P_{ij}$  joining  $i, j$ . The interpretation is that if there is a future demand of  $D_{ij}$  between terminals  $i$  and  $j$ , then we will send  $D_{ij}$  flow along  $P_{ij}$ . A relaxed form of the oblivious routing model would only require a unit flow  $f_{ij}$  for each pair, and we route the demand proportionally according to the flow. We refer to these models as *single-path routing* (SPR) and *multipath routing* (MPR) respectively. In Ben-Ameur and Kerivin [2005], it is shown that the robust design problem is polytime solvable for the MPR model if the universe  $\mathcal{U}$  has a polytime separation oracle. In contrast, the SPR model was recently shown to be hard to approximate to within polylogarithmic factors [Olver and Shepherd 2010] (and using tree embedding techniques, an  $O(\log n)$ -factor approximation is possible [Gupta 2004, private communication]).

### 1.1. The VPN Problem

We have discussed the robust network design model and known results for general universes  $\mathcal{U}$ . As with general robust optimization, there has also been intensive study of specific fixed universes. A class of particular interest (indeed it was well studied before the general robust network design model was even formulated) is the *hose model*. The hose model was independently formulated by Fingerhut et al. [1997], motivated by applications to ATM broadband networks, and by Duffield et al. [1999], motivated by the specification of virtual private networks (VPNs). This latter motivation is particularly important. An enterprise customer wishes to set up a private network (perhaps connecting various company branches), on top of an underlying network of a large telecommunications operator. The customer has specific bandwidth requirements, and in order to satisfy these, the operator will need to reserve capacity on the network for this customer's use. The amount and distribution of this reserved bandwidth depends on the exact customer requirements, and these need to be specified in some way. The simplest way is via point-to-point demands; the customer specifies the bandwidth requirement for each pair of terminals in the VPN, and the network must be able to route all these demands simultaneously. This is known as the *pipe model*. However, the demand pattern across the VPN is unlikely to be fixed; more likely, demands will fluctuate over time. In order for the customer to give a specification in the pipe model that covers all likely circumstances, the worst-case demand between every pair would have to be used.

In the hose model, a different and more flexible description of the customers requirements is given. Rather than specifying the exact demand between each pair of terminals, only a maximal *aggregate* demand at each terminal needs to be specified. More precisely, the customer must specify, for each terminal in the VPN, an upper bound on the total demand terminating at that terminal. The service guarantee provided by the network operator is that any pattern of demands across the network must be routable, as long as they respect these upper bounds.

The model comes in both symmetric and asymmetric flavors; it is the symmetric version that will concern us in this article. The symmetric hose universe is determined by a single *marginal*  $b_i$  for each terminal  $i$ , which gives the upper bound on the total demand terminating at  $i$ . The associated class of hose matrices is

$$\mathcal{H}(\mathbf{b}) = \{D \in \mathbb{R}_+^{|W| \times |W|} : D_{ij} = D_{ji} \text{ and } \sum_{j \in W} D_{ij} \leq b_i \forall i \in W\}; \quad (1)$$

we will write just  $\mathcal{H}$  when  $\mathbf{b}$  is understood. The term “symmetric” is used because we consider demand to be between an *unordered* pair of terminals;  $D_{ij}$  and  $D_{ji}$  refer to the same demand. We thus also require that  $P_{ij} = P_{ji}$ ,  $f_{ij} = f_{ji}$ , as appropriate. There is no such restriction in the asymmetric hose model, where demands are considered to be oriented. Moreover, two distinct marginals  $b_i^-$  and  $b_i^+$  are associated with each terminal  $i$ , representing the upper bounds on the total incoming and outgoing demands, respectively. The asymmetric hose universe is then

$$\mathcal{H}(\mathbf{b}^+, \mathbf{b}^-) = \{D \in \mathbb{R}_+^{|W| \times |W|} : \sum_{j \in W} D_{ij} \leq b_i^+ \text{ and } \sum_{j \in W} D_{ji} \leq b_i^- \forall i \in W\}.$$

The *VPN problem* asks for the cheapest capacity assignment that can support all the corresponding hose matrices.

Erlebach and Rügge [2004] showed that the MPR version of the VPN problem is polynomially solvable, using the ellipsoid method. Subsequently, compact LP formulations were given independently by Altın et al. [2007] and Hurkens et al. [2007]

(see Section 3). Both of these results can easily be extended to the general robust network design model.

The first theoretical results for single-path routing were given by Fingerhut et al. [1997] and (independently, but later) by Gupta et al. [2001]. In both works, a factor 2 approximation algorithm is given based on a particularly simple structure, namely, they find a node  $r$  (not necessarily a terminal) such that the oblivious routing is defined by a shortest path tree  $T$  rooted at  $r$ . The optimal capacity is obtained as follows. For each node  $v$ , route  $b_v$  units from  $v$  on its (shortest) path in  $T$  to  $r$ . The optimal network design is just the aggregate of these capacities for each  $v$ . We call the resulting capacitated tree the *VPN tree* for  $r$ . It had been discussed [Gupta and Kumar 2001, private communication] and subsequently conjectured [Italiano et al. 2006] that this solution actually results in the optimal capacity for the VPN problem; this has become known as the *VPN Conjecture*.

The conjecture has previously been proven for some special cases. Fingerhut et al. [1997] demonstrated it for complete graphs with all edge lengths equal. Hurkens et al. [2007] showed that it holds when the graph is a ring network, as well as all graphs of at most 4 nodes. They in fact showed the stronger result that the optimal MPR solution is always a tree in these cases; their method is to explicitly construct matching primal and dual solutions to a compact LP description of the problem. Grandoni et al. [2008] gave a much shorter proof for ring networks, and their approach will be very important for us in the present work. Independently of this work, Fiorini et al. [2007] gave a proof for outerplanar graphs.

## 1.2. Contributions of This Article

Our main result is that the conjecture is true in general graphs:

**THEOREM 1.1.** *For any graph  $G = (V, E)$  with nonnegative edge costs and an associated class of hose matrices  $\mathcal{H}$ , under the SPR model, there is a minimum cost VPN which is a VPN tree for some node  $r \in V$ .*

This result also settles the complexity of the single-path VPN problem, since the optimal VPN tree can easily be computed in polynomial time by trying all possibilities for the root node (as noted in Fingerhut et al. [1997] and Gupta et al. [2001]). The proof of this result (Section 2) uses a reduction from Grandoni et al. [2008] that shows that a certain *Pyramidal Routing Conjecture* implies the VPN Conjecture. It is this *Pyramidal Routing Conjecture* that we will verify.

Given that the VPN Conjecture holds true, it would be natural to suspect that there was a fractional-implies-integral type result and that in fact the optimal cost VPN under the MPR model is also a tree. This stronger conjecture was in fact posed in Hurkens et al. [2007]. Our second main result is that the strong conjecture does not hold, and we give two simple counterexamples in Section 3.

