

Approximation Algorithms for Disjoint Paths and Related Routing and Packing Problems*

Alok Baveja[†]

Aravind Srinivasan[‡]

Abstract. Given a network and a set of connection requests on it, we consider the maximum edge-disjoint paths and related generalizations and routing problems that arise in assigning paths for these requests. We present improved approximation algorithms and/or integrality gaps for all problems considered; the central theme of this work is the underlying multi-commodity flow relaxation. Applications of these techniques to approximating families of packing integer programs are also presented.

Key words and phrases. Disjoint paths, approximation algorithms, unsplittable flow, routing, packing, integer programming, multicommodity flow, randomized algorithms, rounding, linear programming.

1 Introduction

Due to the emergence and anticipated rapid growth of high-speed integrated networks that provide vast bandwidth (e.g., optical networks) and support heterogeneous applications, considerable attention has been paid recently to edge-disjoint paths, bandwidth allocation, and related algorithmic problems on large-scale networks: see Kleinberg's thesis for background, motivation, and several new results (Kleinberg (1996)). Edge-disjoint paths and related path selection/routing problems have been studied intensely, broadly using two approaches: random walks (e.g., Peleg and Upfal (1989), Broder, Frieze and Upfal (1994), Broder, Frieze and Upfal (1999)), and multi-commodity

*A preliminary version of part of this work appeared in the *Proc. IEEE Symposium on Foundations of Computer Science*, pages 416–425, 1997.

[†]School of Business, Rutgers University, Camden, NJ 08102, USA. E-mail: baveja@crab.rutgers.edu

[‡]Bell Laboratories, Lucent Technologies, 700 Mountain Ave., Murray Hill, NJ 07974-0636, USA. Part of this work was done when the author was with the Dept. of Information Systems & Computer Science, National University of Singapore, Singapore 119260. Supported in part by National University of Singapore Academic Research Fund Grants RP960620 and RP970607. E-mail: srin@research.bell-labs.com.

flow (e.g., Raghavan and Thompson (1987), Leighton and Rao (1988), Leighton and Rao (1996), Kleinberg and Rubinfeld (1996)). By taking the latter approach, we present improved approximation algorithms and bounds for edge-disjoint paths and related problems.

We first recall the notion of an approximation algorithm, a particularly useful approach for computationally hard problems. Given an optimization problem \mathcal{P} and an instance I of \mathcal{P} , let $OPT(I)$ denote the optimal objective-function value for I ; each feasible solution for I will have a non-negative objective-function value, for all problems \mathcal{P} studied here. If \mathcal{P} is a *maximization* problem, an approximation algorithm \mathcal{A} for \mathcal{P} is an efficient (which for our purposes we shall take to be polynomial-time) algorithm that, for some $\lambda \geq 1$, produces a feasible solution of value at least $OPT(I)/\lambda$ for all instances I . \mathcal{A} is called a λ -*approximation algorithm* for \mathcal{P} ; λ is the *approximation guarantee* or *approximation bound* of \mathcal{A} . To preserve the convention that $\lambda \geq 1$ for minimization problems also, we define an algorithm \mathcal{A} to be a λ -approximation algorithm for a *minimization* problem \mathcal{P} , if \mathcal{A} produces a solution of value at most $OPT(I) \cdot \lambda$ for all instances I of \mathcal{P} . Our main goal for all problems studied here is to develop polynomial-time algorithms with improved approximation guarantees, i.e., guarantees that are smaller than ones known before.

Notation. Given an undirected graph $G = (V, E)$, we shall let $n = |V|$, $m = |E|$, and let Δ denote the maximum degree of the vertices of G . The diameter of G (the maximum length of a shortest path between any pair of vertices in G) will be denoted $\text{diam}(G)$. Given $S \subseteq V$, let $\delta(S)$ denote the set of edges of G that have precisely one end-point in S . The *edge-expansion* of G , $\beta = \beta(G)$, is $\min_{S \subseteq V: |S| \leq n/2} |\delta(S)|/|S|$. Expansion is well-known to be of much importance in routing: see, e.g., Leighton (1992), Upfal (1992), and Arora, Leighton and Maggs (1996). Given a non-negative integer k , we define $[k] \doteq \{1, 2, \dots, k\}$, and logarithms will be to the base 2 unless specified otherwise. We now present the problems considered.

(i) Edge-disjoint paths and unsplittable flow. Suppose we are given G and a (multi-)set $\mathcal{T} = \{(s_i, t_i) : 1 \leq i \leq k\}$ of pairs of vertices of G . A sub-(multi-)set \mathcal{T}' of \mathcal{T} is termed *realizable* if the pairs of vertices in \mathcal{T}' can be connected in G by mutually edge-disjoint paths. The classical maximum disjoint paths problem (henceforth MDP) is to find a realizable sub-(multi-)set of \mathcal{T} of maximum cardinality. This is a well-known NP-hard problem (Karp (1975)), and has resisted attempts at showing good approximability. Also, much of the difficulty in admission control and virtual-circuit routing in communication networks stems from the lack of good heuristics for the

MDP: see Chapter 1 of Kleinberg (1996). A natural generalization of the MDP, motivated by networks supporting heterogeneous applications, is to allow each pair (s_i, t_i) to have a *demand* $\rho_i > 0$, and each edge f of G to have a *capacity* $c_f > 0$ (Kleinberg (1996)). (We shall reserve the symbol e for the base of the natural logarithm, so a generic edge in G will be denoted by f .) $\mathcal{T}' \subseteq \mathcal{T}$ is now called *realizable* if each pair of vertices (s_i, t_i) in \mathcal{T}' can send ρ_i flow along a single path in G from s_i to t_i , such that the total flow using any edge f does not exceed c_f . The *unsplittable flow problem* (UFP) is to find a realizable $\mathcal{T}' \subseteq \mathcal{T}$ that maximizes $\sum_{i:(s_i, t_i) \in \mathcal{T}'} \rho_i$ (Kleinberg (1996)). As in Kleinberg (1996) and Kolliopoulos and Stein (1997), we assume that $\forall i \forall f, \rho_i \leq c_f$; i.e., that any single demand can be accommodated on any edge. If all capacities c_f are the same, we call the problem *uniform-capacity* UFP (UCUFP); by scaling, we shall always take the common edge capacity to be 1 for the UCUFP. It is easy to see that the MDP is a special case of the UCUFP wherein $\forall i, \rho_i = 1$.

The MDP seems hard to approximate: the current-best approximation guarantee for arbitrary graphs G is $O(\max\{\sqrt{m}, \text{diam}(G)\})$ (Kleinberg (1996)). No non-trivial approximation bound of the form $F(n, m)$ is known for the UCUFP; an $O(\max\{\sqrt{m}, \text{diam}(G)\} \cdot (\max_i \rho_i^{-1}))$ approximation bound is known (Kleinberg (1996)). However, for special classes of graphs such as the mesh, a family of “densely embeddable” graphs, and “strong” expanders G with sufficiently large $\beta(G) = \Omega(1)$, recent breakthrough results have led to $O(1)$ or $\text{polylog}(n)$ approximations for the MDP (e.g., Awerbuch, Gawlick, Leighton and Rabani (1994), Aumann and Rabani (1995), Kleinberg and Tardos (1995), Kleinberg and Tardos (1998), Kleinberg and Rubinfeld (1996), Kleinberg (1996)).

Our approximation guarantees for the UFP will be to within a small constant factor of our guarantees for the MDP; thus, we shall only state our results for the more general UFP. We in fact work with the general *weighted* case of the UFP, where each (s_i, t_i) has a weight $w_i > 0$, with the objective being to find a realizable $\mathcal{T}' \subseteq \mathcal{T}$ that maximizes $\sum_{i:(s_i, t_i) \in \mathcal{T}'} w_i$. (We assume without loss of generality that $\max_i w_i = 1$.) Let $\alpha(\mathcal{T})$ denote the optimal value for a given UFP instance. We first present an efficient algorithm that outputs a feasible path selection of value: $\Omega(\alpha(\mathcal{T}))$ if $\alpha(\mathcal{T}) \geq m$, and $\Omega(\max\{(\alpha(\mathcal{T}))^2/m, \alpha(\mathcal{T})/\sqrt{m}\})$ if $\alpha(\mathcal{T}) < m$. Even for the MDP, a special case of the UFP, this is better than the current-best $\Omega(\alpha(\mathcal{T})/\max\{\sqrt{m}, \text{diam}(G)\})$ mentioned above. Consider, e.g., constant-degree graphs with diameter $\Theta(n)$: our improvement is by a factor of $\Omega(\sqrt{n})$. Also, for constant-degree graphs, if $\alpha(\mathcal{T}) = \Omega(n/\text{polylog}(n))$, our approximation guarantee is $O(\text{polylog}(n))$, as opposed to the $\Omega(\sqrt{n})$ of Kleinberg (1996).

While it is encouraging that we now have an $O(\sqrt{m})$ approximation for the UFP and an improved approximation for the MDP, the approximation bounds are rather high and reflect the generally felt hardness of these problems. It may be that these problems cannot be approximated to within $n^{o(1)}$ unless $P = NP$ or some such unexpected containment holds, but such a result is not (yet) known. Also, as mentioned in Section 6, our results above are essentially best-possible for general graphs if we use our “linear relaxations” based approach. So, a natural problem is to focus on important classes of graphs for which good approximations are not known: in particular, popular interconnection networks such as the butterfly and meshes of dimension 3 and higher. Also, since graph expansion is well-known to be a key parameter in routing, it is natural to relate the approximability of these problems to $\beta(G)$ (Raghavan and Upfal (1994)). We now describe our next set of results, which follow these leads.

Given an instance (G, \mathcal{T}) of the UFP, a well-known approach is to start with its *multi-commodity flow relaxation*: shipping a flow of at most ρ_i from each s_i to t_i subject to no edge f having to carry a total flow of more than c_f , to maximize the total weighted amount of flow sent. That is, we relax the “integrality” constraint of either choosing one path or none at all for each (s_i, t_i) , to its fractional counterpart. This relaxation is a linear program (LP) and hence can be solved/approximated efficiently, and its optimum $\alpha^*(\mathcal{T})$ is an upper bound on $\alpha(\mathcal{T})$. Suppose we have any feasible fractional flow \mathcal{F} —i.e., an (s_i, t_i) -flow of at most ρ_i for each i , where each edge f carries a total flow of at most c_f —such that the objective function value of \mathcal{F} is at least $\alpha^*(\mathcal{T})/2$. (The constant 2 here can be replaced by any constant greater than 1.) Let d be the maximum length (number of edges) of any flow path in \mathcal{F} , i.e., the length of a longest path that carries nonzero flow from any s_i to the corresponding t_i . This d will be crucial in the sequel: a nice result of Kleinberg and Rubinfeld (1996) shows that we can efficiently compute such an \mathcal{F} that has $d = O(\Delta^2 \beta^{-2} \log^3 n)$, for UCUIFP instances. Since d appears in many of our approximations, we pause to note that $d = O(\text{polylog}(n))$ if Δ and β^{-1} are $O(\text{polylog}(n))$, key examples of which are the butterfly and related hypercubic networks ($\Delta = \Theta(1)$ and $\beta = \Theta(1/\log n)$ for them). For the r -dimensional mesh with $n^{1/r}$ nodes per side, $d = O(r^2 n^{2/r} \log^3 n)$.

Our next main result is that the UFP (and hence the MDP) can be approximated to within a factor of $O(d)$. Thus, an important corollary is that the UCUIFP can be approximated to within $\text{polylog}(n)$ on the butterfly and on related networks; the analogous result for the MDP follows from the approaches of Broder, Frieze and Upfal (1994) and Kleinberg and Rubinfeld (1996).

We next consider a special (“ ϵ -bounded”) case of the UFP studied in Kleinberg (1996), where $\forall i \forall f, \rho_i \leq (1-\epsilon)c_f$. An approximation of $O(\epsilon^{-1} \max\{\sqrt{m}, \text{diam}(G)\})$ is given for this in Kleinberg (1996) and we show an approximation of $O(d^{\min\{1/\epsilon-1, 1\}})$: compare the results as ϵ approaches 1. The quantity $B \doteq 1/(1-\epsilon)$ is called the *edge bandwidth*, and it is well-known that $B = \Omega(\log n)$ suffices to ensure an $O(1)$ approximation for the UCUPF (Raghavan and Thompson (1987)). Our $O(d^{1/\epsilon-1})$ approximation shows that for networks such as the butterfly wherein $d = O(\text{polylog}(n))$ as seen above, even if $B = \Omega(\log \log n)$ (i.e., if $\epsilon = 1 - O(1/\log \log n)$), we can achieve $O(1)$ approximations for the UCUPF. Experimental machines such as the iWarp (Borkar et al. (1988)) and the J-Machine (Noakes, Wallach and Dally (1993)) have $B > 1$, e.g., $B \in \{2, 4\}$; also note that $\log \log n < 6$ for all practical purposes. We also apply our $O(d^{1/\epsilon-1})$ approximation to a “routing within limited time” problem.

(ii) Routing in rounds and all-optical routing. Given an instance (G, \mathcal{T}) of the UCUPF, if \mathcal{T} is not realizable, we may still need to connect each s_i to t_i ; so, a natural question is, how to do this in the minimum number of *rounds*. This minimum, $\chi(\mathcal{T})$, is thus the minimum number of realizable sets into which \mathcal{T} can be partitioned, and is clearly NP-hard to determine. Also, finding $\chi(\mathcal{T})$ is the problem of minimizing the number of wavelengths in all-optical networks, and hence has also received attention (Aggarwal et al. (1996), Raghavan and Upfal (1994), Aumann and Rabani (1995), Kleinberg (1996), Rabani (1996)). We only state here our approximation results for the UCUPF, the same also holding for $\chi(\mathcal{T})$ for the MDP. An approach of Aumann and Rabani (1995) and Kleinberg (1996) shows that a ψ -approximation for $\alpha(\mathcal{T})$ translates to an $O(\psi \log n)$ -approximation for $\chi(\mathcal{T})$. Using this with our approximation algorithms for the UCUPF yields the analogous approximation algorithms for $\chi(\mathcal{T})$; however, we show how further improved approximations can be obtained, in Section 4.

The above, along with a few other results, are our contributions. All of this work starts with the corresponding multi-commodity flow relaxation, and uses some new and some known randomized/deterministic approaches to “round” it suitably. Section 2 starts with the multi-commodity flow relaxation for the UFP, and our basic randomized rounding approach. It then presents probabilistic tools that we will be of use in analyzing such algorithms. Our approximation algorithms for the UFP are shown in Section 3, and Section 4 considers the problem of routing in rounds. Further extensions and applications are presented in Section 5, and Section 6 concludes.

