Cost of Not Arbitrarily Splitting in Routing

Yingjie Bi and Ao Tang
School of Electrical and Computer Engineering, Cornell University
yb236@cornell.edu, atang@ece.cornell.edu

***Abstract*—This paper studies routing performance loss due to traffic split ratio granularity constraints. For any given fineness of traffic splitting capability, we provide upper bounds on the loss of performance. Based on tight convex relaxation, we further develop an efficient approximation algorithm to compute a good routing solution that satisfies given constraints on traffic splitting. The results can be useful for network operators to trade-off between attainable performance and implementation overhead. Some of the mathematical techniques developed here can be of independent interest for studying other similar nonconvex optimization problems.**

I. Introduction

Routing is of central importance in computer networks, with routing optimization extensively investigated in various frameworks [1], [2] such as multicommodity flow problems for traffic engineering. However, in reality, different restrictions can well prevent an optimized routing pattern to be fully realized. For example, multipath routing, in which each source-destination pair can split its traffic arbitrarily, has the potential to achieve the best performance. But it is not commonly used today due to various difficulties of splitting traffic arbitrarily. Instead, the dominating routing protocols in Internet such as Open Shortest Path First (OSPF) only allow a user to choose a single path from the source to the destination, with the exception that traffic may split evenly among equal-cost paths. This paper aims at providing a sound estimation of performance loss due to such inflexibility in splitting traffic, hence provides a solid basis for trade-off between routing performance and implementation overhead.

The single-path routing and multipath routing can be regarded as two extreme cases of a spectrum of routing schemes of different flexibility, and there are different ways to interpret this flexibility. One obvious way is based on the number of paths (W) that are allowed for each source-destination pair. Single-path routing corresponds to the case of $W = 1$, and multipath routing can be understood as the limit case of $W \to \infty$. For any integer $W > 1$, it is expected that the routing solution would perform somewhere between single-path routing and general multipath routing. Routing optimization with these path cardinality constraints is examined in [3].

Here we introduce a different angle to generalize single-path and multipath routing by considering routing problems with certain traffic splitting restrictions. In particular, each

This work was supported by the Office of Naval Research under N00014-17-1-2419, and by the Huawei Innovation Research Program.

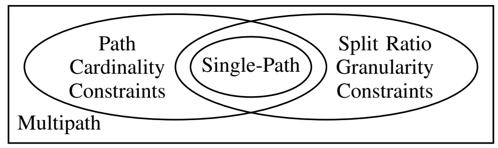


Fig. 1. The relationship among the different types of constraints that can be imposed in the routing optimization problem. Note that the single-path constraint is a special case of both path cardinality constraints and split ratio granularity constraints.

user is only allowed to choose its traffic split ratio as a multiple of $1/p$, where the integer p is the split granularity parameter.[1] Clearly, the case of $p = 1$ is equivalent to single-path routing, while $p \to \infty$ corresponds to multipath routing. Routing optimization with split ratio granularity constraints was studied by heuristic techniques in previous works such as [4], [5]. However, to the best of our knowledge, this is the first time the impact of such granularity constraints is systematically analyzed. The relationship among the different types of constraints mentioned here is summarized in Fig. 1.

More specifically, this paper studies two questions: First, we want to estimate the performance loss due to split ratio granularity constraints. A special case of this question was examined in [6], which estimates the so called *cost of not splitting* in routing, i.e., performance gap between multipath and single-path routing. There the routing optimization is formulated in the network utility maximization framework, and the performance of a routing solution is measured in terms of the aggregate utility of all users it attains. In Section III, we will study the exact same question for the split ratio granularity constraints, which can be viewed as the *cost of not arbitrarily splitting* in routing. The other question is about actually optimizing routing with split ratio granularity constraints (Section IV). Like the problem of finding optimal OSPF link weights [7], [8], we have to resort to approximation solutions. The performance loss between the solution we found and the optimal one can be regarded as another layer of the cost of not arbitrarily splitting in routing.