## 1.3. Further Related Work

*Concave Costs.* Building on our result, Fiorini et al. [2010] consider the generalization of the VPN problem to concave costs. In this model, the cost of purchasing capacity on an edge is given by some concave function of the required bandwidth (rather than a linear function, as in the discussion thus far). They show that the optimal SPR solution is still a tree in this more general model.

*The Asymmetric Hose Model.* As mentioned, only the symmetric VPN problem will concern us in this article. The asymmetric version has also received much attention

however. It is APX-hard [Fingerhut et al. 1997], by reduction from the Steiner tree problem. The first theoretical results [Gupta et al. 2001; Kumar et al. 2002] gave a constant approximation to the problem of finding a best tree solution (however the optimal solution is not always a tree). The first constant factor approximation was due to Gupta et al. [2003], and was subsequently improved by Eisenbrand and Grandoni [2005], Eisenbrand et al. [2007] and Grandoni and Rothvoß [2010]. See also Rothvoß and Sanità [2009] for further work in this model.

*Other Hose Model Variants.* Eisenbrand and Happ [2006] consider a generalization of the symmetric hose model where terminals are divided into groups; in addition to the hose constraints, they disallow demands between terminals in the same group. (As motivation, they suggest a setting where terminals in the same office are already connected via an internal network.) They give a constant factor approximation algorithm for this model.

In Olver and Shepherd [2010], a natural generalization of the symmetric hose model, the *tree demand model*, is considered. The universe is determined by a capacitated tree  $T$ , distinct from the network itself, but with the leaves of  $T$  labelled by the terminal set. A demand is considered feasible if and only if it is routable on  $T$ . One easily sees that the case where  $T$  is a star corresponds exactly to the symmetric hose model. A constant factor approximation algorithm is given.

*Congestion Measure.* A *permutation routing* is a pattern of demands where each terminal sends one packet, and each terminal receives one packet. In early work on parallel computation the following packet routing problem was intensively studied: is there an oblivious routing scheme that routes any permutation routing demand with low, that is,  $O(\log n)$ , congestion?

Viewed as a demand matrix, a permutation routing looks like a collection of node-disjoint cycles. These in turn correspond to (twice) the extreme points of  $\mathcal{H}_1$ , (by which we denote the hose universe defined by unit marginals:  $b_i = 1$  for each terminal  $i$ ). Hence, the problem was to find an oblivious routing scheme to support all demands in  $\mathcal{H}_1$  with low congestion. In celebrated works, Valiant [1982] and Valiant and Brebner [1981], such a randomized oblivious scheme (randomized load balancing) was shown to exist for several popular classes of sparse networks proposed for parallel computation, in particular hypercubes. The randomization in their scheme corresponds to the oblivious templates being fractional (i.e., multipath routing); the weight given to a particular path in the template represents the probability that a packet will be routed along that path. In contrast, it is known that the desired congestion bound of  $O(\log n)$  is not achievable under deterministic (single-path) routing. Borodin and Hopcroft [1982] showed a lower bound of  $\Omega(\sqrt{n}/\Delta^{3/2})$ ,  $\Delta$  being the maximum degree (see also Kaklamani et al. [1991]).

We also mention the seminal work of Leighton and Rao [1999] on flow-cut gaps. The *uniform multiflow* demand on a subset  $X \subseteq V$  of terminals is the demand  $D_{ij} = \frac{1}{|X|}$  for all  $i, j$ . One sees that if the uniform multiflow is routable in a graph, then one can *obliviously* route any hose matrix in  $\mathcal{H}_1$  with congestion 2. Hence, conditions on the existence of uniform multiflows imply conditions on routability of hose matrices. If the uniform demand is routable in a graph  $G$ , then for any cut  $\delta(S)$  with  $S \subseteq V$ , the number of edges must be at least  $\min\{|S \cap X|, |X - S|\}$ . Hence, the graph “expands” relative to  $X$ . Leighton–Rao show that conversely, if a graph has this property, then one may route the uniform demand with congestion  $O(\log n)$  on each edge. Their result is tight.

In some sense, the VPN problem can be viewed as the counterpart to some of these past works, where total cost instead of congestion is the guiding metric.

## 2. PROOF OF THE VPN CONJECTURE

### 2.1. Definitions and Notation

We consider undirected graphs  $G = (V, E)$  with *node* set  $V$  and *edge* set  $E$ . We may abuse notation and simply refer to an edge with endpoints  $u, v$  as  $uv$ . Each edge  $e$  has an associated nonnegative cost  $c(e)$ , representing the cost per unit of bandwidth required for that edge. For any subset  $S$  of nodes, we denote by  $\delta_G(S)$  the set of edges with exactly one endpoint in  $S$  (if the context is clear we simply write  $\delta(S)$ ). For a pair of nodes  $s, t$  an  $s$ - $t$  *flow*  $f$  is (for our purposes) defined in the path formulation: for each simple path  $P$  with endpoints  $s, t$ , it assigns a value  $f(P)$ . The *size* of the flow is  $\sum_P f(P)$ ; if the value is one, we may refer to this as a *unit* flow. We normally think of simple paths as the set of edges on the path; thus, for two paths  $P, Q$ , the notation  $P \Delta Q$  denotes the symmetric difference of the two edge sets represented by paths  $P$  and  $Q$ .

A *VPN problem* instance consists of an undirected graph  $G = (V, E)$  with edge costs  $c(e)$ , a terminal set  $W \subseteq V$ , as well as the integral marginals  $b_i$  for each  $i \in W$  defining a hose universe. We use  $k := |W|$  to denote the number of terminals. As was noted in Fingerhut et al. [1997] and Hurkens et al. [2007], it is sufficient to consider unit marginals, and we assume this from now on.

**LEMMA 2.1.** [FINGERHUT ET AL. 1997; HURKENS ET AL. 2007]. *If the VPN Conjecture holds for instances with unit marginals ( $b_i = 1$  for all  $i \in W$ ), then it holds for all instances.*

We consider two variants of the VPN problem.

- (*MPR*). In the *multipath* routing problem, the solution is specified by a unit  $i$ - $j$  flow  $f_{ij}$ , for each pair  $i, j \in W$ . The routing template  $\mathcal{P}$  is defined by  $\mathcal{P} := \{f_{ij} : i, j \in W\}$ . Given a hose matrix  $D$  satisfying the marginals, the flow from  $i$  to  $j$  will be routed proportionally according to this template, that is, any  $i$ - $j$ -path  $P$  carries a flow of  $D_{ij}f_{ij}(P)$ .
- (*SPR*). In the *single-path* routing problem, the flows in the routing template are restricted to be unsplittable, that is, each pair  $i, j$  routes along a single path  $P_{ij}$ . We denote the routing template by  $\mathcal{P} = \{P_{ij} : i \neq j \in W\}$ .