2 Multi-commodity flow relaxation and probabilistic tools

Consider the UFP, and recall that $w_i > 0$ is the weight given to (s_i, t_i) . (By scaling, we assume that $\max_i w_i = 1$.) Let $\text{Ed}(u)$ denote the set of edges incident at vertex u . There is a natural integer programming (IP) formulation for a given weighted instance (G, \mathcal{T}) of the UFP:

maximize $\sum_{i \in [k]} w_i x_i$, subject to

$$x_i = \sum_{f \in \text{Ed}(s_i)} y_{i,f}, \quad i = 1, 2, \dots, k; \quad (1)$$

$$x_i = \sum_{f \in \text{Ed}(t_i)} y_{i,f}, \quad i = 1, 2, \dots, k; \quad (2)$$

$$y_{i,f} = \sum_{f' \in \text{Ed}(u): f' \neq f} y_{i,f'}, \quad i = 1, 2, \dots, k, \quad u \notin \{s_i, t_i\}, \quad f \in \text{Ed}(u); \quad (3)$$

$$\sum_{i \in [k]} \rho_i y_{i,f} \leq c_f, \quad f \in E; \quad (4)$$

$$x_i \in \{0, 1\}, \quad i = 1, 2, \dots, k; \quad (5)$$

$$y_{i,f} \in \{0, 1\}, \quad i = 1, 2, \dots, k, \quad f \in E. \quad (6)$$

The variable x_i is 1 if we choose to connect s_i to t_i by a path, and is 0 otherwise. The variable $y_{i,f}$ is 1 if the path chosen (if any) for (s_i, t_i) , passes through f ; it is 0 otherwise. The constraints (1), (2) and (3) model these. Next, (4) is the capacity constraint on edge f . Finally, (5) and (6) are the integrality constraints on the variables. If we relax (5) and (6), and let each x_i and $y_{i,f}$ to be a real lying in $[0, 1]$, we get a natural LP (multi-commodity flow) relaxation.

From now on, we let $\{x_i^*, y_{i,f}^* \in [0, 1] : i \in [k], f \in E\}$ denote a generic feasible (not necessarily optimal) solution to the LP relaxation, with objective function value $y^* = \sum_i w_i x_i^*$. We first transform this flow in polynomial time into a set of flow-paths, using the well-known “flow decomposition” idea (Ahuja, Magnanti and Orlin (1993)). Thus, we now have, for each i , a set $\{P_{i,1}, P_{i,2}, \dots, P_{i,\ell_i}\}$ of (s_i, t_i) -paths, where $\ell_i \leq m$. Each path $P_{i,j}$ carries a positive amount $\rho_i z_{i,j}^*$ of flow from s_i to t_i , where $x_i^* = \sum_j z_{i,j}^*$ and where, by the capacity constraints, we have

$$\forall f \in E, \quad \sum_{(i,j): f \in P_{i,j}} \rho_i z_{i,j}^* \leq c_f. \quad (7)$$

As mentioned in the introduction, we shall let d denote the length of a longest $P_{i,j}$; this shall be a key parameter for us. We will occasionally view each $P_{i,j}$ as a set of edges.

Theorems 3.1, 3.2, 3.3 and Corollary 3.1 will show how to use randomization to “round” the above fractional solution well. The rounding process will be to choose a suitable $\gamma > 1$, and,

independently for each i, j , to define $z_{i,j}$ to be 1 with probability $z_{i,j}^*/\gamma$, and $z_{i,j}$ to be 0 with probability $1 - z_{i,j}^*/\gamma$. Let us call this “ γ -rounding”. For each i : (i) if at least one of the $z_{i,j}$ ’s is 1, we shall arbitrarily choose one of them, say $z_{i,k}$, and connect s_i to t_i by path $P_{i,k}$; (ii) if all the $z_{i,j}$ ’s are 0, we do not connect s_i to t_i . An alternative approach as in Raghavan and Thompson (1987) and Srinivasan (1996), might have been to randomly round *at most* one $z_{i,j}$ to 1, for each i : it seems wasteful to potentially round more than one $z_{i,j}$ to 1, for any given i . However, we shall see that it buys us sufficient independence and desirable correlations.

Given some $f \in E$, note that the pairs $\{(s_i, t_i) : i \in [k]\}$ request a total demand of **at most**

$$R_f \doteq \sum_{(i,j): f \in P_{i,j}} \rho_i z_{i,j} \tag{8}$$

from f . (We have used the phrase “at most” here because even if some $z_{i,j}$ gets defined as one, the path $P_{i,j}$ may not be used if some other $z_{i,k}$ was also one, as seen above.) From (7), we get

$$\mathbf{E}[R_f] = \sum_{(i,j): f \in P_{i,j}} \rho_i z_{i,j}^*/\gamma \leq c_f/\gamma. \tag{9}$$

For any $f \in E$, let E_f denote the “bad” event that $R_f > c_f$; if $\overline{E_f}$ holds, then edge f ’s capacity is not violated. Let “r.v.” denote the phrase “random variable”; all r.v.s that we consider will take on values from *finite* domains. Our focus from now on will be to show that for a suitable γ , γ -rounding will, with positive probability, avoid all the bad events E_f **and** keep the objective function reasonably high; note that the underlying collection of r.v.s is $\mathcal{R} \doteq \{z_{i,j} : i \in [k], j \in [\ell_i]\}$. Section 2.1 presents tools that help bound probabilities such as $\Pr[E_f]$ for *individual* edges f . Section 2.2 then shows how to analyze the correlations involved and get constructive results (i.e., how to set values for the r.v.s in \mathcal{R} appropriately).

Note that \mathcal{R} is the set of underlying independent binary r.v.s in our basic algorithm above. For the sake of generality, Sections 2.1 and 2.2 present probabilistic results for an arbitrary sequence $\vec{X} = (X_1, X_2, \dots, X_\ell)$ of binary, independent r.v.s. The sequence \vec{X} can thus be associated with any fixed ordering of the variables in \mathcal{R} . The reader is asked to keep this correspondence in mind; whenever we specify key tools in Sections 2.1 and 2.2, we also point out how they will help in our analyses that begin in Section 3.

2.1 Large-deviation bounds

As mentioned above, suppose $\vec{X} = (X_1, X_2, \dots, X_\ell)$ is a sequence of binary, independent r.v.s. We will say that an event Z is an *assignment event* w.r.t. \vec{X} , iff Z is of the form “ $\bigwedge_{i \in S} (X_i = b_i)$ ”, for any $S \subseteq [\ell]$ and any sequence of values b_i . Now, even conditional on any assignment event Z , we can still view the X_i ’s as independent: it is just that for all $i \in S$, $X_i = b_i$ with probability 1. This simple idea, which we will call the “independence view”, will be useful. In our setting (where \vec{X} is associated with our set \mathcal{R}), most of our algorithms will be based on rounding the r.v.s $z_{i,j}$ one-by-one, deterministically. Thus, our assignment events will be settings of the variables in some subset of \mathcal{R} ; we will need to condition on such assignment events, to determine how to round the remaining r.v.s. In this process, the independence view will simplify things greatly for us.

We recall the Chernoff-Hoeffding bounds (see, e.g., Motwani and Raghavan (1995)):

Theorem 2.1 *Let $\vec{X} = (X_1, \dots, X_\ell)$ be a sequence of independent r.v.s, with each $X_i \in \{0, 1\}$, and let $u_1, \dots, u_\ell \in [0, 1]$. Define $R = \sum_i u_i X_i$, and let $\mathbf{E}[R] = \mu$. Also let Z be any assignment event w.r.t. \vec{X} . Then:*

(i) For any $\delta \geq 0$,

$$\Pr[R \geq \mu(1 + \delta) | Z] \leq \frac{\mathbf{E}[(1 + \delta)^R | Z]}{(1 + \delta)^{\mu(1 + \delta)}} = \frac{\prod_{i \in [\ell]} \mathbf{E}[(1 + \delta)^{u_i X_i} | Z]}{(1 + \delta)^{\mu(1 + \delta)}}; \quad (10)$$

$$\Pr[R \geq \mu(1 + \delta)] \leq \frac{\prod_{i \in [\ell]} \mathbf{E}[(1 + \delta)^{u_i X_i}]}{(1 + \delta)^{\mu(1 + \delta)}} \leq G(\mu, \delta) \doteq (e^\delta / (1 + \delta))^{(1 + \delta)\mu}. \quad (11)$$

(ii) If $0 \leq \delta < 1$, then

$$\Pr[R \leq \mu(1 - \delta) | Z] \leq \frac{\mathbf{E}[(1 - \delta)^R | Z]}{(1 - \delta)^{\mu(1 - \delta)}} = \frac{\prod_{i \in [\ell]} \mathbf{E}[(1 - \delta)^{u_i X_i} | Z]}{(1 - \delta)^{\mu(1 - \delta)}}; \quad (12)$$

$$\Pr[R \leq \mu(1 - \delta)] \leq \frac{\prod_{i \in [\ell]} \mathbf{E}[(1 - \delta)^{u_i X_i}]}{(1 - \delta)^{\mu(1 - \delta)}} \leq H(\mu, \delta) \doteq e^{-\mu\delta^2/2}. \quad (13)$$

Note from (8) that for any edge $f \in E$, R_f is an r.v. such as the r.v. R of Theorem 2.1. Thus, bounds such as (10) and (11) will help in bounding $\Pr[E_f]$.

We next record some simple properties of the above functions G and H :

Lemma 2.1 (i) *If $\mu_1(1 + \delta_1) = \mu_2(1 + \delta_2)$ where $\mu_1 \leq \mu_2$ and $\mu_1, \mu_2, \delta_1, \delta_2 \geq 0$, then $G(\mu_1, \delta_1) \leq G(\mu_2, \delta_2)$.* (ii) *If $\mu_1 \geq \mu_2 \geq 0$ and $\delta_1, \delta_2 \in [0, 1]$ are such that $\mu_1(1 - \delta_1) = \mu_2(1 - \delta_2)$, then $H(\mu_1, \delta_1) \leq H(\mu_2, \delta_2)$.*

Proof: (i) Suppose we hold μ_1 and δ_1 fixed; let $a = \mu_1(1 + \delta_1)$. Then, $0 \leq \delta_2 \leq \delta_1$ and $\mu_2 = a/(1 + \delta_2)$. Now, $\ln(G(\mu_2, \delta_2)) = a(\delta_2/(1 + \delta_2) - \ln(1 + \delta_2))$ has a derivative (w.r.t. δ_2) equaling $-a\delta_2/(1 + \delta_2)^2$. Thus, in the range $[0, \delta_1]$, $\ln(G(\mu_2, \delta_2))$ (and hence $G(\mu_2, \delta_2)$) is minimized when $\delta_2 = \delta_1$. (ii) We need to show that $\mu_1\delta_1^2 \geq \mu_2\delta_2^2$; as above, this is accomplished by holding μ_1 and δ_1 fixed and then showing that $\mu_2\delta_2^2$ is maximized when $\delta_2 = \delta_1$. ■

We also need a simple proposition:

Proposition 2.1 *If $\sum_{i=1}^N a_i \leq b$ where $a_i \geq 0$ for each i , then $\sum_{i < j} a_i a_j \leq \binom{N}{2} (b/N)^2 < b^2/2$.*

Proof: If $a_u < a_v$ for some u, v , then setting $a_u := a_u + \epsilon$ and $a_v := a_v - \epsilon$ for any $\epsilon \in (0, a_v - a_u)$, increases the value of $\sum_{i < j} a_i a_j$. Thus, subject to the condition that $\sum_{i=1}^N a_i = b'$ for any $0 \leq b' \leq b$, $\sum_{i < j} a_i a_j$ is maximized when each a_i equals b'/N . Thus, $\sum_{i < j} a_i a_j \leq \binom{N}{2} (b/N)^2 < b^2/2$. ■

2.2 Correlation inequalities and well-behaved estimators

We now introduce ways to analyze the *correlations* among our events such as $\overline{E_f}$; we will then present Theorem 2.3, which will let us extract efficient algorithms from our probabilistic arguments. As in Section 2.1, we present the results for a general sequence $\vec{X} \doteq (X_1, X_2, \dots, X_\ell)$ of binary independent r.v.s; we will point out specializations to our set \mathcal{R} of r.v.s, as we go along. All events and r.v.s considered in this subsection will be assumed to be completely determined by the value of \vec{X} .

The FKG inequality. This powerful inequality that originated in statistical physics (Fortuin, Ginibre and Kasteleyn (1971)) is summarized as follows for our purposes. Given $\vec{a} = (a_1, \dots, a_\ell) \in \{0, 1\}^\ell$ and $\vec{b} = (b_1, \dots, b_\ell) \in \{0, 1\}^\ell$, let us say that $\vec{a} \preceq \vec{b}$ iff $a_i \leq b_i$ for all i . Define an event \mathcal{A} to be *increasing* iff: for all $\vec{a} \in \{0, 1\}^\ell$ such that \mathcal{A} holds when $\vec{X} = \vec{a}$, \mathcal{A} also holds when $\vec{X} = \vec{b}$, for any \vec{b} such that $\vec{a} \preceq \vec{b}$. Analogously, event \mathcal{A} is said to be *decreasing* iff: for all $\vec{a} \in \{0, 1\}^\ell$ such that \mathcal{A} holds when $\vec{X} = \vec{a}$, \mathcal{A} also holds when $\vec{X} = \vec{b}$, for any $\vec{b} \preceq \vec{a}$. In our setting with \mathcal{R} playing the role of \vec{X} , note, e.g., that any ‘‘capacity non-violation’’ event such as $\overline{E_f}$ is a decreasing event, while the objective function $\sum_i w_i \bigvee_j (z_{i,j} = 1)$ is increasing. Thus, non-trivial ways of bounding correlations among increasing and decreasing events will help.