We use  $\mathcal{R}_{ij}$  to denote the set of all simple  $i$ - $j$ -paths. A routing template induces a VPN solution as follows. A bandwidth requirement  $u(e)$  on each edge is chosen as small as possible but allowing all hose demands to be met; thus,

$$u(e) := \max_{D \in \mathcal{H}} \sum_{\{i,j\} \subseteq W} f_{ij}(e) D_{ij},$$

where  $f_{ij}(e) := \sum_{P \in \mathcal{R}_{ij}: e \in P} f_{ij}(P)$ . Note that the sum is over *unordered* pairs of terminals. The cost of the solution is then given by  $C_{\text{VPN}}(\mathcal{P}) := \sum_{e \in E} c(e)u(e)$ .

### 2.2. Pyramidal Routing

A *Pyramidal Routing* (PR) *problem* instance, as introduced in Grandoni et al. [2008], is defined by an undirected graph  $G$  with edge costs, and a set of  $k$  terminals  $W$ . In addition, one node  $t \in W$  is specified as the *root*. A *routing template* consists of a set  $\mathcal{P}_t$  of simple  $i$ - $t$  paths  $P_{it}$ , one for each  $i \in W \setminus \{t\}$ . Define  $l(e, \mathcal{P}_t)$  to be the total flow through edge  $e$ , that is,  $l(e, \mathcal{P}_t) := |\{i \in W : e \in P_{it}\}|$ . The bandwidth requirement  $y(e, \mathcal{P}_t)$  is instead given by the function

$$y(e, \mathcal{P}_t) := \min\{l(e, \mathcal{P}_t), k - l(e, \mathcal{P}_t)\}.$$

Note that the function on the right hand side is concave in  $l(e, \mathcal{P}_t)$ , but unusually, is not monotone. The total cost is then  $C_{\text{PR}}(\mathcal{P}_t) := \sum_{e \in E} c(e)y(e, \mathcal{P}_t)$ .

One can also define an analogous *fractional pyramidal routing problem*, where instead of paths  $P_{it}$ , the routing template consists of a set of unit  $i$ - $t$  flows  $f_{it}$ .

Grandoni et al. [2008] show that the VPN Conjecture is implied by the following conjecture:

*Pyramidal Routing Conjecture.* For every integral Pyramidal Routing instance, there exists a minimal cost solution that is a tree.

The connection between Pyramidal Routing and the VPN problem may seem strange at first. But as we show later in this section, a Pyramidal Routing solution can be interpreted as a solution to the VPN problem of a particular type (that we call *truncated hub templates*) which gives the pyramidal cost function a natural interpretation. Thus, the result of Grandoni et al. [2008] may be interpreted as saying that an optimal solution for VPN can be found amongst the truncated hub templates, each of which corresponds to a PR solution.

The crucial part of their argument is the following lemma, the proof of which we reproduce here for completeness. Their proof is an extension of one given in Gupta et al. [2001, Theorem 3.1].

LEMMA 2.2. [GRANDONI ET AL. 2008]. *Given a VPN instance, and an SPR routing template  $\mathcal{P} = \{P_{ij} : i \neq j \in W\}$ , there exists a terminal  $t \in W$  so that  $C_{\text{VPN}}(\mathcal{P}) \geq C_{\text{PR}}(\mathcal{P}_t)$ , where  $\mathcal{P}_t = \{P_{it} : i \in W \setminus \{t\}\}$ .*

PROOF. The strategy of the proof is to derive a lower bound for  $u(e)$  for each  $e$  in the instance by judiciously selecting a demand matrix, which will then give us the desired lower bound on  $C_{\text{VPN}}(\mathcal{P})$ . Fix an edge  $e$ . The choice  $D^e$  of demand matrix for  $e$  is given by

$$D_{ij}^e := \begin{cases} \frac{1}{k} \left( \frac{y(e, \mathcal{P}_i)}{l(e, \mathcal{P}_i)} + \frac{y(e, \mathcal{P}_j)}{l(e, \mathcal{P}_j)} \right) & \text{if } e \in P_{ij}, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Here,  $\mathcal{P}_j = \{P_{ij} : i \in W \setminus \{j\}\}$  for each  $j \in W$ , and depends on the SPR template  $\mathcal{P}$ .  $\square$

CLAIM 2.3.  $D^e$  is a valid hose demand matrix for all edges.

PROOF. We need to show that  $\sum_{j \in W} D_{ij}^e \leq 1$  for all  $i \in W$ . We have

$$\begin{aligned} \sum_{j \in W} D_{ij}^e &= \sum_{j \in W: e \in P_{ij}} \frac{1}{k} \left( \frac{y(e, \mathcal{P}_i)}{l(e, \mathcal{P}_i)} + \frac{y(e, \mathcal{P}_j)}{l(e, \mathcal{P}_j)} \right) \\ &\leq \sum_{j \in W: e \in P_{ij}} \frac{1}{k} \left( \frac{k - l(e, \mathcal{P}_i)}{l(e, \mathcal{P}_i)} + \frac{l(e, \mathcal{P}_j)}{l(e, \mathcal{P}_j)} \right) \\ &= \sum_{j \in W: e \in P_{ij}} \frac{1}{k} \cdot \frac{k}{l(e, \mathcal{P}_i)} \\ &= 1. \end{aligned} \quad \square$$

CLAIM 2.4. *For every edge  $e$ , we have*

$$u(e) \geq \frac{1}{k} \sum_{i \in W} y(e, \mathcal{P}_i). \quad (3)$$

PROOF. The claim follows from the definitions of  $D_{ij}^e$  and  $l(e, \mathcal{P}_i) = |\{j \in W : e \in P_{ij}\}|$ .

$$\begin{aligned} u(e) &\geq \sum_{\{i,j\}:e \in P_{ij}} \frac{1}{k} \left( \frac{y(e, \mathcal{P}_i)}{l(e, \mathcal{P}_i)} + \frac{y(e, \mathcal{P}_j)}{l(e, \mathcal{P}_j)} \right) \\ &= \sum_{i \in W} \sum_{j \in W: e \in P_{ij}} \frac{1}{k} \cdot \frac{y(e, \mathcal{P}_i)}{l(e, \mathcal{P}_i)} \\ &= \frac{1}{k} \sum_{i \in W} y(e, \mathcal{P}_i). \quad \square \end{aligned}$$

The theorem now follows by multiplying the inequality in (3) by  $c(e)$  and summing over all  $e \in E$ :

$$\begin{aligned} \sum_{e \in E} c(e)u(e) &\geq \sum_{e \in E} c(e) \frac{1}{k} \sum_{i \in W} y(e, \mathcal{P}_i) \\ &= \frac{1}{k} \sum_{i \in W} \sum_{e \in E} c(e)y(e, \mathcal{P}_i) \\ &\geq \min_{i \in W} \sum_{e \in E} c(e)y(e, \mathcal{P}_i). \end{aligned}$$

We have seen that the VPN cost is lower bounded by some Pyramidal Routing cost. In fact, a converse result holds too: the cost of optimal SPR solutions can also be upper bounded by the cost of an associated Pyramidal Routing problem. To do this, for each solution  $\mathcal{P}_t$  to a PR problem instance with root  $t$ , we define an oblivious routing template, called the *truncated hub template* associated with  $\mathcal{P}_t$ . This is defined as the template  $\mathcal{Q} = \{Q_{ij} : i, j \in W\}$ , where  $Q_{ij}$  is any  $i$ - $j$ -path in the component of  $P_{it} \Delta P_{jt}$  (symmetric difference) containing  $i$  and  $j$ . Note that since  $i$  and  $j$  are the only odd-degree nodes in  $P_{it} \Delta P_{jt}$ , they will indeed be in the same component.

Figure 1 shows two examples of how the truncated hub template is constructed from the pyramidal routing template. Note that while intuitively the truncated hub template is obtained by “shortcutting” the pyramidal routing, this is not always the case, as is shown in the second example.

LEMMA 2.5. *Given a solution  $\mathcal{P}_t$  to a PR problem instance with root  $t$ , the capacity on any edge  $e$  required by its truncated hub template  $\mathcal{Q}$  is at most  $y(e, \mathcal{P}_i)$ . In particular,  $C_{\text{VPN}}(\mathcal{Q}) \leq C_{\text{PR}}(\mathcal{P}_t)$ .*

PROOF. Let  $D \in \mathcal{H}$  be any valid demand matrix. Consider any edge  $e$  and define the set of nodes which route through  $e$  by  $R_e := \{i \in W : e \in P_{it}\}$ ; note that  $l(e, \mathcal{P}_t) = |R_e|$ .



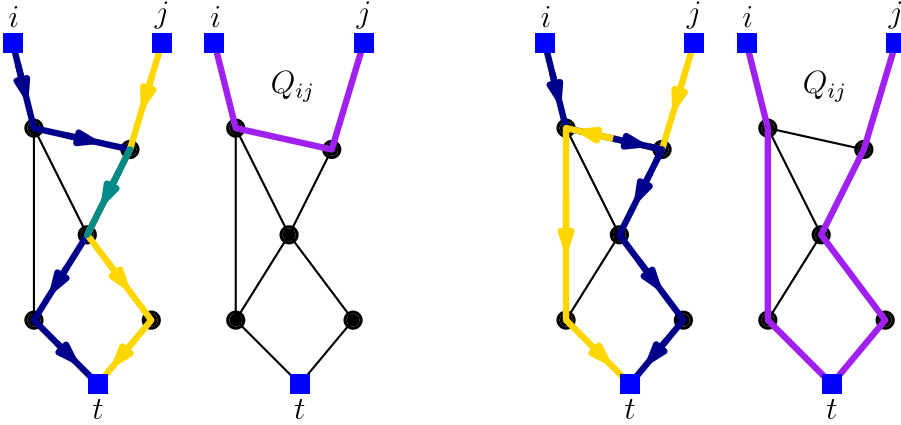


Fig. 1. Two examples of how the truncated hub template routing  $Q_{ij}$  is constructed from the pyramidal routing template.

Now notice that if we have a pair  $i, j \in W$  where  $e \in Q_{ij}$ , then exactly one of  $i$  and  $j$  is in  $R_e$ , because of the symmetric difference construction. So we have

$$\begin{aligned}
 \sum_{\{i,j\}:e \in Q_{ij}} D_{ij} &\leq \sum_{i \in R_e} \sum_{j \notin R_e} D_{ij} \\
 &\leq \sum_{i \in R_e} \sum_{j \in W} D_{ij} \\
 &\leq \sum_{i \in R_e} 1 \\
 &= l(e, \mathcal{P}_t).
 \end{aligned}$$

Similarly,

$$\sum_{\{i,j\}:e \in Q_{ij}} D_{ij} \leq \sum_{j \notin R_e} \sum_{i \in R_e} D_{ij} \leq \sum_{j \notin R_e} 1 = k - l(e, \mathcal{P}_t).$$

Thus, we have that

$$\sum_{\{i,j\}:e \in Q_{ij}} D_{ij} \leq \min\{l(e, \mathcal{P}_t), k - l(e, \mathcal{P}_t)\} = y(e, \mathcal{P}_t).$$

But then the required capacity on edge  $e$  is

$$u(e) = \max_{D \in \mathcal{H}} \sum_{\{i,j\}:e \in Q_{ij}} D_{ij} \leq y(e, \mathcal{P}_t),$$

as required. □

Note that by the proof of Lemma 2.2 the optimal SPR cost is at least a weighted average of costs  $C_{PR}(\mathcal{P}_t)$ . The preceding lemma shows that it is also at most the cost of any given optimal PR solution. Thus, we have the following new result. It effectively shows that, in order to find an optimal VPN, one may restrict to truncated hub-routing templates.

**THEOREM 2.6.** *For any pair  $t, t' \in W$ , the optimal solutions to the PR problems with root  $t$  and  $t'$  are the same, and have the same value as the optimal solution of the associated SPR problem. Hence, the PR conjecture and the VPN Conjecture are equivalent.*

It should be noted, however, that no such correspondence holds between the MPR problem and the fractional PR problem. As we prove later, the optimal MPR solution may be cheaper than the optimal SPR solution in some cases; but note the following.

**LEMMA 2.7.** *There exists an optimal solution to a fractional Pyramidal Routing instance that is integral, that is,  $f_{it}$  induces an  $i$ - $t$ -path for all  $i \in W \setminus \{t\}$ .*

**PROOF.** We show this by proving that the problem consists of minimizing a concave function over a 0-1 polytope  $\mathcal{B}$ . A routing template  $\mathcal{P}_t = \{f_{it} : i \in W\}$  induces a vector  $\mathbf{x}^{\mathcal{P}_t}$  in  $\mathbb{R}^{k \times |E|}$ , by setting  $x_{i,e}^{\mathcal{P}_t} = f_{it}(e)$ . Let the polytope  $\mathcal{B} \subset \mathbb{R}^{k \times |E|}$  be the set of vectors induced by feasible solution templates. By the max-flow min-cut theorem, the extreme points of  $\mathcal{B}$  are 0-1 vectors.

The objective function is  $C_{\text{PR}}(\mathcal{P}) = \sum_{e \in E} c(e)y(e, \mathcal{P}_t)$ . As noted before,  $y(e, \mathcal{P}_t)$  is a concave function in the load on  $e$ . Since the load is a linear function of the routing template (thought of as a vector  $\mathbf{f} = (f_{it})_{i \in W}$  of flows),  $y(e, \mathcal{P}_t)$  is concave over  $\mathcal{B}$ ; since the sum of concave functions remains concave,  $C_{\text{PR}}(\mathcal{P})$  is too. It is well known that a minimizer of a concave objective always exists at a vertex of the polytope, which corresponds to an integral routing template.  $\square$

### 2.3. A Reduction to $T$ -Joins

We begin with an instance of the Pyramidal Routing problem, with root  $t$ . Let  $\mathcal{P}_t$  be a routing template for this instance.

*Definition 2.8.* Call an edge  $e \in E$  *heavy* if  $l(e, \mathcal{P}_t) \geq k/2$ .