The FKG inequality proves the ‘‘intuitively plausible’’ idea that any set of increasing events are *positively correlated* with each other; analogously for any set of decreasing events. Similarly, any

increasing event is negatively correlated with any set of decreasing events; each decreasing event is negatively correlated with any set of increasing events. The formal statement is:

Theorem 2.2 (FKG inequality) *Let I_1, I_2, \dots, I_t be any collection of increasing events and D_1, D_2, \dots, D_t be any sequence of decreasing events (each I_i and D_i completely determined by \vec{X}). Then for any $i \in [t]$ and any $S \subseteq [t]$,*

$$(i) \Pr[I_i | \bigwedge_{j \in S} I_j] \geq \Pr[I_i] \text{ and } \Pr[D_i | \bigwedge_{j \in S} D_j] \geq \Pr[D_i];$$

$$(ii) \Pr[I_i | \bigwedge_{j \in S} D_j] \leq \Pr[I_i] \text{ and } \Pr[D_i | \bigwedge_{j \in S} I_j] \leq \Pr[D_i].$$

Well-behaved estimators. Suppose \mathcal{E} is some event (determined completely by \vec{X} , as assumed above). A random variable g is said to be a *well-behaved* estimator for \mathcal{E} (w.r.t. \vec{X}) iff it satisfies the following three properties, for all $t \leq \ell$, for all $T = \{i_1, i_2, \dots, i_t\} \subseteq [\ell]$, and for all $b_1, b_2, \dots, b_t \in \{0, 1\}$ for which $\Pr[\bigwedge_{s=1}^t (X_{i_s} = b_s)] > 0$:

$$(P1) \mathbf{E}[g | \bigwedge_{s=1}^t (X_{i_s} = b_s)] \text{ is efficiently computable (i.e., in time polynomial in } \ell);$$

$$(P2) \Pr[\mathcal{E} | \bigwedge_{s=1}^t (X_{i_s} = b_s)] \leq \mathbf{E}[g | \bigwedge_{s=1}^t (X_{i_s} = b_s)]; \text{ and}$$

$$(P3) \text{ if } \mathcal{E} \text{ is increasing, then for all } i_{t+1} \in ([\ell] - T) \text{ for which } \Pr[X_{i_{t+1}} = 1] \in (0, 1),$$

$$\mathbf{E}[g | (X_{i_{t+1}} = 0) \wedge \bigwedge_{s=1}^t (X_{i_s} = b_s)] \leq \mathbf{E}[g | (X_{i_{t+1}} = 1) \wedge \bigwedge_{s=1}^t (X_{i_s} = b_s)].$$

Taking g to be the indicator variable for \mathcal{E} will satisfy (P2) and (P3), but not necessarily (P1). So the intuition here is that we want to approximate quantities such as $\Pr[\mathcal{E} | \bigwedge_{s=1}^t (X_{i_s} = b_s)]$ “well” (in the sense of (P2) and (P3)) by an *efficiently computable* value ($\mathbf{E}[g | \bigwedge_{s=1}^t (X_{i_s} = b_s)]$).

Suppose E_1, E_2, \dots, E_{m+1} are all “bad” events: we would like to find an assignment for \vec{X} that avoids **all** of the E_i . (For us, E_i , $i = 1, 2, \dots, m$, could represent the “capacity violation” events for the m edges of G . Finally, E_{m+1} could be the event that, say, the objective function is less than half its expected value. We aim for a rounding that avoids all of these bad events.) We now follow an approach of Srinivasan (1995) based on the method of conditional probabilities (Erdős and Selfridge (1973), Spencer (1987), Raghavan (1988)), that shows a useful sufficient condition for this.

Theorem 2.3 *Suppose $\vec{X} = (X_1, X_2, \dots, X_\ell)$ is a sequence of independent r.v.s X_i , with $X_i \in \{0, 1\}$ for each i . Let E_1, E_2, \dots, E_{m+1} be events and g_1, g_2, \dots, g_{m+1} be r.v.s, all completely determined by \vec{X} , such that each g_i is a well-behaved estimator for E_i (w.r.t. \vec{X}). Suppose further that*

E_1, E_2, \dots, E_m (and not necessarily E_{m+1}) are all increasing events (as functions of \vec{X}). Then, if

$$1 - \left(\prod_{i \in [m]} (1 - \min\{\mathbf{E}[g_i], 1\}) \right) + \mathbf{E}[g_{m+1}] < 1 \quad (14)$$

holds, we can efficiently construct a deterministic assignment for \vec{X} under which none of E_1, E_2, \dots, E_{m+1} hold.

Proof: We first upper-bound $\Pr[\exists i \in [m] : E_i] = 1 - \Pr[\bigwedge_{i \in [m]} \overline{E_i}]$ as

$$\begin{aligned} 1 - \Pr[\bigwedge_{i \in [m]} \overline{E_i}] &\leq 1 - \prod_{i \in [m]} \Pr[\overline{E_i}] \quad (\text{by the FKG inequality, since } \overline{E_1}, \dots, \overline{E_m} \text{ are all decreasing}) \\ &= 1 - \prod_{i \in [m]} (1 - \Pr[E_i]) \leq 1 - \prod_{i \in [m]} (1 - \min\{\mathbf{E}[g_i], 1\}); \end{aligned}$$

the last inequality is a consequence of (P2). Also, $\Pr[E_{m+1}] \leq \mathbf{E}[g_{m+1}]$ by (P2). Thus,

$$\Pr[\exists i \in [m+1] : E_i] \leq \Pr[\exists i \in [m] : E_i] + \Pr[E_{m+1}] \leq 1 - \left(\prod_{i \in [m]} (1 - \min\{\mathbf{E}[g_i], 1\}) \right) + \mathbf{E}[g_{m+1}] < 1,$$

by (14). Thus, we know that there exists a value of \vec{X} that avoids all of E_1, E_2, \dots, E_{m+1} .

How to find such a value for \vec{X} ? For any $t \leq \ell$ and any $\vec{b} = (b_1, b_2, \dots, b_t) \in \{0, 1\}^t$, define

$$h(t, \vec{b}) = 1 - \left(\prod_{i \in [m]} (1 - \min\{\mathbf{E}[g_i | \bigwedge_{j \in [t]} (X_j = b_j)], 1\}) \right) + \mathbf{E}[g_{m+1} | \bigwedge_{j \in [t]} (X_j = b_j)].$$

We first round the r.v.s X_i for which $\Pr[X_i = 1] \in \{0, 1\}$, in the obvious way. Renumber the r.v.s so that $\Pr[X_i = 1] \in \{0, 1\}$ is true precisely for the indices $i = 1, 2, \dots, t'$, for some t' . For $i = 1, 2, \dots, t'$, define $\xi_i = 0$ if $\Pr[X_i = 0] = 1$, and $\xi_i = 1$ if $\Pr[X_i = 1] = 1$. Then, for $i = 1, 2, \dots, t'$, it is easy to see that $h(i, (\xi_1, \xi_2, \dots, \xi_i)) = h(i-1, (\xi_1, \xi_2, \dots, \xi_{i-1}))$. Thus we see from (14) that $h(t', (\xi_1, \xi_2, \dots, \xi_{t'})) < 1$.

Next, suppose $t' \leq t \leq \ell - 1$, and that $(b_1, b_2, \dots, b_t) \in \{0, 1\}^t$ with $b_1 = \xi_1, b_2 = \xi_2, \dots, b_{t'} = \xi_{t'}$. A result of Srinivasan (1995) that will be of crucial use to us is that

$$\min_{r \in \{0, 1\}} h(t+1, (b_1, b_2, \dots, b_t, r)) \leq h(t, \vec{b}). \quad (15)$$

Consider the following algorithm, which starts with \vec{b} initialized to the list $(\xi_1, \xi_2, \dots, \xi_{t'})$:

for $i := t' + 1$ **to** ℓ **do**:

$$b_i := \operatorname{argmin}_{r \in \{0, 1\}} h(i, (b_1, b_2, \dots, b_{i-1}, r));$$

$$\vec{b} := (b_1, b_2, \dots, b_i).$$

Starting with the initial condition $h(t', (\xi_1, \xi_2, \dots, \xi_{t'})) < 1$ and using (15), it is easy to see by induction on i that we maintain the property $h(i, \vec{b}) < 1$. Setting $i = \ell$, this implies, by (P2), that the final vector \vec{b} indeed is such that all of the E_i are avoided when $\vec{X} = \vec{b}$. ■

We shall use Theorem 2.3; note that it works even if $\Pr[\bigwedge_{i \in [m+1]} \overline{E_i}]$ is very small. We also remark that we will often need to show that an r.v. $g = g(\vec{X})$ is such that for any assignment event \mathcal{A} w.r.t. X , we can efficiently compute $\mathbf{E}[g(\vec{X}) \mid \mathcal{A}]$ (e.g., we may want to verify property (P1) of well-behaved estimators). The “independence view” is very useful in such situations. In all such situations, g will be a sum of terms (r.v.s) of the form $\prod_{i=1}^v f_{u_i}(X_{u_i})$, where all the u_i s are distinct. Even conditional on \mathcal{A} , the independence view lets us compute the expectation of this term as $\prod_{i=1}^v \mathbf{E}[f_{u_i}(X_{u_i})]$; thus, it will suffice if each $\mathbf{E}[f_{u_i}(X_{u_i})]$ is efficiently computable, which will always be the case for us.

3 Approximation algorithms for the UFP

The main results of Section 3 are Theorems 3.1, 3.2, 3.3, and Corollary 3.1. We start with Lemma 3.1, which bounds $\max_{f \in E} \Pr[E_f]$. Its part (ii) is useful when all demands ρ_i are “small”, while its part (iii) helps handle the case where all demands are “large”.

Lemma 3.1 *Fix any $f \in E$. We have:*

(i) $\Pr[E_f] \leq 1/\gamma$.

(ii) *Suppose, for some $\lambda \geq 1$, that for all $P_{i,j}$ passing through f , we have $\rho_i \leq c_f/\lambda$. Then, there is a well-behaved estimator u_f for E_f (w.r.t. \mathcal{R}), such that $\mathbf{E}[u_f] \leq G(\lambda/\gamma, \gamma - 1)$, where G is defined in the statement of Theorem 2.1.*

(iii) *Suppose, for some $\eta \in (0, 1]$, that for all $P_{i,j}$ passing through f , we have $\rho_i \geq c_f\eta$. Then, there is a well-behaved estimator v_f for E_f (w.r.t. \mathcal{R}), with $\mathbf{E}[v_f] \leq (\eta\gamma)^{-2}/2$.*

Proof: Bound (i) follows directly from Markov’s inequality, using (9). To see (ii), define $\rho'_i = \lambda\rho_i/c_f$ and $R'_f \doteq \sum_{(i,j): f \in P_{i,j}} \rho'_i z_{i,j}$. Since $\rho'_i \in [0, 1]$ for all $P_{i,j}$ passing through f , R'_f is a sum of independent r.v.s, each taking values in $[0, 1]$. Also, (9) shows that $\mu \doteq \mathbf{E}[R'_f]$ is at most λ/γ . Define $\delta > 0$ by $\mu(1 + \delta) = \lambda$; let

$$u_f \doteq \frac{\prod_{(i,j): f \in P_{i,j}} (1 + \delta)^{\rho'_i z_{i,j}}}{(1 + \delta)^\lambda}.$$

Now, $E_f \equiv (R'_f > \lambda)$. Using (10), (11), part (i) of Lemma 2.1 and the independence view, we can check that u_f is a well-behaved estimator for E_f (w.r.t. \mathcal{R}), with $\mathbf{E}[u_f] \leq G(\lambda/\gamma, \gamma - 1)$.

As for (iii), note that $R_f > 1$ holds only if $\sum_{(i,j): f \in P_{i,j}} z_{i,j} \geq 2$; this holds even conditional on any assignment event w.r.t. \mathcal{R} . Define

$$v_f \doteq \sum_{(i,j),(i',j'): (i,j) \neq (i',j'), f \in (P_{i,j} \cap P_{i',j'})} z_{i,j} \cdot z_{i',j'}. \quad (16)$$

(Each choice of distinct pairs (i, j) and (i', j') is counted exactly once—not twice—in the sum in (16).)

The independence view helps show that v_f is a well-behaved estimator for E_f (w.r.t. \mathcal{R}). Also, since $\rho_i \geq c_f \eta$ for all $P_{i,j}$ passing through f , we see from (9) that $\sum_{(i,j): f \in P_{i,j}} \mathbf{E}[z_{i,j}] \leq (\eta\gamma)^{-1}$. This, in combination with Proposition 2.1, shows that $\mathbf{E}[v_f] \leq (\eta\gamma)^{-2}/2$. ■

The next definition and lemma handle the objective function.

Definition 3.1 For each $i \in [k]$, let Z_i be the r.v. that is 1 if at least one of the $z_{i,j}$'s is 1, and let Z_i be 0 if all the $z_{i,j}$'s are 0. For any $T \subseteq [k]$, define $y_T^* \doteq \sum_{i \in T} w_i x_i^*$ and $Z(T) \doteq \sum_{i \in T} w_i Z_i$.

Lemma 3.2 For any $T \subseteq [k]$ and $\psi \in (0, 1)$, there is a well-behaved estimator $W_{\psi, T}$ for the event “ $Z(T) < y_T^*(1 - \psi)(1 - 1/e)/\gamma$ ” (w.r.t. \mathcal{R}), such that $\mathbf{E}[W_{\psi, T}] \leq H(y_T^*(1 - 1/e)/\gamma, \psi)$ (the function H is from the statement of Theorem 2.1).

Proof: For each $i \in T$, $\Pr[Z_i = 1] = 1 - \prod_j (1 - z_{i,j}^*/\gamma) \geq 1 - e^{-(\sum_j z_{i,j}^*)/\gamma} = 1 - e^{-x_i^*/\gamma}$; the last term is at least $x_i^*(1 - 1/e)/\gamma$ since for all $y \in [0, 1]$, $1 - e^{-y} \geq y(1 - 1/e)$. Thus, $\mathbf{E}[Z(T)] = \sum_{i \in T} w_i \mathbf{E}[Z_i] \geq \sum_{i \in T} w_i x_i^*(1 - 1/e)/\gamma = y_T^*(1 - 1/e)/\gamma$. Let $\mu_1 = \mathbf{E}[Z(T)]$ and let $\delta_1 \in (0, 1)$ be given by $\mu_1(1 - \delta_1) = y_T^*(1 - \psi)(1 - 1/e)/\gamma$. Defining

$$W_{\psi, T} = \frac{\prod_{i \in T} (1 - \delta_1)^{w_i Z_i}}{(1 - \delta_1)^{(1 - \delta_1)\mu_1}},$$

we can see using part (ii) of Theorem 2.1 and of Lemma 2.1 that $W_{\psi, T}$ is a well-behaved estimator for the event “ $Z(T) < y_T^*(1 - \psi)(1 - 1/e)/\gamma$ ” (w.r.t. \mathcal{R}), with $\mathbf{E}[W_{\psi, T}] \leq H(y_T^*(1 - 1/e)/\gamma, \psi)$. (Since “ $Z(T) < y_T^*(1 - \psi)(1 - 1/e)/\gamma$ ” is **not** an increasing event, we need not check property (P3) of well-behaved estimators.) ■

As mentioned in Section 1, we make the usual *balance assumption* that $(\max_i \rho_i) \leq (\min_f c_f)$; thus, by scaling all the demands and capacities uniformly, we may assume

$$\forall i, \rho_i \leq 1; \quad \forall f, c_f \geq 1. \quad (17)$$

We partition $[k]$ into $S_0 = \{i \in [k] : \rho_i \leq 1/2\}$ and $S_1 = \{i \in [k] : \rho_i > 1/2\}$.

3.1 The first approximation algorithm

We now present our first approximation result for $\alpha(\mathcal{T})$.