Let  $H$  be the set of heavy edges determined by  $\mathcal{P}_t$ . Note that

$$y(e) = \begin{cases} l(e, \mathcal{P}_t) & \text{if } e \notin H, \\ k - l(e, \mathcal{P}_t) & \text{if } e \in H. \end{cases}$$

Recall that for any  $T \subseteq V$  with  $|T|$  is even, a  $T$ -join of  $G$  is a set  $J$  of edges such that the odd degree nodes in the subgraph defined by  $J$  is precisely  $T$ .

Let  $T'$  be the set of odd degree vertices in the subgraph induced by  $H$ . Now define  $T := T' \Delta \{t\}$ , and  $T_u := T \Delta \{u\}$  for all  $u \in W$ . Note that  $|T'|$  is even, so  $|T|$  is odd, and so  $|T_u|$  is even for all  $u \in W$ . Let  $M_u$  be the minimum cost  $T_u$ -join on  $G$ .  $C(M_u) := \sum_{e \in M_u} c(e)$  is defined to be the cost of  $M_u$ . Define  $C'(u) := C(P_{ut} \Delta H)$ . This we think of as being  $u$ 's contribution to the total cost of the Pyramidal Routing. Notice that  $u$  pays for light edges on its path, and heavy edges not on its path. We also have

$$\begin{aligned} \sum_{u \in W} C'(u) &= \sum_{u \in W} \sum_{e \in P_{ut} \Delta H} c(e) \\ &= \sum_{e \in E \setminus H} l(e, \mathcal{P}_t) \cdot c(e) + \sum_{e \in H} (k - l(e, \mathcal{P}_t))c(e) \\ &= C_{\text{PR}}(\mathcal{P}_t). \end{aligned}$$

So this really is a division of the total cost between the terminals.

LEMMA 2.9. *A lowerbound for the cost  $C_{\text{PR}}(\mathcal{P}_t)$  of solution  $\mathcal{P}_t$  is  $\sum_{u \in W} C(M_u)$ .*

PROOF. We want to show  $C'(u) \geq C(M_u)$  for all  $u \in W$ . Consider the symmetric difference  $H_u := P_{ut} \Delta H$ . Since  $P_{ut}$  has even degree at every node except  $u$  and  $t$ , and  $H$  is a  $T'$ -join, it follows that  $H_u$  is a  $T_u$ -join. So  $C(H_u) \geq C(M_u)$ . But by definition,  $C'(u) = C(H_u)$ . Thus,

$$C_{\text{PR}}(\mathcal{P}_t) = \sum_{u \in W} C'(u) = \sum_{u \in W} C(H_u) \geq \sum_{u \in W} C(M_u). \quad \square$$

Note that the right-hand side of this inequality depends on  $H$  only via  $T$ . We also have the following pleasant result. Define the truncated template  $\mathcal{Q} = \{Q_{uv} : u, v \in W\}$ , where  $Q_{uv}$  is any  $u$ - $v$  path contained within the component of  $M_u \Delta M_v$  containing  $u$  and  $v$ . Such a component must exist, because  $u$  and  $v$  are the only odd-degree nodes in  $M_u \Delta M_v$ . Then we have the following.

THEOREM 2.10. *The truncated template  $\mathcal{Q}$  satisfies*

$$C_{\text{VPN}}(\mathcal{Q}) \leq \sum_{u \in W} C(M_u).$$

The proof is very similar to the proof of Lemma 2.5, and so we omit it.

Note that if we had  $T = \{v\}$ , then  $M_u$  is a shortest path from  $u$  to  $v$ . So if  $v$  is the centre of the minimum cost VPN tree,  $\sum_{u \in W} C(M_u)$  is exactly the cost of the optimal tree solution. In the next section, we show that no other choice of  $T$  improves upon this.

#### 2.4. A $T$ -Join Inequality

For each node  $v$  in  $G$ , define  $C_{\text{SP}}(v)$  to be the cost of the VPN tree from the terminals to  $v$ , that is,

$$C_{\text{SP}}(v) := \sum_{u \in W} SP(u, v),$$

where  $SP(u, v)$  is the cost of the shortest path (with respect to  $c(e)$ ) between  $u$  and  $v$  in  $G$ .

We prove the following inequality, which in turn proves the Pyramidal Routing Conjecture by the reduction in the previous section.

THEOREM 2.11. *Consider a VPN instance given by  $G(V, E)$ , nonnegative cost function  $c$  on edges, and terminals  $W \subseteq V$ . Fix any terminal  $t \in W$  and any subset  $T \subseteq V$  of odd cardinality, and let  $M_u$  be a  $T_u$ -join for each  $u \in W$ , where  $T_u := T \Delta \{u\}$ . Then, there exists a node  $v \in V$  so that*

$$\sum_{u \in W} C(M_u) \geq C_{\text{SP}}(v).$$

In fact, we prove the following slightly stronger theorem.

THEOREM 2.12. *With the setup of the previous theorem, let  $F$  be the multigraph obtained by taking the disjoint union of the  $M_u$ 's. Then there exists a node  $v \in V$  so that there are edge-disjoint paths from all the vertices in  $W$  to  $v$ .*

Theorem 2.11 shows that the weight of the edges in the multiset  $\cup_u M_u$  is at least the weight of the edges in  $C_{\text{SP}}(v)$ . Theorem 2.12 shows that in fact the latter set is contained in the former one, immediately implying Theorem 2.11.

We provide two proofs of this theorem. The first was given in an earlier version of this paper, and is completely elementary and self-contained. The second, subsequently

communicated to us by András Sebő [2008, private communication], is somewhat shorter, but makes use of the properties of Gomory–Hu trees.

We fix some notation before getting into the proofs. Call a set  $S \subseteq V$  *T-even* (respectively *T-odd*) if  $|S \cap T|$  is even (respectively, odd). Note that since  $|T|$  is odd, exactly one of  $S$  and  $V \setminus S$  is *T-even* for any  $S \subseteq V$ .

*2.4.1. First proof of Theorem 2.12.* We need the following lemma.

**LEMMA 2.13.** *For any set  $S \subseteq V$  which is *T-even*,  $|\delta_F(S)| \geq |S \cap W|$ .*

**PROOF.** Consider any  $u \in S \cap W$ . Since  $S$  is *T-even*, it is  $T_u$ -odd. Thus, since  $M_u$  is a  $T_u$ -join, it must cross  $S$ . The lemma now follows from the definition of  $F$  as the disjoint union of the  $M_u$ 's.  $\square$

**PROOF OF THEOREM 2.12.** Construct the graph  $F'$  from  $F$  by adding an extra node  $s$ , and edges  $su$  for all  $u \in W$ . The statement of the theorem is equivalent to showing that there exists a node  $v$  so that there is an  $s$ - $v$  flow of size  $k$  in  $F'$ , taking all the edges to have unit capacity.

Define  $D_z$  to be the side of a minimum  $s$ - $z$  cut containing  $z \in V$ . Suppose for a contradiction that  $|\delta_{F'}(D_z)| < k$  for all  $z \in V$ , since otherwise we have a valid routing by the max-flow min-cut theorem. Since  $|\delta_{F'}(D_z)| = |\delta_F(D_z)| + |D_z \cap W|$ , this gives

$$|\delta_F(D_z)| < k - |D_z \cap W| \quad \forall z \in V. \quad (4)$$

Note that  $D_z$  is *T-even*; otherwise  $V \setminus D_z$  is *T-even*, and so by Lemma 2.13,

$$|\delta_F(D_z)| = |\delta_F(V \setminus D_z)| \geq |W \setminus D_z| = k - |D_z \cap W|,$$

which contradicts (4).

It is enough to show that arbitrary intersections of the  $D_u$ 's are *T-even*. To see this, note that it implies that  $\cup_{u \in V} D_u$  is *T-even*, by the inclusion-exclusion principle. But  $u \in D_u$ , so  $\cup_{u \in V} D_u = V$ , which is *T-odd*. This contradiction implies the result.

Hence, it remains to prove that intersections of  $D_u$ 's are always *T-even*. Suppose inductively that all intersections of less than  $l$  sets, each of type  $D_z$  for some  $z \in V$ , are *T-even*. This is true for  $l = 2$ , since  $D_z$  is *T-even* for all  $z \in V$ . Now suppose for a contradiction that for some arbitrary collection  $D_1, \dots, D_l$ , the set  $D_1 \cap D_2 \cap \dots \cap D_l$  is *T-odd*. Let  $D'_i = D_i \setminus (\cup_{j \neq i} D_j)$ . Claim 2.14 follows from the inclusion-exclusion principle, and our inductive assumption.

**CLAIM 2.14.** *Under the assumptions in the preceding paragraph,  $D'_i$  is *T-odd* for all  $i$ .*

**PROOF.** Note that  $D'_i = D_i \setminus (\cup_{j \neq i} (D_j \cap D_i))$ . Now by the inclusion-exclusion principle, we have

$$\begin{aligned} \left| \bigcup_{j \neq i} (D_j \cap D_i) \cap T \right| &= \sum_{j \neq i} |(D_j \cap D_i) \cap T| - \sum_{j_1 < j_2: j_1, j_2 \neq i} |(D_{j_1} \cap D_{j_2} \cap D_i) \cap T| + \dots \\ &\quad + (-1)^{l-1} |(D_1 \cap D_2 \cap \dots \cap D_l) \cap T|. \end{aligned}$$

The last term on the right is odd and the rest are even by our assumptions, and thus  $\cup_{j \neq i} (D_j \cap D_i)$  is *T-odd*; since  $D_i$  is *T-even*, it follows that  $D'_i$  is *T-odd*.  $\square$

CLAIM 2.15.

$$\sum_{i=1}^l |\delta_{F'}(D_i)| \geq \sum_{i=1}^l |\delta_{F'}(D'_i)|$$

PROOF. Consider any edge  $e$  that contributes to the right-hand side. If it has endpoints in  $D'_i$  and  $D'_j$  for some  $i \neq j$ , then it will contribute twice to the right-hand side; such an edge will also contribute at least (in fact, precisely) twice to the left-hand side, in  $\delta_{F'}(D_i)$  and  $\delta_{F'}(D_j)$ . If  $e$  has an endpoint in  $D'_i$  only, and not in any other  $D'_j$ , then it is counted once on the right-hand side, and at least once on the left-hand side.  $\square$

Now we have

$$\begin{aligned} \sum_{i=1}^l |\delta_{F'}(D_i)| &\geq \sum_{i=1}^l |\delta_{F'}(D'_i)| && \text{by Claim 2.15} \\ &\geq \sum_{i=1}^l (k - |D'_i \cap W|) && \text{as } D'_i \text{ is } T\text{-odd} \\ &\geq \sum_{i=1}^l (k - |D_i \cap W|) && \text{as } D'_i \subseteq D_i. \end{aligned}$$

But this is a contradiction because (4) implies

$$\sum_{i=1}^l |\delta_{F'}(D_i)| < \sum_{i=1}^l (k - |D_i \cap W|).$$

So our assumption that  $D_1 \cap \dots \cap D_l$  is  $T$ -odd was incorrect, and the proof is complete.  $\square$

2.4.2. *A Shorter Proof.* The following is a somewhat shorter proof of Theorem 2.12, due to András Sebő [2008, private communication]. We actually prove the following theorem, which in turn implies Theorem 2.12.

**THEOREM 2.16.** *Let  $K$  be the complete graph on node set  $V$ , and let  $T \subseteq V$  be of even size. Let  $H$  be the edge-disjoint union of any  $k$   $T$ -joins of  $K$ . Then for any node  $s \in T$ , there is another node  $v \in V$  such that there are  $k$  edge-disjoint  $s$ - $v$ -paths in  $H$ .*

Note that taking the underlying graph be the complete graph allows for any set of  $T$ -joins, and thus makes the theorem more general. Before proving the theorem, let us see why it implies Theorem 2.12. Construct  $F'$  as described in the earlier proof of Theorem 2.12, by adding a new node  $s$  connected to all the terminals; Theorem 2.12 asserts that there is a node  $v \in V$  such that there are  $|W|$  edge-disjoint  $s$ - $v$ -paths in  $F'$ .

Let  $M'_u := M_u \cup \{us\}$  for all  $u \in W$ . By construction,  $F'$  is the disjoint union of  $M'_u$  for  $u \in W$ . But  $M'_u$  is a  $T_s$ -join, since  $M_u$  is a  $T_u$ -join and

$$T_u \Delta \{u\} \Delta \{s\} = T \Delta \{s\} = T_s.$$

Thus,  $F'$  is the disjoint union of  $k$  different  $T_s$ -joins. Now applying Theorem 2.16 with the multigraph  $F'$  playing the role of  $H$  and  $T_s$  playing the role of  $T$  gives the conclusion of Theorem 2.12.

**PROOF OF THEOREM 2.16.** By the max-flow min-cut theorem it suffices to prove that there is a node  $v$  such that the minimum  $s$ - $v$  cut in  $H$  has size at least  $k$ . Recall

that a  $T$ -cut is a set  $\delta(S)$  for some  $S \subseteq V$  such that  $|S \cap T|$  is odd. A basic property of  $T$ -cuts and  $T$ -joins is that every  $T$ -cut intersects every  $T$ -join (see, e.g., Cook et al. [1997]).

Observe that for any  $T$ -cut  $S$  in  $H$ , we have  $|\delta_H(S)| \geq k$ . This is because  $H$  is the disjoint union of  $k$   $T$ -joins, and each of them must intersect  $\delta_H(S)$ .