Theorem 3.1 *There is an absolute constant $c_3 > 0$ such that for any weighted instance of the UFP, we can produce a feasible path-selection in deterministic polynomial-time, with the objective function value $\sum_{i \in [k]} w_i Z_i$ being at least $c_3 \cdot \min\{(\alpha^*(\mathcal{T}))^2/m, \alpha^*(\mathcal{T})\}$.*

Proof: Let $\{x_i^*, y_{i,f}^* \in [0, 1] : i \in [k], f \in E\}$ be an optimal solution to the LP relaxation, computed in polynomial time. Thus, $y^* = \sum_i w_i x_i^* = \alpha^*(\mathcal{T})$. We treat the “small” and “large” demands separately. Define $S = S_0$ if $y_{S_0}^* \geq y_{S_1}^*$, and $S = S_1$ otherwise; thus, $y_S^* \geq y^*/2$. We now show how to do a feasible path selection for S with objective function value $\Omega(\min\{(y_S^*)^2/m, y_S^*\}) = \Omega(\min\{(y^*)^2/m, y^*\})$. In what follows, u_f, v_f and $W_{\psi, \mathcal{T}}$ are the r.v.s from Lemmas 3.1 and 3.2.

First suppose $S = S_0$. Since $c_f \geq 1$ for all f , we can, in the notation of part (ii) of Lemma 3.1, take $\lambda = 2$ here. Suppose we do a γ -rounding (for a suitable $\gamma > 1$ to be chosen below). By part (ii) of Lemma 3.1, $\mathbf{E}[u_f] \leq e^2/\gamma^2$ holds for all $f \in E$. Also, by Lemma 3.2,

$$\mathbf{E}[W_{0.5, S}] \leq H(y_S^*(1 - 1/e)/\gamma, 0.5) \leq H(y^*(1 - 1/e)/(2\gamma), 0.5),$$

since $y_S^* \geq y^*/2$. A simple calculation shows that for a suitably large absolute constant $c_4 > 0$, $\gamma = c_4 \cdot \max\{m/y^*, 1\}$ satisfies: (i) $e^2/\gamma^2 < 1$ and (ii) $(1 - e^2/\gamma^2)^m > H(y^*(1 - 1/e)/(2\gamma), 0.5)$; thus we have

$$1 - \left(\prod_{f \in E} (1 - \min\{\mathbf{E}[u_f], 1\}) \right) + \mathbf{E}[W_{0.5, S}] < 1. \quad (18)$$

Note from Lemmas 3.1 and 3.2 that u_f and $W_{0.5, S}$ are well-behaved estimators for the events E_f and “ $Z(S) < y_S^*(1 - 1/e)/(2\gamma)$ ” respectively. Also, all the events E_f are increasing. Thus, (18) and Theorem 2.3 combine to show that if $S = S_0$, we can efficiently do a feasible path selection for S with objective function value $\Omega(\min\{(y^*)^2/m, y^*\})$.

The case $S = S_1$ needs a little more work. Fix any $f \in E$. If $c_f \geq 2$, we can, as above, take $\lambda = 2$; $\mathbf{E}[u_f] \leq e^2/\gamma^2$ holds as above. If $c_f < 2$, then since $\rho_i > 1/2$ for all $i \in S_1$, we can take $\eta = 1/4$ in the notation of part (iii) of Lemma 3.1; by part (iii) of Lemma 3.1, we have $\mathbf{E}[v_f] \leq 8/\gamma^2$. Thus, whether $c_f \geq 2$ or not, we have a well-behaved estimator s_f for E_f , with $\mathbf{E}[s_f] \leq 8/\gamma^2$. The

desired analog of (18) now is

$$1 - \left(\prod_{f \in E} (1 - \min\{\mathbf{E}[s_f], 1\}) \right) + \mathbf{E}[W_{0.5, S}] < 1.$$

Thus, it suffices if $8/\gamma^2 < 1$ and $(1 - 8/\gamma^2)^m > H(y^*(1 - 1/e)/(2\gamma), 0.5)$, which are once again assured by taking $\gamma = c_4 \cdot \max\{m/y^*, 1\}$ for a suitably large constant c_4 . ■

Corollary 3.1 *Given any weighted instance of the UFP we can efficiently do a feasible path selection with objective function value $\Omega(\alpha^*(\mathcal{T})/\sqrt{m})$.*

Proof: If $\alpha^*(\mathcal{T}) \geq \sqrt{m}$, we are done by Theorem 3.1. So suppose $\alpha^*(\mathcal{T}) < \sqrt{m}$. We may assume without loss of generality that there is an (s_i, t_i) -path in G for each i . Recalling that $\max_i w_i = 1$, simply choose a j such that $w_j = 1$ and connect s_j to t_j ; this ensures an objective function value equaling 1, which is $\Omega(\alpha^*(\mathcal{T})/\sqrt{m})$. ■

So we can efficiently do a path-selection of value $\Omega(\min\{\alpha^*(\mathcal{T}), \max\{(\alpha^*(\mathcal{T}))^2/m, \alpha^*(\mathcal{T})/\sqrt{m}\})$.

3.2 The second approximation: parametrization by d

Recall from the introduction that d denotes the maximum length of the flow-paths $P_{i,j}$. Since d can be assumed “small” in some cases (e.g., for expanders), we would like approximations parametrized by d . Our approach of Section 3.1 does not seem to help directly with this, as it treats the objective function *separately* from the “capacity non-violation” constraints. More precisely, we now wish to lower-bound probabilities such as $\Pr[z_{i,j} = 1 \mid \bigwedge_{f \in E} \overline{E_f}]$. Part (ii) of the FKG inequality shows that this quantity is *upper-bounded* by $\Pr[z_{i,j} = 1]$, but what we require is a lower bound. (We will actually require a lower bound on $\Pr[z_{i,j} = 1 \mid (\mathcal{A} \wedge \bigwedge_{f \in E} \overline{E_f})]$, where \mathcal{A} is an assignment event; the independence view will let us handle this extra conditioning on \mathcal{A} .)

We start with Lemma 3.3, which is motivated by the work of Boppana and Spencer (1989); this will then help us prove Lemma 3.4, which handles the above-mentioned lower-bounding problem. To appreciate Lemma 3.3, the reader is asked to associate: (i) $\{y_1, \dots, y_\ell\}$ with the r.v.s $\{z_{i,j}\}$, (ii) $\{U_1, \dots, U_t\}$ with the “capacity violation” events $\{E_f\}$, and (iii) \mathcal{A} with any assignment event w.r.t. \mathcal{R} . Finally, suppose y_u gets associated with the r.v. $z_{i,j}$. Then, S_u will be the set of edges contained in the path $P_{i,j}$; note that $|S_u| \leq d$. With these associations, Lemma 3.3 will be of help

to us: e.g., the independence view shows that the term “ $\Pr[(y_u = 1) \mid (\mathcal{A} \wedge \bigwedge_{r \in ([t] - S_u)} \overline{U}_r)]$ ” in the statement of Lemma 3.3, simply equals $\Pr[(y_u = 1) \mid \mathcal{A}]$.

Lemma 3.3 *Let y_1, \dots, y_ℓ be arbitrary r.v.s, each taking values in $\{0, 1\}$; let $\mathcal{A}, U_1, \dots, U_t$ be any events. For any $S_1, S_2, \dots, S_\ell \subseteq [t]$, $\Pr[(\exists j : y_j = 1) \mid (\mathcal{A} \wedge \bigwedge_{r=1}^t \overline{U}_r)]$ is at least*

$$\sum_{j=1}^{\ell} (\Pr[(y_j = 1) \mid (\mathcal{A} \wedge \bigwedge_{r \in ([t] - S_j)} \overline{U}_r)] - \sum_{s \in S_j} \Pr[((y_j = 1) \wedge U_s) \mid (\mathcal{A} \wedge \bigwedge_{r \in ([t] - S_j)} \overline{U}_r)]) - \sum_{1 \leq j < j' \leq \ell} \Pr[(y_j = y_{j'} = 1) \mid (\mathcal{A} \wedge \bigwedge_{r=1}^t \overline{U}_r)].$$

Proof: By inclusion-exclusion, $\Pr[(\exists j : y_j = 1) \mid (\mathcal{A} \wedge \bigwedge_{r=1}^t \overline{U}_r)]$ is at least

$$\sum_{j=1}^{\ell} \Pr[(y_j = 1) \mid (\mathcal{A} \wedge \bigwedge_{r=1}^t \overline{U}_r)] - \sum_{1 \leq j < j' \leq \ell} \Pr[(y_j = y_{j'} = 1) \mid (\mathcal{A} \wedge \bigwedge_{r=1}^t \overline{U}_r)]. \quad (19)$$

Fix an arbitrary j . Using the inequality $\Pr[A \mid (B \wedge C)] \geq \Pr[(A \wedge B) \mid C]$, we lower-bound $\Pr[(y_j = 1) \mid (\mathcal{A} \wedge \bigwedge_{r=1}^t \overline{U}_r)]$ by $\Pr[((y_j = 1) \wedge \bigwedge_{s \in S_j} \overline{U}_s) \mid (\mathcal{A} \wedge \bigwedge_{r \in ([t] - S_j)} \overline{U}_r)]$; this, in turn, can be lower-bounded using inclusion-exclusion, by the expression

$$\Pr[(y_j = 1) \mid (\mathcal{A} \wedge \bigwedge_{r \in ([t] - S_j)} \overline{U}_r)] - \sum_{s \in S_j} \Pr[((y_j = 1) \wedge U_s) \mid (\mathcal{A} \wedge \bigwedge_{r \in ([t] - S_j)} \overline{U}_r)].$$

Substituting this into (19) completes the proof. ■

As in the proof of Theorem 3.1, define $S = S_0$ if $y_{S_0}^* \geq y_{S_1}^*$, and $S = S_1$ otherwise; thus, $y_S^* \geq y^*/2$. For all $i \in ([k] - S)$, we remove the pair (s_i, t_i) and all paths $P_{i,j}$ from consideration; the capacity constraints (7) continue to hold, of course. Thus, any index i of an (s_i, t_i) pair, flow path $P_{i,j}$, or r.v.s $z_{i,j}$ and Z_i that we refer to for the rest of Section 3.2, implies that $i \in S$. Our next approximation algorithm, presented in Theorem 3.2, shows how to do a feasible path selection for S with objective function value $\Omega(y_S^*/d) = \Omega(y^*/d)$.

Define, for all (f, i, j) such that $f \in P_{i,j}$,

$$R_f(i, j) \doteq R_f - \rho_i z_{i,j} = \sum_{(r,s) \neq (i,j): f \in P_{r,s}} \rho_r z_{r,s}. \quad (20)$$

Lemma 3.4 For any set X of pairs (i, j) and any set of $\{0, 1\}$ values $\{b_{i,j} \in \{0, 1\} : (i, j) \in X\}$, let \mathcal{A} denote the assignment event that “ $\forall (i, j) \in X, z_{i,j} = b_{i,j}$ ”. Suppose $\Pr[\bigwedge_{f \in E} \overline{E_f} \mid \mathcal{A}] > 0$. Then, for any i , $\mathbf{E}[Z_i \mid (\mathcal{A} \wedge (\bigwedge_{f \in E} \overline{E_f}))]$ is at least

$$\left(\sum_j \Pr[z_{i,j} = 1 \mid \mathcal{A}] (1 - \sum_{f \in P_{i,j}} \Pr[R_f(i, j) > (c_f - \rho_i) \mid \mathcal{A}]) \right) - \sum_{j < j'} \Pr[(z_{i,j} = z_{i,j'} = 1) \mid \mathcal{A}].$$

Proof: Recall that $z_{i,j}$ is the indicator for path $P_{i,j}$ being chosen by the randomized rounding process. For notational convenience, we define, for all j ,

$$g(j) \doteq \Pr[z_{i,j} = 1 \mid (\mathcal{A} \wedge (\bigwedge_{f' \notin P_{i,j}} \overline{E_{f'}}))] - \sum_{f \in P_{i,j}} \Pr[(z_{i,j} = 1) \wedge (R_f > c_f) \mid (\mathcal{A} \wedge (\bigwedge_{f' \notin P_{i,j}} \overline{E_{f'}}))]. \quad (21)$$

By Definition 3.1, $(Z_i = 1) \equiv (\exists j : z_{i,j} = 1)$. By Lemma 3.3,

$$\mathbf{E}[Z_i \mid (\mathcal{A} \wedge (\bigwedge_{f' \in E} \overline{E_{f'}}))] \geq \left(\sum_j g(j) \right) - \sum_{j < j'} \Pr[z_{i,j} = z_{i,j'} = 1 \mid (\mathcal{A} \wedge (\bigwedge_{f' \in E} \overline{E_{f'}}))]. \quad (22)$$

We now proceed to lower-bound (22). Recall that even conditional on \mathcal{A} , our “independence view” lets us take all the $z_{i,j}$ ’s as *independent*: it is just that for all $(i, j) \in X$, $\Pr[z_{i,j} = b_{i,j}] = 1$. Fix j arbitrarily. By the independence view, it is not hard to deduce that

$$\Pr[z_{i,j} = 1 \mid (\mathcal{A} \wedge (\bigwedge_{f' \notin P_{i,j}} \overline{E_{f'}}))] = \Pr[z_{i,j} = 1 \mid \mathcal{A}]. \quad (23)$$

Next, consider any $f \in P_{i,j}$. Once again via our independence view, even conditional on \mathcal{A} , “ $z_{i,j} = 1$ ” and “ $R_f > c_f$ ” are increasing events, while “ $\bigwedge_{f' \notin P_{i,j}} \overline{E_{f'}}$ ” is decreasing. Thus, the independence view and FKG inequality show that

$$\begin{aligned} \Pr[(z_{i,j} = 1) \wedge (R_f > c_f) \mid (\mathcal{A} \wedge (\bigwedge_{f' \notin P_{i,j}} \overline{E_{f'}}))] &\leq \Pr[(z_{i,j} = 1) \wedge (R_f > c_f) \mid \mathcal{A}] \\ &= \Pr[(z_{i,j} = 1) \wedge (R_f(i, j) > c_f - \rho_i) \mid \mathcal{A}]. \end{aligned} \quad (24)$$

Since $R_f(i, j)$ involves only variables other than $z_{i,j}$ (see (20)), the independence view yields

$$\Pr[(z_{i,j} = 1) \wedge (R_f(i, j) > (c_f - \rho_i)) \mid \mathcal{A}] = \Pr[z_{i,j} = 1 \mid \mathcal{A}] \cdot \Pr[R_f(i, j) > (c_f - \rho_i) \mid \mathcal{A}]. \quad (25)$$

The independence view and the FKG inequality show, for $j \neq j'$, that

$$\Pr[z_{i,j} = z_{i,j'} = 1 \mid (\mathcal{A} \wedge (\bigwedge_{f' \in E} \overline{E_{f'}}))] \leq \Pr[z_{i,j} = z_{i,j'} = 1 \mid \mathcal{A}]. \quad (26)$$