Now consider a Gomory–Hu tree  $\mathcal{T}$  of  $H$  (see, e.g., Cook et al. [1997] for definition and properties of these trees). Removing  $s$  from  $\mathcal{T}$  breaks it into connected components. At least one of these connected components must be  $T$ -odd because  $|T|$  is even and  $s \in T$ ; call this component  $C$ . Let  $sv$  be the edge connecting  $s$  to  $C$ . Then the cut defined by the edge  $sv$  and  $\mathcal{T}$  is a  $T$ -cut and hence has size at least  $k$ . But since it is also a cut given by the Gomory–Hu tree  $\mathcal{T}$  it is a minimum  $s$ - $v$  cut. Thus, we have shown that the minimum cut between  $s$  and  $v$  has size at least  $k$ , and so there are  $k$  edge-disjoint  $s$ - $v$ -paths.  $\square$

### 3. THE MULTIPATH VPN CONJECTURE

It has been conjectured [Hurkens et al. 2007] (cf. Erlebach and Rüegg [2004]) that the multipath version of the VPN problem, where routing templates may be fractional, also has an optimal solution in the form of a tree. In this section, we show that this conjecture is false. Moreover, we will give an example that shows that the ratio between the integral and fractional optima can be as large as  $9/8$ .

A similar phenomenon occurs more forcefully in the context of permutation packet routing with the congestion measure, as discussed in Section 1.3. There, the congestion can be dramatically worse for deterministic routing in comparison with randomized (essentially, fractional) routing.

Before describing the counterexamples, it will be useful to introduce the compact LP formulation for the multipath VPN problem, as derived in Altın et al. [2007] and Hurkens et al. [2007]:

$$\begin{aligned} \min \quad & \sum_{e \in E} \sum_{i \in W} c(e) y_i(e) \\ \text{such that} \quad & y_i(e) + y_j(e) \geq f_{ij}(e) & \forall e \in E, \{i, j\} \subseteq W, \\ & f_{ij} \text{ is a unit } i\text{-}j\text{-flow} & \forall \{i, j\} \subseteq W \\ & y_i(e) \geq 0 & \forall i \in W, e \in E \end{aligned}$$

The flows  $f_{ij}$  are defined on the bidirection of  $G$ , and we use  $f_{ij}(e)$  to refer to the amount of flow on edge  $e$ , irrespective of direction.

This formulation yields a convenient description of the capacity reservation of a solution as a cost sharing between the terminals. We will think of  $y_i$  as the capacity paid for by terminal  $i$ . We see that  $\mathbf{y} := \{y_i : i \in W\}$  is a valid capacity reservation if and only if for every pair  $i \neq j \in W$ ,  $y_i + y_j$  supports a unit  $i$ - $j$ -flow. Thus, we may think of  $\mathbf{y}$  as specifying a MPR solution, of cost

$$C(\mathbf{y}) = \sum_{i \in W} C(y_i) = \sum_{e \in E} \sum_{i \in W} c(e) y_i(e).$$

We begin with the smallest counterexample we know of (and likely the smallest possible). It consists of only 7 nodes, and is shown in Figure 2(a); edge lengths are indicated. Half a unit of capacity is bought on each edge; the colors show which terminal pays for each edge, with the edges paid for by terminal 1 shown in bold. Since  $y_i + y_j$  has enough capacity for a unit  $i$ - $j$  flow for every  $i, j \in W$ , this is indeed a feasible MPR solution of cost  $15/2$ . The cheapest shortest path tree on the other hand

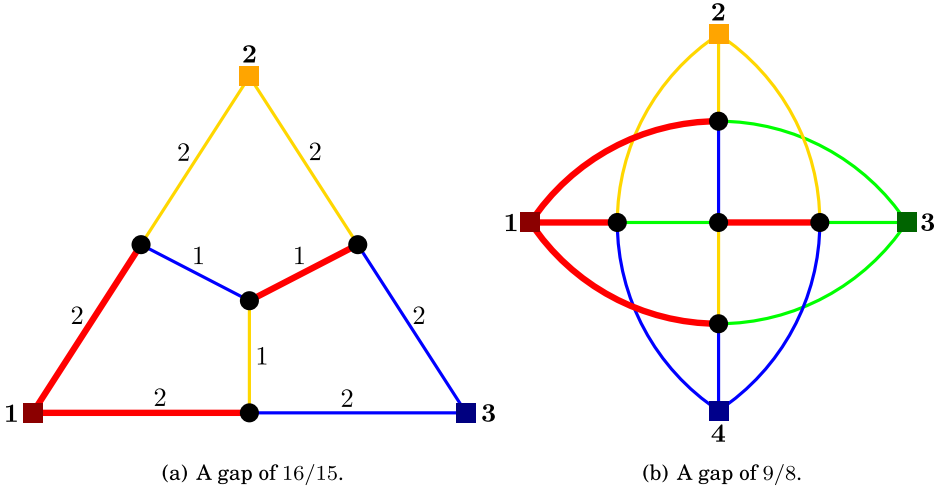


Fig. 2. Two small gap examples. In each, the edges paid for by terminal 1 (red) are shown in bold.

(found just by checking all possible choices of root) has cost 8. By Theorem 1.1, this is the cost of the optimal SPR solution, and so we have a counterexample. The ratio between  $OPT_{SPR}$  and  $OPT_{MPR}$  in this example is  $16/15$ .

This example also shows that the multipath VPN Conjecture is false even for planar graphs.

We can obtain a slightly larger gap with a slightly larger construction, using the same idea. In Figure 2(b), all edges have cost 1. For the MPR solution,  $1/3$  capacity is bought on each edge; the edges paid for by terminal 1 (red) are shown in bold, and the rest are symmetric. This is again a feasible MPR solution, of cost  $16/3$ . The cheapest tree solution on the other hand costs 6; this gives a ratio  $OPT_{SPR}/OPT_{MPR} = 9/8$ .

#### 4. CONCLUSION

We have seen that the ratio between the single-path and multipath optima can be as large as  $9/8$ . Conversely, it is explicit in Gupta et al. [2001] and implicit in Fingerhut et al. [1997] that the gap between MPR and SPR is at most 2. We suspect that the integrality gap is significantly smaller than 2, but this is currently an open question.

Having resolved the VPN Conjecture, we would like to conclude by suggesting a new conjecture to take its place. Recall the tree demand problem discussed in our section on related work: for a given capacitated tree  $T$ , with leaves labelled by the terminals, a demand is considered feasible if and only if it is routable in  $T$ . In Olver and Shepherd [2010], we suggest a natural algorithm for the problem: it finds the cheapest solution in the form of a *hierarchical hubbing*. A hierarchical hubbing solution for  $T$  is defined by some mapping  $\phi : V(T) \rightarrow V(G)$ , satisfying  $\phi(i) = i$  for all  $i \in W$ . An edge  $uv \in E(T)$  is then mapped to a shortest path between  $\phi(u)$  and  $\phi(v)$ . Finally, this determines a routing template by setting  $P_{ij}$  to the image of the unique  $i$ - $j$ -path in  $T$  under the mapping  $\phi$ . Notice that in the case where  $T$  is a star, this reduces to a shortest path tree solution. The resulting solution is “tree-like” in the sense that it preserves the structure of the tree  $T$  in the solution, even though the support of the solution need no longer be a tree in general.

*Generalized VPN Conjecture.* For any instance of the tree demand problem, for a given capacitated tree  $T$ , there is an optimal SPR solution in the form of a hierarchical hubbing.