Substituting (21), (23), (24), (25) and (26) into (22) completes the proof. \blacksquare

Now, suppose $\gamma \geq 2$ is such that

$$\forall i \in S \forall j \forall f \in P_{i,j}, \Pr[R_f(i,j) > (c_f - \rho_i)] \leq 1/(2d). \quad (27)$$

Then, since $\gamma > 1$, part (i) of Lemma 3.1 shows that $\Pr[E_f] < 1$ for all $f \in E$. So, by the FKG inequality, $\Pr[\bigwedge_{f \in E} \overline{E_f}] \geq \prod_{f \in E} \Pr[\overline{E_f}] > 0$. Recall that $\forall i, \sum_j \Pr[z_{i,j} = 1] = x_i^*/\gamma$. Thus, by Proposition 2.1, $\sum_{j < j'} \Pr[z_{i,j} = z_{i,j'} = 1] \leq (x_i^*/\gamma)^2/2 \leq x_i^*/(4\gamma)$, the last inequality following since $x_i^* \in [0, 1]$ and $\gamma \geq 2$. Next, since each $P_{i,j}$ has at most d edges, we see, by setting $X = \phi$ and \mathcal{A} to be the tautology in Lemma 3.4, that for all $i \in S$,

$$\mathbf{E}[Z_i \mid \bigwedge_{f \in E} \overline{E_f}] \geq \left(\sum_j \Pr[z_{i,j} = 1](1 - d/(2d)) \right) - x_i^*/(4\gamma) = x_i^*/(2\gamma) - x_i^*/(4\gamma) = x_i^*/(4\gamma).$$

Thus, $\mathbf{E}[Z(S) \mid \bigwedge_{f \in E} \overline{E_f}] \geq (\sum_{i \in S} w_i x_i^*)/(4\gamma) = y_S^*/(4\gamma)$. This yields

Corollary 3.2 *Suppose $\gamma \geq 2$ is such that (27) holds. Then, $\Pr[\bigwedge_{f \in E} \overline{E_f}] > 0$, and $\mathbf{E}[Z(S) \mid \bigwedge_{f \in E} \overline{E_f}] \geq y_S^*/(4\gamma) \geq y^*/(8\gamma)$. Thus, there exists a way of rounding the $z_{i,j}$'s so that no edge's capacity is violated, with objective function value at least $y^*/(8\gamma)$.*

We are now led to our next main theorem:

Theorem 3.2 *We can round the $z_{i,j}$'s in deterministic polynomial time to produce a feasible path selection that has objective function value $Z = \sum_{i \in [k]} w_i Z_i \geq c' y^*/d$, for an absolute constant $c' > 0$.*

Proof: There are 2 cases: $S = S_0$ or $S = S_1$. Suppose $S = S_0$. Choose $\gamma = 4d$. For any $i \in S$, $\rho_i \leq 1/2$ since $S = S_0$; also, $c_f \geq 1$ for all $f \in E$, by (17). Thus, for any f, i, j ,

$$\begin{aligned} \Pr[R_f(i,j) > (c_f - \rho_i)] &\leq \Pr[R_f(i,j) > c_f/2] \\ &\leq (2/c_f) \cdot \sum_{(r,s) \neq (i,j): r \in S, f \in P_{r,s}} \rho_r \Pr[z_{r,s} = 1] \quad (\text{Markov's ineq.}) \quad (28) \\ &\leq (2/c_f) \cdot \mathbf{E}[R_f] \leq (2/c_f) \cdot (c_f/\gamma) = 1/(2d). \quad (29) \end{aligned}$$

Thus, Corollary 3.2 implies the *existence* of a rounding that does a feasible path selection for $S = S_0$, with objective function value $\Omega(y^*/d)$. We now show how to turn this into an efficient algorithm, by rounding the $z_{i,j}$'s one-by-one: to this end, we build on the method of conditional probabilities.

Assume inductively, for some set X of pairs (i, j) and some set $Y = \{b_{i,j} \in \{0, 1\} : (i, j) \in X\}$, that we have already rounded $z_{i,j}$ to $b_{i,j}$, $\forall (i, j) \in X$. Let \mathcal{A} denote the event that “ $\forall (i, j) \in X$, $z_{i,j} = b_{i,j}$ ”. ($X = Y = \phi$ when we start our algorithm.) As in the proof of Lemma 3.4, even conditional on \mathcal{A} , we can take the independence view that all the $z_{i,j}$ ’s are independent: it is just that $\forall (i, j) \in X$, $\Pr[z_{i,j} = b_{i,j}] = 1$. Suppose we define $h(\mathcal{A})$ to be

$$\sum_{i \in S} w_i \left(\sum_j \Pr[z_{i,j} = 1] \left(1 - \sum_{f \in P_{i,j}} \frac{2}{c_f} \cdot \sum_{(r,s) \neq (i,j): r \in S, f \in P_{r,s}} \rho_r \Pr[z_{r,s} = 1] \right) - \sum_{j < j'} \Pr[z_{i,j} = 1] \cdot \Pr[z_{i,j'} = 1] \right), \quad (30)$$

with the convention that “ $\forall (i, j) \in X$, $\Pr[z_{i,j} = b_{i,j}] = 1$ ”. Now, (28) holds even when the probabilities in its l.h.s. and r.h.s. are computed conditional on \mathcal{A} . Thus, Lemma 3.4 and our independence view show that

$$\mathbf{E}[Z(S) \mid (\mathcal{A} \wedge (\bigwedge_{f \in E} \overline{E_f}))] \geq h(\mathcal{A}). \quad (31)$$

Our inductive requirements on X and \mathcal{A} are: **(Q1)** $\Pr[\bigwedge_{f \in E} \overline{E_f} \mid \mathcal{A}] > 0$, and **(Q2)** $h(\mathcal{A}) \geq y_S^*/(4\gamma)$. By (28), (29) and the proof of Corollary 3.2, (Q1) and (Q2) are satisfied initially, when $X = \phi$. Given that some \mathcal{A} satisfies (Q1) and (Q2), we now show how to add one or more new elements to X and update \mathcal{A} appropriately, to maintain (Q1) and (Q2). If we have rounded all the $z_{i,j}$ ’s and still maintain (Q1) and (Q2), we see by (31) that we have made a feasible path selection of value at least $y_S^*/(4\gamma)$.

Let $(i, j) \notin X$ be arbitrary, with $i \in S$. Define \mathcal{A}_0 to be “ $\mathcal{A} \wedge (z_{i,j} = 0)$ ”, and \mathcal{A}_1 to be “ $\mathcal{A} \wedge (z_{i,j} = 1)$ ”. For any $f \in E$, define $cd_{\mathcal{A}}(f) \doteq \sum_{(r,s) \in X: f \in P_{r,s}} \rho_r b_{r,s}$. This is the *committed demand* on f : since we have decided to round $z_{r,s}$ to $b_{r,s}$ for each $(r, s) \in X$, this is the demand that we have so far committed on f . Now, if $\exists f \in P_{i,j} : cd_{\mathcal{A}}(f) > c_f - \rho_i$, then clearly,

$$\Pr[\bigwedge_{f \in E} \overline{E_f} \mid \mathcal{A}_1] = 0. \quad (32)$$

Otherwise, suppose that $\forall f \in P_{i,j}$, $cd_{\mathcal{A}}(f) \leq c_f - \rho_i$. Then, since there is a nonzero probability that all the yet-unrounded variables get rounded to 0 ($\forall (r, s)$, $z_{r,s}^*/\gamma < 1$ since $\gamma > 1$),

$$\Pr[\bigwedge_{f \in E} \overline{E_f} \mid \mathcal{A}_0] > 0, \text{ and } \Pr[\bigwedge_{f \in E} \overline{E_f} \mid \mathcal{A}_1] > 0. \quad (33)$$

Our algorithm is now as follows. There are two possibilities:

(i) $\exists(i, j) \notin X : (i \in S) \wedge (\forall f \in P_{i,j}, cd_{\mathcal{A}}(f) \leq c_f - \rho_i)$. Now,

$$h(\mathcal{A}) = \mathbf{E}[(\sum_{i' \in S} w_{i'} \sum_{j'} z_{i',j'} (1 - \sum_{f \in P_{i',j'}} (2/c_f) \cdot \sum_{(r,s) \neq (i',j'): r \in S, f \in P_{r,s}} \rho_r z_{r,s})) - \sum_{i' \in S} w_{i'} \sum_{j' < j''} z_{i',j'} \cdot z_{i',j''}] \quad (34)$$

by the independence view. Thus, $h(\mathcal{A})$ is a convex combination of $h(\mathcal{A}_0)$ and $h(\mathcal{A}_1)$: so there exists an *efficiently computable* $u \in \{0, 1\}$ such that $h(\mathcal{A}_u) \geq h(\mathcal{A})$. Furthermore, $\Pr[\bigwedge_{f \in E} \overline{E_f} \mid \mathcal{A}_u] > 0$, from (33). Thus, adding (i, j) to X and rounding $z_{i,j}$ to u (i.e., setting $\mathcal{A} := \mathcal{A}_u$), will maintain (Q1) and (Q2). We next consider case (ii), the complement of case (i):

(ii) $\forall(i, j) \notin X : (i \in S) \rightarrow \exists f \in P_{i,j}, cd_{\mathcal{A}}(f) > c_f - \rho_i$. Then, by (32), rounding any yet-unrounded (i, j) to 1 will make E_f true for some $f \in E$. But, we are assured by (31), (Q1) and (Q2) that there exists a rounding of the yet-unrounded variables, such that $\overline{E_f}$ holds for all $f \in E$ and such that $Z(S) \geq y_S^*/(4\gamma)$. Thus, the only possible valid rounding for all remaining variables is to set them to 0, and this is guaranteed to be a feasible path selection ($\overline{E_f}$ holds for all $f \in E$) with objective function value at least $y_S^*/(4\gamma)$. Thus, in both cases (i) and (ii), we can successfully keep rounding the variables; this is clearly a deterministic polynomial-time algorithm that outputs a feasible path selection of value $\Omega(y_S^*/d)$.

Remark: In case (i) above, suppose $\Pr[z_{i,j} = 0] = 1$. Then, $h(\mathcal{A}) = h(\mathcal{A}_0)$, and so our proof above shows that we may just round $z_{i,j}$ to zero. Similarly, if $\Pr[z_{i,j} = 1] = 1$, then we may round $z_{i,j}$ to one, without decreasing the value of $h(\mathcal{A})$. It is also easy to see that if we do such an obvious rounding to all the variables $z_{i,j}$ such that $\Pr[z_{i,j} = 1] \in \{0, 1\}$ *before* rounding any other variable, we will violate neither (Q1) nor (Q2). Thus, our algorithm remains correct if we started by rounding all variables $z_{i,j}$ such that $\Pr[z_{i,j} = 1] \in \{0, 1\}$: those with $\Pr[z_{i,j} = 1] = 0$ are rounded to zero, and those with $\Pr[z_{i,j} = 1] = 1$ are rounded to one.

We next consider the case $S = S_1$. We now choose $\gamma = 8d$. Consider any f, i, j . If $c_f \geq 2$, then since $\rho_i \leq 1$ by (17), we get, exactly as in (28) and (29), that

$$\Pr[R_f(i, j) > (c_f - \rho_i)] \leq (2/c_f) \cdot \sum_{(r,s) \neq (i,j): r \in S, f \in P_{r,s}} \rho_r \Pr[z_{r,s} = 1] \leq (2/c_f) \cdot (c_f/\gamma) = 1/(4d).$$

On the other hand, if $c_f < 2$, we have $\rho_i \in (1/2, 1]$ and $c_f \in [1, 2)$ since $i \in S_1$ and from (17). Thus, by the definition of $R_f(i, j)$, we see that if $R_f(i, j) > 0$, then $R_f(i, j) > 1/2$. So

$$\Pr[R_f(i, j) > (c_f - \rho_i)] \leq \Pr[R_f(i, j) > 1/2]$$

$$\begin{aligned}
&\leq \Pr[R_f(i, j) > c_f/4] \\
&\leq (4/c_f) \cdot \sum_{(r,s) \neq (i,j): r \in S, f \in P_{r,s}} \rho_r \Pr[z_{r,s} = 1] \quad (\text{Markov's inequality}) \\
&\leq (4/c_f) \cdot (c_f/\gamma) = 1/(2d).
\end{aligned}$$

Hence, irrespective of the value of c_f , we have

$$\Pr[R_f(i, j) > (c_f - \rho_i)] \leq (4/c_f) \cdot \sum_{(r,s) \neq (i,j): r \in S, f \in P_{r,s}} \rho_r \Pr[z_{r,s} = 1] \leq 1/(2d).$$

Thus, by replacing the term “ $(2/c_f)$ ” in the expression (34) by the term “ $(4/c_f)$ ”, we can apply our above approach for the case $S = S_0$, to the case $S = S_1$. \blacksquare

3.3 The case of small demands

We now consider situations where the demands are noticeably smaller than the capacities.

Theorem 3.3 *Suppose $\exists \epsilon \in [1/2, 1) \forall i \forall f, \rho_i \leq (1 - \epsilon)c_f$. Then we can efficiently do a feasible path selection for the UFP, with objective function value $\Omega(y^*/d^{1/\epsilon-1})$.*

Proof: Take $\gamma = e(2d)^{1/\epsilon-1}/\epsilon$ and conduct a γ -rounding. We now take $S = [k]$, i.e., we now consider all the given (s_i, t_i) pairs. Fix any (f, i, j) . We have $\Pr[R_f(i, j) > c_f - \rho_i] \leq \Pr[R_f(i, j) > \epsilon c_f]$ since $\rho_i \leq (1 - \epsilon)c_f$. We now mimic our proof approach for part (ii) of Lemma 3.1 to bound $\Pr[R_f(i, j) > \epsilon c_f]$. Define $\rho'_{r,f} = \rho_r / ((1 - \epsilon)c_f)$ and $R'_f(i, j) \doteq \sum_{(r,s) \neq (i,j): f \in P_{r,s}} \rho'_{r,f} z_{r,s}$. $R'_f(i, j)$ is a sum of independent r.v.s, each taking values in $[0, 1]$; by (9), $\mu_{f,i,j} \doteq \mathbf{E}[R'_f(i, j)]$ is at most $((1 - \epsilon)\gamma)^{-1}$. Define $\delta_{f,i,j} > 0$ by $\mu_{f,i,j}(1 + \delta_{f,i,j}) = \epsilon/(1 - \epsilon)$. Thus, by part (i) of Theorem 2.1,

$$\begin{aligned}
\Pr[R_f(i, j) > \epsilon c_f] &= \Pr[R'_f(i, j) > \mu_{f,i,j}(1 + \delta_{f,i,j})] \\
&\leq \frac{\mathbf{E}[\prod_{(r,s) \neq (i,j): f \in P_{r,s}} (1 + \delta_{f,i,j})^{\rho'_{r,f} z_{r,s}}]}{(1 + \delta_{f,i,j})^{\mu_{f,i,j}(1 + \delta_{f,i,j})}} \\
&\leq G(\mu_{f,i,j}, \delta_{f,i,j}) \\
&\leq G(((1 - \epsilon)\gamma)^{-1}, \gamma\epsilon - 1) \quad (\text{Lemma 2.1, part (i)}) \\
&\leq (e/(\epsilon\gamma))^{\epsilon/(1-\epsilon)} = 1/(2d).
\end{aligned}$$

Thus, by Corollary 3.2, there exists a rounding that leads to a feasible path selection of value $\Omega(y^*/\gamma) = \Omega(y^*/d^{1/\epsilon-1})$.

The above calculation also shows how this can be made constructive by a small modification to the proof of Theorem 3.2. Suppose, as in the proof of Theorem 3.2, \mathcal{A} denotes the event that “ $\forall(i, j) \in X$, $z_{i,j} = b_{i,j}$ ” (for some setting of X and the $b_{i,j}$ ’s). In place of (30) and (34), we now define

$$h(\mathcal{A}) = \mathbf{E} \left[\left(\sum_{i \in S} w_i \sum_j z_{i,j} \left(1 - \sum_{f \in P_{i,j}} \frac{\prod_{(r,s) \neq (i,j): f \in P_{r,s}} (1 + \delta_{f,i,j})^{\rho'_{r,f} z_{r,s}}}{(1 + \delta_{f,i,j})^{\mu_{f,i,j} (1 + \delta_{f,i,j})}} \right) \right) - \sum_{i \in S} w_i \sum_{j < j'} z_{i,j} \cdot z_{i,j'} \right],$$

with the convention that “ $\forall(i, j) \in X$, $\Pr[z_{i,j} = b_{i,j}] = 1$ ”. Note that $h(\mathcal{A})$ is efficiently computable, by the independence view; it equals

$$\left(\sum_{i \in S} w_i \sum_j \mathbf{E}[z_{i,j}] \left(1 - \sum_{f \in P_{i,j}} \frac{\prod_{(r,s) \neq (i,j): f \in P_{r,s}} \mathbf{E}[(1 + \delta_{f,i,j})^{\rho'_{r,f} z_{r,s}}]}{(1 + \delta_{f,i,j})^{\mu_{f,i,j} (1 + \delta_{f,i,j})}} \right) \right) - \sum_{i \in S} w_i \sum_{j < j'} \mathbf{E}[z_{i,j}] \cdot \mathbf{E}[z_{i,j'}],$$

under the convention that “ $\forall(i, j) \in X$, $\Pr[z_{i,j} = b_{i,j}] = 1$ ”. We can now follow the proof of Theorem 3.2. ■

Note next that for any parameter D , it is easy to add the constraint that “all flow paths are of length at most D ” to our LP relaxation; we now present a few applications based on this.

(i) Bandwidth requirement for expander graphs. Recall that a fractional flow is feasible for us iff the total demand using any edge f is at most c_f . Suppose all capacities c_f are the same (say, 1, by scaling); i.e., suppose we have a UCUFP instance. Then, it is shown in Kleinberg and Rubinfeld (1996) that if there is a feasible fractional flow shipping a flow of f_i ($\in [0, \rho_i]$) from s_i to t_i for each i , then there exists a feasible fractional flow shipping a flow of at least $f_i/2$ for each i , such that all flow paths are of length at most $d_0 \doteq a\Delta^2\beta^{-2}\log^3 n$, for some absolute constant a . Suppose we add the constraint that “all flow paths are of length at most d_0 ” to our LP relaxation; an optimal flow thus produced is guaranteed to have objective function value $y^* \geq \alpha^*(\mathcal{T})/2$, by the above result of Kleinberg and Rubinfeld (1996). Thus, we may replace d by d_0 in Theorems 3.2 and 3.3, leading to the applications discussed in §1. In particular, for hypercubic networks such as the butterfly, there are $O(\text{polylog}(n))$ approximations for the UCUFP, and $\Omega(\log \log n)$ edge bandwidth suffices to derive $O(1)$ approximations. Appropriate improvements also hold for other networks with “moderate” d_0 , such as meshes of dimension 3 or more.

We remark that even the *single-source* version of the UFP (wherein all the source vertices are the same) is an interesting NP-hard problem (Kleinberg (1996)); good constant-factor approximation algorithms have been developed for it (Kleinberg (1996), Kolliopoulos and Stein (1997)).

(ii) Routing within limited time. Recall that store-and-forward routing (*packet-switching*) is a common alternative to circuit switching. Here, packets move from source to destination through the network G , with each edge traversal taking unit time and each edge carrying at most one packet per time step. Edges of G are allowed to queue packets. Suppose we are given an arbitrary network G and some collection $\mathcal{T} = \{(s_i, t_i) : 1 \leq i \leq k\}$ of vertex pairs. The i th packet has some given weight $w_i \geq 0$ and needs to be routed from s_i to t_i . Given a time bound T , what is a maximum-weight collection of packets that can be routed in store-and-forward fashion, within time T ? We now use Theorem 3.3 to develop a bicriteria approximation algorithm for this problem; since the case $T = 1$ is trivial, we assume that $T \geq 2$.

Our first task is to find a suitably large-weight $\mathcal{T}' \subseteq \mathcal{T}$ and to identify one (s_i, t_i) -path (routing path) for each $(s_i, t_i) \in \mathcal{T}'$. Now, whatever paths the packets use in any solution to the problem, their *congestion* (maximum number of paths using any given edge) and *dilation* (maximum length of any of the routing paths) must clearly be at most T . So, if we view the problem as a UFP instance with all demands being one, all edge-capacities being T and all flow-paths of length at most T , the optimal value y^* of the corresponding LP relaxation will be an upper bound on the optimal value OPT for the given problem. In the LP relaxation, we equivalently view the problem as a UCUFP instance with all demands being $1/T$. In the notation of Theorem 3.3, we $\epsilon = 1 - 1/T$ and $d \leq T$; thus by Theorem 3.3, we can efficiently do a path-selection with objective function value $\Omega(y^*/T^{1/(T-1)})$. Since $T \geq 2$, this quantity is $\Omega(y^*)$, i.e., $\Omega(OPT)$.

The congestion and dilation of the paths chosen are bounded by T . We now use results of Leighton, Maggs and Rao (1994) and Leighton, Maggs and Richa (1996) to convert such a path-selection into an $O(T)$ -length routing schedule using constant-sized queues. Thus we get

Corollary 3.3 *There are absolute constants $c_0 < 1$ and $c_1, c_2 > 1$ such that the following holds. For any instance of the “routing within limited time” problem with time bound T and optimal solution weight OPT , we can efficiently route a collection of packets of total weight at least $c_0 \cdot OPT$ within $c_1 T$ time, using c_2 -sized edge-queues.*

(iii) Routing along short paths. In practice, it is often required that no routing path have length more than a small constant D such as 5, 6 etc. As seen above, we can add such constraints to our LP; thus, if the specified D is a constant, Theorem 3.2 provides a *constant-factor* approximation algorithm for the corresponding UFP instances (this application does not require the demands to

be noticeably smaller than the capacities).

4 Routing in Rounds

Given a network G with unit edge-capacities and a collection $\mathcal{T} = \{(s_i, t_i) : i \in [k]\}$ of request pairs with (s_i, t_i) having a demand $\rho_i \in [0, 1]$, we now consider the problem of “routing in rounds” introduced in §1; we present improved approximation algorithms and bounds for $\chi(\mathcal{T})$. Let $\chi^*(\mathcal{T})$ be the smallest positive real such that $\forall i$, a fractional flow of ρ_i can be shipped from s_i to t_i with each edge carrying a total flow of at most $\chi^*(\mathcal{T})$. It is easily seen that $\chi^*(\mathcal{T}) \leq \chi(\mathcal{T})$; we thus have an LP relaxation and hence, $\chi^*(\mathcal{T})$ and the corresponding fractional flow can be found in polynomial time. Let $y^* \geq \chi^*(\mathcal{T})$ be some real such that we have a fractional flow \mathcal{F} routing ρ_i units of flow for each i , with the total flow on any edge being at most y^* . Let the flow-paths in \mathcal{F} for (s_i, t_i) be denoted $P_{i,1}, P_{i,2}, \dots$, with $P_{i,j}$ carrying a flow of $\rho_i z_{i,j}^*$ as before (thus, $\sum_j z_{i,j}^* = 1$). Also as before, we let d denote the length of a longest flow path in \mathcal{F} .

As mentioned in the introduction, results of Aumann and Rabani (1995) and Kleinberg (1996) show that a ψ -approximation for $\alpha(\mathcal{T})$ translates to an $O(\psi \log n)$ -approximation for $\chi(\mathcal{T})$. Thus, our results for the UFP directly lead to $O(d \log n)$ and $O(\sqrt{m} \log n)$ approximations for $\chi(\mathcal{T})$, the former again implying a polylog(n) approximation, e.g., for the butterfly. However, Theorems 4.1 and 4.2 present further improved approximations for $\chi(\mathcal{T})$.

The following lemma will be useful in proving Theorems 4.1 and 4.2:

Lemma 4.1 *Let $\delta \in (0, 1/2]$ be arbitrary. In deterministic polynomial time, we can do a feasible “routing in rounds” of all (s_i, t_i) such that $\rho_i \leq \delta$, using $O(\chi(\mathcal{T})d^{\delta/(1-\delta)} \log n)$ rounds.*

Proof: It is known that all pairs (s_i, t_i) with $\rho_i \leq 1/\log n$ can be routed in $O(\chi(\mathcal{T}))$ rounds in deterministic polynomial time (Raghavan and Thompson (1987), Raghavan (1988)). We route all these “low demand” pairs first. We next consider $\mathcal{P}_0 = \{(s_i, t_i) : \rho_i \in [1/\log n, \delta]\}$. A result of Kleinberg (1996) shows that we can turn a ν -approximation algorithm for $\alpha(\mathcal{T})$ efficiently into an $O(\nu(\log n + \log(\max_i \rho_i^{-1})))$ -approximation algorithm for routing in rounds. Since $\rho_i \leq \delta$ if $(s_i, t_i) \in \mathcal{P}_0$, we see from Theorem 3.3 that $\alpha(\mathcal{P}_0)$ can be approximated to within $O(d^{\delta/(1-\delta)})$. As $\max_{i: (s_i, t_i) \in \mathcal{P}_0} \rho_i^{-1} = O(\log n)$, we can efficiently route all of \mathcal{P}_0 in $O(\chi(\mathcal{T})d^{\delta/(1-\delta)} \log n)$ rounds. ■

Our main results are as follows:

Theorem 4.1 *We can do a feasible “routing in rounds” in deterministic polynomial time, the number of rounds being: (i) $O(\chi(\mathcal{T})d^\delta \log n + d(y^* + \log n))$, for arbitrary \mathcal{T} and for any desired fixed $\delta \in (0, 1)$, and (ii) $O(\eta^{-1}d(y^* + \log n))$, if $\forall i, \rho_i \geq \eta$.*

Proof: We prove part (ii) first. Suppose we conduct the “standard” randomized rounding of Raghavan and Thompson (1987): independently for each i , we pick **exactly one** $P_{i,j}$, the probability of picking $P_{i,j}$ being $z_{i,j}^*$. (Note that this is different from our randomized rounding approach of §3.) Then, the results of Raghavan and Thompson (1987) and Raghavan (1988) yield an efficient derandomization of this that ensures that the total demand using any edge is at most $O(y^* + \log n)$. Thus, since $\rho_i \geq \eta$ for all i , each edge is used by at most $O(\eta^{-1}(y^* + \log n))$ paths. Now, we follow an idea from Aumann and Rabani (1995), and Aggarwal et al. (1996). Construct an undirected graph G' with the chosen paths as nodes and with two nodes being adjacent in G' iff the corresponding paths have a common edge in G . Since each chosen path can meet at most $O(\eta^{-1}(y^* + \log n))$ paths per edge, and it has at most d edges, the maximum degree Δ' of G' is $O(\eta^{-1}d(y^* + \log n))$. Thus, we can efficiently color the vertices of G' using at most $\Delta' + 1 = O(\eta^{-1}d(y^* + \log n))$ colors: in other words, we can do a feasible routing in rounds using at most $O(\eta^{-1}d(y^* + \log n))$ rounds.

We now prove part (i). Let $\delta' \in (0, 1/2]$ be the constant such that $\delta'/(1 - \delta') = \delta$. Thus, by Lemma 4.1, we can efficiently do a feasible “routing in rounds” of all (s_i, t_i) such that $\rho_i \leq \delta'$, using $O(\chi(\mathcal{T})d^\delta \log n)$ rounds. Next, all (s_i, t_i) such that $\rho_i > \delta'$ can be handled as in part (ii). Thus, the total number of rounds is $O(\chi(\mathcal{T})d^\delta \log n + d(y^* + \log n))$. ■

A flow \mathcal{F} which has $y^* \leq 2\chi^*(\mathcal{T})$ and $d = O(\Delta^2\beta^{-2}\log^3 n)$ can be efficiently computed, using the work of Kleinberg and Rubinfeld (1996). Thus, we get

Corollary 4.1 *Theorem 4.1 remains true if we replace y^* by $\chi^*(\mathcal{T})$ and d by $\Delta^2\beta^{-2}\log^3 n$.*

Our next theorem shows an approximation independent of d :

Theorem 4.2 *We can do a feasible “routing in rounds” using $O(\chi(\mathcal{T}) \cdot \sqrt{m(1 + (\log n)/\chi(\mathcal{T}))})$ rounds, in deterministic polynomial time.*

Proof: Let us start with an optimal solution to the LP relaxation, and let d be the length of a longest path among the flow paths $\{P_{i,j}\}$. Setting $\delta = 1/4$ in Lemma 4.1 and observing that $d \leq m$ w.l.o.g., we first route all (s_i, t_i) with $\rho_i \leq 1/4$, using $O(\chi(\mathcal{T})m^{1/3} \log n)$ rounds.