Since the optimal hierarchical hubbing solution can be found in polynomial time by means of a dynamic program, this would make the tree demand problem polynomially solvable. The interested reader may refer to Olver and Shepherd [2010] for more details on this problem.

## ACKNOWLEDGMENTS

We are grateful to the Grandoni et al. [2008] who provided a preprint of their paper. Their ideas substantially altered our approach to the VPN problem. We are also grateful to András Sebó for communicating his short proof of Theorem 2.12. Finally, we are grateful to the reviewers for their insightful comments on several matters of fact and taste.

## REFERENCES

- Altun, A., Amaldi, E., Belotti, P., and Pinar, M. C. 2007. Provisioning virtual private networks under traffic uncertainty. *Networks* 49, 1, 100–115.
- Ben-Ameur, W. and Kerivin, H. 2003. New economical virtual private networks. *Comm. ACM* 46, 6, 69–73.
- Ben-Ameur, W. and Kerivin, H. 2005. Routing of uncertain traffic demands. *Optim. Eng.* 6, 3, 283–313.
- Ben-Tal, A. and Nemirovski, A. 1998. Robust convex optimization. *Math. Oper. Res.* 23, 4, 769–805.
- Ben-Tal, A. and Nemirovski, A. 1999. Robust solutions of uncertain linear programs. *Oper. Res. Lett.* 25, 1, 1–13.
- Ben-Tal, A., El Ghaoui, L., and Nemirovski, A. 2009. *Robust Optimization*. Princeton Series in Applied Mathematics. Princeton University Press.
- Bertsimas, D. and Sim, M. 2003. Robust discrete optimization and network flows. *Math. Prog.* 98, 1, 49–71.
- Borodin, A. and Hopcroft, J. E. 1982. Routing, merging and sorting on parallel models of computation. In *Proceedings of the 14th Annual ACM Symposium on Theory of Computing (STOC)*. ACM, New York, 338–344.
- Chekuri, C., Oriolo, G., Scutellà, M. G., and Shepherd, F. B. 2007. Hardness of robust network design. *Networks* 50, 1, 50–54.
- Cook, W. J., Cunningham, W. H., Pulleyblank, W. R., and Schrijver, A. 1997. *Combinatorial Optimization*, 1st Ed. Wiley-Interscience.
- Duffield, N. G., Goyal, P., Greenberg, A., Mishra, P., Ramakrishnan, K. K., and van der Merwe, J. E. 1999. A flexible model for resource management in virtual private networks. In *Proceedings of SIGCOMM*. ACM, New York, 95–108.
- Eisenbrand, F. and Grandoni, F. 2005. An improved approximation algorithm for virtual private network design. In *Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM, Philadelphia, PA, 928–932.
- Eisenbrand, F. and Happ, E. 2006. Provisioning a virtual private network under the presence of non-communicating groups. In *Proceedings of the 6th International Conference on Algorithms and Complexity*. Lecture Notes in Computer Science, vol. 3998, Springer, 105–114.
- Eisenbrand, F., Grandoni, F., Oriolo, G., and Skutella, M. 2007. New approaches for virtual private network design. *SIAM J. Comput.* 37, 3, 706–721.
- El Ghaoui, L. and Lebret, H. 1997. Robust solutions to least-squares problems with uncertain data. *SIAM J. Matrix Anal. Appl.* 18, 1035–1064.
- El Ghaoui, L., Oustry, F., and Lebret, H. 1998. Robust solutions to uncertain semidefinite programs. *SIAM J. Optim.* 9, 1, 33–52.
- Erlebach, T. and Rüegg, M. 2004. Optimal bandwidth reservation in hose-model VPNs with multi-path routing. In *Proceedings of INFOCOM*. 2275–2282.
- Fingerhut, J. A., Suri, S., and Turner, J. S. 1997. Designing least-cost nonblocking broadband networks. *J. Algor.* 24, 2, 287–309.
- Fiorini, S., Oriolo, G., Sanità, L., and Theis, D. O. 2007. The VPN tree routing conjecture for outerplanar networks. arXiv:0711.2623v3 [math.OC].
- Fiorini, S., Oriolo, G., Sanità, L., and Theis, D. O. 2010. The VPN problem with concave costs. *SIAM J. Disc. Math.* 24, 3, 1080–1090.
- Grandoni, F., Kaibel, V., Oriolo, G., and Skutella, M. 2008. A short proof of the VPN tree routing conjecture on ring networks. *Oper. Res. Lett.* 36, 3, 361–365.



- Grandoni, F. and Rothvoß, T. 2010. Network design via core detouring for problems without a core. In *Proceedings of the 37th International Colloquium on Automata, Languages and Programming (ICALP)*. Lecture Notes in Computer Science, vol. 6198, 490–502.
- Gupta, A., Kleinberg, J., Kumar, A., Rastogi, R., and Yener, B. 2001. Provisioning a virtual private network: A network design problem for multicommodity flow. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing (STOC)*. ACM Press, New York, 389–398.
- Gupta, A., Kumar, A., and Roughgarden, T. 2003. Simpler and better approximation algorithms for network design. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing (STOC)*. ACM, New York, 365–372.
- Hurkens, C. A. J., Keijsper, J. C. M., and Stougie, L. 2007. Virtual private network design: A proof of the tree routing conjecture on ring networks. *SIAM J. Disc. Math.* 21, 2, 482–503.
- Italiano, G., Leonardi, S., and Oriolo, G. 2006. Design of trees in the hose model: The balanced case. *Oper. Res. Lett.* 34, 6, 601–606.
- Kaklamanis, C., Krizanc, D., and Tsantilas, T. 1991. Tight bounds for oblivious routing in the hypercube. *Theory Comput. Syst.* 24, 1, 223–232.
- Kouvelis, P. and Yu, G. 1997. *Robust Discrete Optimization and Its Applications (Nonconvex Optimization and Its Applications)*, 1st Ed. Springer.
- Kumar, A., Rastogi, R., Silberschatz, A., and Yener, B. 2002. Algorithms for provisioning virtual private networks in the hose model. *IEEE/ACM Trans. Netw.* 10, 4, 565–578.
- Leighton, T. and Rao, S. 1999. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *J. ACM* 46, 6, 787–832.
- Olver, N. and Shepherd, F. B. 2010. Approximability of robust network design. In *Proceedings of the 21st ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1097–1105.
- Räcke, H. 2002. Minimizing congestion in general networks. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 43–52.
- Räcke, H. 2008. Optimal hierarchical decompositions for congestion minimization in networks. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing (STOC)*. ACM, New York, 255–264.
- Rothvoß, T. and Sanità, L. 2009. On the complexity of the asymmetric VPN problem. In *Proceedings of the 12th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*. 326–338.
- Valiant, L. G. 1982. A scheme for fast parallel communication. *SIAM J. Comput.* 11, 2, 350–361.
- Valiant, L. G. and Brebner, G. J. 1981. Universal schemes for parallel communication. In *Proceedings of the 13th Annual ACM Symposium on Theory of Computing (STOC)*. ACM, New York, 263–277.

Received December 2010; revised January 2012; accepted March 2013