Next, consider $\mathcal{P}_1 = \{(s_i, t_i) : \rho_i \in (1/4, 1]\}$. We first weed out very long paths; this is “filtering” in the usage of Lin and Vitter (1992). (The basic idea is to take each i such that $(s_i, t_i) \in \mathcal{P}_1$, and let D_i be the total amount of flow for the pair (s_i, t_i) . Then, one can see that the total flow on paths $P_{i,j}$ of length more than $2D_i/\rho_i$ is “small”. Hence, we can effectively ignore such paths.) Formally, we let $|P_{i,j}|$ denote the number of edges in path $P_{i,j}$; define, for each i such that $(s_i, t_i) \in \mathcal{P}_1$, $D_i \doteq \rho_i \sum_j z_{i,j}^* |P_{i,j}|$. D_i is thus the *total* amount of flow for the pair (s_i, t_i) , summed over all edges. Call a flow-path $P_{i,j}$ *long* iff $|P_{i,j}| > 2D_i/\rho_i$. Since the $z_{i,j}^*$ ’s are non-negative and $\sum_j z_{i,j}^* = 1$, we see that for any i , $h_i \doteq \sum_{j: P_{i,j} \text{ long}} z_{i,j}^* < 1/2$. We now set $z_{i,j}^{**} := 0$ for all *long* $P_{i,j}$, and re-scale all the $z_{i,j}^*$ corresponding to non-long $P_{i,j}$ by setting $z_{i,j}^{**} := z_{i,j}^*/(1 - h_i)$. Since $h_i < 1/2$, we see that $z_{i,j}^{**} \leq 2z_{i,j}^*$. Hence, the maximum edge-congestion now is at most $2\chi^*(\mathcal{T})$; crucially, all remaining flow paths are non-long.

For a parameter x to be chosen soon, partition \mathcal{P}_1 into two sets: (a) pairs (s_i, t_i) , all of whose flow paths now are of length at most x , and (b) pairs (s_i, t_i) that have at least one flow path of length more than x . By Theorem 4.1(ii), the pairs in (a) can be routed efficiently in $r_1 = O((\chi(\mathcal{T}) + \log n)x)$ rounds. Next, since each remaining flow path is non-long, any i for which (s_i, t_i) falls into category (b) must have $x \leq 2D_i/\rho_i$; thus, $D_i \geq x\rho_i/2 \geq x/8$. Recall that D_i is the total flow contributed by (s_i, t_i) over all edges, and that the edge-congestion, summed over all edges, is at most $2m\chi^*(\mathcal{T}) \leq 2m\chi(\mathcal{T})$. Thus, the number of pairs in category (b) can be at most $r_2 = 16m\chi(\mathcal{T})/x$; so they can trivially be routed within $\lceil r_2 \rceil$ rounds. Now, to balance r_1 and r_2 , we pick $x = \sqrt{m\chi(\mathcal{T})/(\chi(\mathcal{T}) + \log n)}$. Thus, \mathcal{P}_1 can be routed efficiently in $r_1 + r_2 = O(\chi(\mathcal{T})\sqrt{m \cdot (1 + (\log n)/\chi(\mathcal{T}))})$ rounds. ■

5 Extensions and applications

5.1 Column-restricted packing problems, vertex-disjoint paths

In recent interesting work, Kolliopoulos and Stein (1998) have studied a family of NP-hard integer linear programs, which they call *column-restricted packing integer programs* (henceforth referred to as CIPs). The variables in such a problem are x_1, x_2, \dots, x_N ; the objective is to maximize $\sum_j w_j x_j$ subject to: (i) a system of m linear constraints $Ax \leq b$, and (ii) integrality constraints $x_j \in \{0, 1, \dots, d_j\}$ for each j (some of the given integers d_j could equal infinity). We have furthermore that: **(C1)** all the w_j ’s lie in $[0, 1]$ (scaled w.l.o.g. so that $\max_j w_j = 1$); **(C2)** all entries in A

and b are non-negative, with $A_{i,j} \leq b_i$ for all i, j , and, importantly, **(C3)** all nonzero entries in any given column of A are the same (hence the name *column-restricted*). A key parameter will be d , the maximum number of nonzero entries in any column of A .

By presenting approximation algorithms for CIPs, Kolliopoulos and Stein (1998) show, among other results, that a fractional solution to a UFP instance can be rounded to an integral solution of value

$$\Omega(\max\{\frac{(y^*)^2}{m \log^3 m}, \frac{y^*}{\sqrt{m} \log m}, \frac{y^*}{d}\}), \quad (35)$$

in the case where $y^* \leq m$. This bound is slightly weaker than our $\Omega(\max\{(y^*)^2/m, y^*/\sqrt{m}, y^*/d\})$ bound given by Theorems 3.1 and 3.2 and Corollary 3.1, for the case where $y^* \leq m$.

We start with Theorem 5.1 that generalizes Theorems 3.1 and 3.2 and Corollary 3.1; we will then use it to present an approximation algorithm for CIPs. Given an event E , let $\mathbf{Ind}(E)$ denote its *indicator*: 1 if E occurred, and 0 otherwise. We will also need a small extension of **(C1)** and **(C3)**: suppose the variables are indexed as $z_{u,v}^*$, instead of as x_j . Let z^* be the vector of the $z_{u,v}^*$'s (arranged in any fixed order). Then, two useful properties of the vector w and the linear system $Az^* \leq b$ will be: **(C1')** for each u , the values $w_{u,1}, w_{u,2}, \dots$ are all the same, and are denoted by w_u ; all the w_u 's lie in $[0, 1]$ (scaled so that $\max_u w_u = 1$); **(C3')** for each u , all nonzero entries in the columns of A that are associated with $z_{u,1}^*, z_{u,2}^*, \dots$, are the same.

Theorem 5.1 *Suppose we have non-negative reals $\{z_{u,v}^* : u \in [k], v \in [\ell_u]\}$ where the ℓ_u are some given integers. Let b, w be vectors and A be an m -row matrix that satisfy **(C1')**, **(C2)**, and **(C3')**; let d be the maximum number of nonzero entries in any column of A . Suppose we also have **(D1)** $\forall u, \sum_v z_{u,v}^* \leq 1$ and **(D2)** $Az^* \leq b$, where z^* is the vector of the $z_{u,v}^*$'s (arranged in any fixed order). Let $y^* = \sum_u w_u \sum_v z_{u,v}^*$. Then, we can efficiently round all the $z_{u,v}^*$'s (each $z_{u,v}^*$ is rounded to some $z_{u,v} \in \{0, 1\}$) such that **(D1)** and **(D2)** are satisfied, **and** with $\sum_u w_u \mathbf{Ind}(\bigvee_v (z_{u,v} = 1))$ being at least: (i) $\Omega(y^*)$ if $y^* \geq m$, and (ii) $\Omega(\max\{(y^*)^2/m, y^*/\sqrt{m}, y^*/d\})$ if $y^* < m$.*

Proof: The proof is essentially identical to those of Theorems 3.1 and 3.2 and Corollary 3.1; so we merely give a sketch. As before, the basic step is as follows: for an appropriate $\gamma > 1$, set $z_{u,v}$ to 1 with probability $z_{u,v}^*/\gamma$ (and $z_{u,v} := 0$ with probability $1 - z_{u,v}^*/\gamma$), independently for all u and v . For each u , we then do the following: if more than one $z_{u,v}$ is 1, one of them is chosen arbitrarily and all others are re-set to 0.

We then proceed exactly the same way as before: scale the entries of A and b so that $\min_i b_i = 1$. Let $\rho_u \in [0, 1]$ be the “common nonzero value” referred to in (C3’). As in the proofs of Theorems 3.1 and 3.2, we partition $[k]$ into $S_0 = \{u \in [k] : \rho_u \leq 1/2\}$ and $S_1 = \{u \in [k] : \rho_u > 1/2\}$. For $T \subseteq [k]$, let $y_T^* \doteq \sum_{u \in T} w_u \sum_v z_{u,v}^*$. Define $S = S_0$ if $y_{S_0}^* \geq y_{S_1}^*$, and $S = S_1$ otherwise; thus, $y_S^* \geq y^*/2$. For all $u \in ([k] - S)$, we remove all variables $z_{u,v}^*$ from consideration (i.e., drive them all down to 0); we will now focus on rounding $z_{u,v}^*$ for all $u \in S$ and for all $v \in [\ell_u]$.

The proof of Theorem 3.1 directly implies that by choosing $\gamma = \Theta(\max\{m/y^*, 1\})$, we can efficiently round so that the objective function, $\sum_{u \in S} w_u \mathbf{Ind}(\bigvee_j (z_{u,v} = 1))$, is $\Omega((y^*)^2/m)$. The proof of Corollary 3.1 can then be invoked to show a rounding with objective function value $\Omega(y^*/\sqrt{m})$. Finally, we note that the proof of Theorem 3.2 does not require the underlying path-edge structure, but simply requires properties (C1’), (C2) and (C3’). Thus, we can also efficiently round to objective function value $\Omega(y^*/d)$. \blacksquare

The *vertex-disjoint paths* problem is essentially the same as the MDP, except that no two paths chosen must intersect at a *vertex*. The integer programming formulation and LP relaxation for this problem is very much similar to that for the MDP; we now have n constraints, where n is the number of vertices in the underlying graph G (one constraint per vertex). Let y^* and d respectively denote the value of an optimal solution to the LP relaxation and the length of a longest (s_i, t_i) -flow path in such a solution. Kolliopoulos and Stein (1998) show how to round such a fractional solution into an integral solution of value $\Omega(\max\{(y^*)^2/n, y^*/\sqrt{n}, y^*/d\})$. Since the problem of rounding the fractional flow-paths is modeled by Theorem 5.1, Theorem 5.1 also shows such a result.

Given an instance of a CPIP, let y^* denote the optimal value of its LP relaxation, which is obtained by relaxing the integrality constraints to “ $\forall j, x_j \in [0, d_j]$ ”. We now apply Theorem 5.1 to get

Corollary 5.1 *The fractional optimum of a given CPIP can be efficiently rounded to an integral solution of value: (i) $\Omega(y^*)$ if $y^* \geq m$, and (ii) $\Omega(\max\{(y^*)^2/m, y^*/\sqrt{m}, y^*/d\})$ if $y^* < m$.*

Proof: Let $\{x_1^*, \dots, x_N^*\}$ be an optimal solution computed for the LP relaxation of the given CPIP. Partition $[N]$ into $S_0 = \{j : x_j^* \geq 1\}$ and $S_1 = \{j : x_j^* < 1\}$, and, for $i = 1, 2$, let $y_i^* = \sum_{j \in S_i} w_j x_j^*$. Now, if $y_0^* \geq y^*/2$, we could simply set $x_j := \lfloor x_j^* \rfloor$ for all $j \in S_0$ and $x_j := 0$ for all $j \in S_1$, to get an integral solution of value at least $y_0^*/2 = \Omega(y^*)$. So, suppose $y_1^* > y^*/2$; we set $x_j := 0$ for all $j \in S_0$, and focus on rounding $\{x_j^* : j \in S_1\}$. We will show how to do this to get an integral solution

of value: (i) $\Omega(y_1^*)$ if $y_1^* \geq m$, and (ii) $\Omega(\max\{(y_1^*)^2/m, y_1^*/\sqrt{m}, y_1^*/d\})$ if $y^* < m$. (This will clearly give us the desired result.) Such a solution directly follows from Theorem 5.1: our problem now is in fact a CPIP instance with $\ell_u = 1$ for all u . ■

5.2 Improved approximations for a class of packing problems

Packing integer programs (PIPs) which subsume CPIPs, satisfy (C1) and (C2) but not necessarily (C3). PIPs model numerous problems in combinatorial optimization such as independent sets and matchings in graphs and hypergraphs, certain classes of scheduling problems etc. We adopt the terminology used for CPIPs, for PIPs also. A useful parametrization is

$$B = \min_{i,j: A_{i,j} \neq 0} b_i/A_{i,j}; \quad (36)$$

note that $B \geq 1$. By scaling A and b uniformly, we assume that $\max_{i,j} A_{i,j} = 1$ and $\min_i b_i = B$.

The current-best results for PIPs (Srinivasan (1995), Srinivasan (1996)) are that for a certain absolute constant $K_0 \in (0, 1)$, we can efficiently compute an integral solution of value

$$\Omega(y^*) \text{ if } y^* \geq m, \text{ and } \Omega(\max\{y^*(K_0 y^*/m)^{1/(B-1)}, y^*(K_0/d)^{1/(B-1)}\}) \text{ otherwise}; \quad (37)$$

these papers have also shown that we can replace $(B - 1)$ by B in (37), if A has only zeroes and ones. Kolliopoulos and Stein (1998) have generalized this by showing that $(B - 1)$ can be replaced by $\lfloor B \rfloor$ in (37) if A satisfies (C3); note that A satisfying (C3), i.e., A being column-restricted, is a less stringent requirement than that A has only zeroes and ones.

Returning to general PIPs, note that for “large” values of B —in particular, if B grows as $\Omega(\log d)$ —the approximation bound of (37) is reasonable. However, the $\Omega(\cdot)$ notation in (37) hides a constant $a_0 > 1$; thus, even if B is arbitrarily large, the bound (37) does not imply $(1 + \epsilon)$ -approximations for an arbitrarily small $\epsilon > 0$. This leads us to the next question we address: can we approximate PIP instances by a factor arbitrarily close to 1, if B is sufficiently large? We answer this now in the affirmative, via the approach we used for proving Theorem 3.2.

Theorem 5.2 *There is an absolute constant $a_1 > 0$ such that for any parameter $0 < \epsilon \leq 1$, PIP instances with $B \geq (a_1/\epsilon^2) \ln(2d/\epsilon)$ can be approximated to within $1 + \epsilon$.*

Proof: We first simplify things by making all entries of b equal B : multiply the i th constraint of “ $Ax \leq b$ ” by $B/b_i \leq 1$. (Note that all the $A_{i,j}$ continue to lie in $[0, 1]$ after this.) Thus we

assume $b_i = B$ for all i . Start with an optimal fractional solution $\{x_1^*, \dots, x_N^*\}$ to the given PIP, and choose $\gamma = 1 + \epsilon/3$. For $j = 1, 2, \dots, N$, define independent random variables $z_j \in \{0, 1\}$, with $\Pr[z_j = 1] = x_j^*/\gamma - \lfloor x_j^*/\gamma \rfloor$ (and with $\Pr[z_j = 0] = 1 - \Pr[z_j = 1]$). Next define $x_j = \lfloor x_j^*/\gamma \rfloor + z_j$; note that $\mathbf{E}[x_j] = x_j^*/\gamma$. This is our notion of γ -rounding now, which we proceed to analyze. Let E_i be the bad event that $\sum_j A_{i,j}x_j > B$. Since $\gamma > 1$, $\Pr[E_i] < 1$. Since each $\overline{E_i}$ is decreasing as a function of the z_j , the FKG inequality shows as before that $\Pr[\bigwedge_{i \in [m]} \overline{E_i}] \geq \prod_{i \in [m]} \Pr[\overline{E_i}] > 0$.

Next, define $s_i = \sum_j A_{i,j} \lfloor x_j^*/\gamma \rfloor$ for each $i \in [m]$; note that E_i is the event “ $\sum_j A_{i,j}z_j > (B - s_i)$ ”. For each j , let $S(j)$ denote the set of constraints (in the constraint set $Ax \leq b$) in which x_j appears with a nonzero coefficient. In exactly the same way as was proved for Lemma 3.4, we have, setting $X = \phi$ and \mathcal{A} to be the tautology in the statement of Lemma 3.4,

$$\begin{aligned} \mathbf{E}[z_j \mid (\bigwedge_{i \in [m]} \overline{E_i})] &\geq \Pr[z_j = 1] \cdot (1 - \sum_{i \in S_j} \Pr[\sum_{j' \neq j} A_{i,j'}z_{j'} > (B - s_i - A_{i,j})]) \\ &\geq \Pr[z_j = 1] \cdot (1 - \sum_{i \in S_j} \Pr[\sum_{j' \neq j} A_{i,j'}z_{j'} > (B - s_i - 1)]). \end{aligned} \quad (38)$$

Fix j , and $i \in S_j$. We first upper-bound the probability of $\mathcal{E}_i \equiv (\sum_{j' \neq j} A_{i,j'}z_{j'} > (B - s_i - 1))$. Define $\mu_i = \mathbf{E}[\sum_{j' \neq j} A_{i,j'}z_{j'}]$. Since $Ax^* \leq b$, it can be seen that $s_i + \mu_i \leq B/\gamma$. If $\mu_i = 0$, then $\Pr[\mathcal{E}_i] = 0$; so suppose $s_i < B/\gamma$ and $\mu_i > 0$. Define δ_i by $\mu_i(1 + \delta_i) = B - s_i - 1$. Since $\mu_i > 0$, $s_i + \mu_i \leq B/\gamma$ and $B \geq (a_1/\epsilon^2) \ln(2d/\epsilon)$ by assumption, we see that $\delta_i \geq 0$ as long as, e.g., $a_1 \geq 6$. Next, (11) shows that $\Pr[\mathcal{E}_i] \leq G(\mu_i, \delta_i)$; since $s_i + \mu_i \leq B/\gamma$, part (i) of Lemma 2.1 shows that $G(\mu_i, \delta_i) \leq G(u_i, v_i)$, where $u_i = B/\gamma - s_i$ and $u_i(1 + v_i) = B - s_i - 1$. So,

$$\Pr[\mathcal{E}_i] \leq G(u_i, v_i) = e^{B(\gamma-1)/\gamma} \cdot e^{-1} \cdot (1 + \frac{\gamma - B\gamma + B}{(B - s_i - 1)\gamma})^{B - s_i - 1}. \quad (39)$$

Note that $(\gamma - B\gamma + B)/\gamma < 0$, if $a_1 \geq 6$. For any fixed $\nu > 0$, the function $\lambda \mapsto (1 - \nu\lambda)^{1/\lambda}$ decreases in the range $0 < \lambda < 1/\nu$. Thus, if we view s_i as a variable constrained to lie in $[0, B/\gamma]$ and keep all other parameters fixed, then $G(u_i, v_i)$ is maximized when $s_i = 0$. So,

$$\Pr[\mathcal{E}_i] \leq \frac{e^{B(\gamma-1)/\gamma} \cdot e^{-1} \cdot (1 + 1/(B-1))^{B-1}}{\gamma^{B-1}} \leq \frac{e^{B-B/\gamma}}{\gamma^{B-1}} = (1 + \xi)e^{B(\xi/(1+\xi) - \ln(1+\xi))},$$

where $\xi = \epsilon/3$ (recall that $\gamma = 1 + \xi$). Also, by definition of d , $|S(j)| \leq d$. Substituting these ideas into (38), we get

$$\mathbf{E}[z_j \mid (\bigwedge_{i \in [m]} \overline{E_i})] \geq \Pr[z_j = 1] \cdot (1 - d \cdot (1 + \xi)e^{B(\xi/(1+\xi) - \ln(1+\xi))}). \quad (40)$$

Now, the function $t \mapsto \ln(1+t) - t/(1+t) - 9t^2/32$ vanishes at $t = 0$; its derivative is $t((1+t)^{-2} - 9/16)$, which is non-negative for $t \in [0, 1/3]$. Thus, this function is non-negative for $t \in [0, 1/3]$; since $\xi \in [0, 1/3]$, this implies that $(1 + \xi)e^{B(\xi/(1+\xi) - \ln(1+\xi))} \leq (1 + \xi)e^{-9B\xi^2/32}$, which is at most ξ/d if B is large enough as specified in the statement of the theorem.

Plugging this into (40), we get $\mathbf{E}[z_j \mid (\bigwedge_{i \in [m]} \overline{E_i})] \geq \Pr[z_j = 1](1 - \xi)$. Thus, by linearity of expectation, $\mathbf{E}[(\sum_j w_j z_j) \mid (\bigwedge_{i \in [m]} \overline{E_i})] \geq (1 - \xi) \sum_j w_j \Pr[z_j = 1]$. So there exists a rounding avoiding all the events E_i , with $\sum_j w_j z_j \geq (1 - \xi) \sum_j w_j \Pr[z_j = 1]$. This can be turned into a deterministic polynomial-time algorithm by mimicking our proofs of Theorems 3.2 and 3.3. Thus, the final solution $\sum_j w_j (\lfloor x_j^*/\gamma \rfloor + z_j)$ is at least as high as

$$\sum_j w_j (\lfloor x_j^*/\gamma \rfloor + (1 - \xi) \cdot (x_j^*/\gamma - \lfloor x_j^*/\gamma \rfloor)) \geq \sum_j w_j (1 - \xi) x_j^*/\gamma \geq \sum_j w_j x_j^*/(1 + \epsilon) = y^*/(1 + \epsilon).$$

■

6 Conclusions

We have presented improved approximation algorithms for a few classes of routing and packing problems, based on LP relaxations that basically involve multicommodity flow. We have essentially taken this approach to the limit for both our approximation algorithms for the UFP: there are families of instances of even the MDP wherein $d = \Theta(\sqrt{m})$, $\alpha(\mathcal{T}) = \Theta(1)$, and $\alpha^*(\mathcal{T}) = \Theta(\sqrt{m})$ (Kleinberg (1996)). In fact, recent work of Guruswami et al. (1999) has shown that these approximations are essentially best-possible: for directed graphs, they have shown that there cannot be an $O(m^{1/2-\epsilon})$ -approximation algorithm even for the MDP for any constant $\epsilon > 0$, unless $P = NP$.

One interesting question is to what extent correlation inequalities such as the FKG inequality and that of Lemma 3.3, are required to obtain our results. Suppose the capacities are “small”; e.g., suppose the maximum demand and the minimum capacity are the same. Then, as shown by Guruswami et al. (1999), simple greedy algorithms as well as direct randomized rounding algorithms can yield approximations similar to ours. At the other extreme, if all capacities are $\Omega(\log m)$ times larger than all demands, work of Raghavan and Thompson (1987) shows that the UFP can be approximated to within a constant factor. It is in the intermediate range where the capacities are “moderately” (say, thrice) larger than the demands, that these other approaches do not seem to provide results such as Theorem 3.3. It would be interesting if simpler methods than ours lead to the results reported in this work.

Some other algorithmic directions are to study the complexity of these problems on specific interesting classes of networks, and to develop good distributed algorithms. It is also of interest to develop robust code for many of these problems.

Acknowledgements. Jon Kleinberg’s thesis (Kleinberg (1996)) has been a major inspiration for this work. We thank Jon Kleinberg, Stavros Kolliopoulos, Tom Leighton, Satish Rao, Cliff Stein and Eli Upfal for valuable discussions, Baruch Schieber and the FOCS 1997 program committee member(s) and referee(s) for their helpful feedback, and Philip Long for his careful reading and many helpful comments. Part of the second author’s work was completed while visiting the Institute for Mathematical Sciences, Chennai-600013, India; he thanks his hosts at IMSc for their hospitality. Finally, we thank the referees and the Associate Editor for their detailed suggestions, which have greatly enhanced the quality of this work.

References

- Aggarwal A., A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, M. Sudan (1996). Efficient routing in optical networks. *Journal of the ACM* **46** 973–1001.
- Ahuja, R. K., T. L. Magnanti, J. B. Orlin (1993). *Network flows: theory, algorithms, and applications*, Prentice Hall, Englewood Cliffs, New Jersey.
- Arora, S., F. T. Leighton, B. M. Maggs (1996). On-line algorithms for path selection in a non-blocking network. *SIAM J. Comput.* **25** 600–625.
- Aumann Y., Y. Rabani (1995). Improved bounds for all-optical routing. *Proc. ACM-SIAM Symposium on Discrete Algorithms*, 567–576.
- Awerbuch, B., R. Gawlick, F. T. Leighton, Y. Rabani (1994). On-line admission control and circuit routing for high performance computing and communication. *Proc. IEEE Symposium on Foundations of Computer Science*, 412–423.
- Awerbuch, B., F. T. Leighton. Improved approximations for the multi-commodity flow problem and local competitive routing in networks. *Proc. ACM Symposium on the Theory of Computing*, 489–496.

- Boppana, R. B., J. H. Spencer (1989). A useful elementary correlation inequality. *Journal of Combinatorial Theory, Series A* **50** 305–307.
- Borkar, S., R. Cohn, G. Cox, S. Gleason, T. Gross, H. T. Kung, M. Lam, B. Moore, C. Peterson, J. Pieper, L. Rankin, P. S. Tseng, J. Sutton, J. Urbanski, J. Webb (1988). iWarp, an integrated solution to high-speed parallel computing. *Proc. International Symposium on Supercomputing*, 330–339.
- Broder, A., A. Frieze, E. Upfal (1994). Existence and construction of edge-disjoint paths on expander graphs. *SIAM J. Comput.* **23** 976–989.
- Broder, A., A. Frieze, E. Upfal (1999). Static and dynamic path selection on expander graphs: a random walk approach. *Random Structures & Algorithms* **14** 87–109.
- Cole, R. J., B. M. Maggs, R. K. Sitaraman (1995). Routing on butterfly networks with random faults. *Proc. IEEE Symposium on Foundations of Computer Science*, 558–570.
- Erdős, P., J. L. Selfridge (1973). On a combinatorial game. *Journal of Combinatorial Theory, Series A* **14** 298–301.
- Fortuin, C. M., J. Ginibre, P. N. Kasteleyn (1971). Correlational inequalities for partially ordered sets. *Communications of Mathematical Physics* **22** 89–103.
- Guruswami, V., S. Khanna, R. Rajaraman, B. Shepherd, M. Yannakakis (1999). Near-optimal hardness results and approximation algorithms for edge-disjoint paths and related problems. *Proc. ACM Symposium on Theory of Computing*, 19–28.
- Karp, R. M. (1975). On the computational complexity of combinatorial problems. *Networks* **5** 45–68.
- Kleinberg, J. (1996). Approximation algorithms for disjoint paths problems. PhD Thesis, Department of EECS, MIT, 1996.
- Kleinberg, J., R. Rubinfeld (1996). Short paths in expander graphs. *Proc. IEEE Symposium on*

Foundations of Computer Science, 86–95.

Kleinberg, J., É. Tardos (1995). Disjoint paths in densely embedded networks. Proc. IEEE Symposium on Foundations of Computer Science, 52–61.

Kleinberg, J., É. Tardos (1998). Approximations for the disjoint paths problem in high-diameter planar networks. Journal of Computer and System Sciences **57** 61–73.

Kolliopoulos, S. G., C. Stein (1997). Improved approximation algorithms for unsplittable flow problems. Proc. IEEE Symposium on Foundations of Computer Science, 426–435.

Kolliopoulos, S. G., C. Stein (1998). Approximating disjoint-path problems using greedy algorithms and packing integer programs. Proc. MPS Conference on Integer Programming and Combinatorial Optimization, 153–168, Lecture Notes in Computer Science **1412**, Springer-Verlag.

Leighton, F. T. (1992). Introduction to Parallel Algorithms and Architectures: Arrays • Trees • Hypercubes, Morgan Kaufmann, San Mateo, California.

Leighton, F. T., B. M. Maggs, S. B. Rao (1994). Packet routing and jobshop scheduling in $O(\text{congestion} + \text{dilation})$ steps. Combinatorica **14** 167–186.

Leighton, F. T., B. M. Maggs, A. Richa (1996). Fast algorithms for finding $O(\text{congestion} + \text{dilation})$ packet routing schedules. Technical Report CMU-CS-96-152, School of Computer Science, Carnegie-Mellon University.

Leighton, F. T., S. B. Rao (1988). An approximate max-flow min-cut theorem for uniform multi-commodity flow problems with applications to approximation algorithms. Proc. IEEE Symposium on Foundations of Computer Science, 422–431.

Leighton, F. T., S. B. Rao (1996). Circuit switching: a multi-commodity flow based approach. Proc. Workshop on Randomized Parallel Computing.

Lin, J. H., J. S. Vitter (1992). ϵ -approximations with minimum packing constraint violation. Proc. ACM Symposium on the Theory of Computing, 771–782.

- Motwani, R., J. Naor, P. Raghavan (1997). Randomized approximation algorithms in combinatorial optimization. D. S. Hochbaum, ed., *Approximation Algorithms for NP-Hard Problems*, PWS Press.
- Motwani, R., P. Raghavan (1995). *Randomized algorithms*. Cambridge University Press.
- Noakes, M. D., D. A. Wallach, W. J. Dally (1993). The J-Machine multicomputer: an architectural evaluation. *Proc. Annual International Symposium on Computer Architecture*, 224–235.
- Peleg, D., E. Upfal (1989). Disjoint paths on expander graphs. *Combinatorica* **9** 289–313.
- Rabani, Y. (1996). Path coloring on the mesh. *Proc. IEEE Symposium on Foundations of Computer Science*, 400–409.
- Raghavan, P., E. Upfal (1994). Efficient all-optical routing. *Proc. ACM Symposium on the Theory of Computing*, 134–143.
- Raghavan, P., C. D. Thompson (1987). Randomized rounding: a technique for provably good algorithms and algorithmic proofs. *Combinatorica* **7** 365–374.
- Raghavan, P. (1988). Probabilistic construction of deterministic algorithms: approximating packing integer programs. *Journal of Computer and System Sciences* **37** 130–143.
- Spencer, J. H. (1987). *Ten Lectures on the Probabilistic Method*. SIAM, Philadelphia.
- Srinivasan, A. (1995). Improved approximations of packing and covering problems. *Proc. ACM Symposium on the Theory of Computing*, 268–276.
- Srinivasan, A. (1996). An extension of the Lovász Local Lemma, and its applications to integer programming. *Proc. ACM-SIAM Symposium on Discrete Algorithms*, 6–15.
- Upfal, E. (1992). An $O(\log N)$ deterministic packet routing scheme. *Journal of the ACM* **39** 55–70.