II. Model and Notation

Consider a network with L uni-directional links and N users (source-destination pairs). The capacity of link l is c_l, for $l =$

[1] There are other possible split ratio granularity constraints. For example, in [4] the constraint is that the split ratio is a multiple of $1/q$, where q is an arbitrary integer bounded by some fixed number Q.

978-1-5090-6501-1/17/$31.00 ©2017 IEEE

$1, \ldots, L$. For each user i, there are K^i predetermined paths from its source to its destination which are represented by an $L \times K^i$ matrix R^i, where $R^i_{lk} = 1$ if path k of user i passes through link l and $R^i_{lk} = 0$ otherwise. The rate vector x^i of user i is a $K^i \times 1$ vector whose kth entry x^i_k denotes the sending rate of user i through its path k. Each user i also has a utility $U^i(\cdot)$ as a function of the total sending rate $\|x^i\|_1$ on its all available paths. We assume that the utility function $U^i(\cdot)$ is continuous, concave and increasing [9].

The network utility maximization (NUM) problem is to maximize the aggregate utility of all users. In multipath routing, the only constraints imposed on the rate vectors x^i are nonnegativity and link capacity constraints, and the NUM problem can be formulated as the following convex optimization problem:

$$\begin{aligned} \max \quad & \sum_{i=1}^{N} U^i\left(\|x^i\|_1\right) \\ \text{s. t.} \quad & \sum_{i=1}^{N} R^i x^i \le c, \\ & x^i \in \mathcal{I}_{K^i}, \quad \forall i = 1, \ldots, N. \end{aligned} \tag{1}$$

Here the set $\mathcal{I}_K$ is given by

$$\mathcal{I}_K = \{x \in \mathbb{R}^K | 0 \le x_k \le \|c\|_\infty, k = 1, \ldots, K\},$$

where $\|c\|_\infty$ is the maximum capacity among all links in the network. The constraints $x_k \le \|c\|_\infty$ above are redundant since they are implied by the link capacity constraints. The reason for including these constraints in problem (1) is to make $\mathcal{I}_K$ a bounded set, simplifying our later discussion. In addition, in the remaining of this paper, it is assumed without loss of generality that $\|c\|_\infty = 1$.

The multipath problem (1) does not model other practical restrictions on the sending rates. As a general way to deal with these additional restrictions, we introduce a set $\mathcal{S}_K \subseteq \mathcal{I}_K$ for the following variant of the NUM problem:

$$\begin{aligned} \max \quad & \sum_{i=1}^{N} U^i\left(\|x^i\|_1\right) \\ \text{s. t.} \quad & \sum_{i=1}^{N} R^i x^i \le c, \\ & x^i \in \mathcal{S}_{K^i}, \quad \forall i = 1, \ldots, N. \end{aligned} \tag{2}$$

For instance, in the NUM problem with single-path routing, $\mathcal{S}_K$ will contain all the vectors in $\mathcal{I}_K$ with at most one positive component.

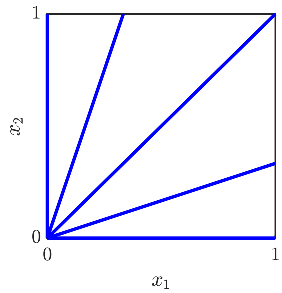
In the case of split ratio granularity constraints studied in this paper, the split ratio needs to be a multiple of $1/p$, where p is a given positive integer. Hence the following definition for the set $\mathcal{S}_K$ will be adopted in problem (2):

$$\mathcal{S}_K = \{0\} \cup \left\{ x \in \mathcal{I}_K \middle| x \ne 0, \frac{p x_k}{\|x\|_1} \in \mathbb{Z}, k = 1, \ldots, K \right\}.$$

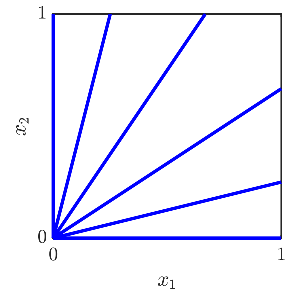
For convenience, usually we will work with an equivalent definition for the set $\mathcal{S}_K$: A rate vector $x \in \mathcal{I}_K$ satisfies the split ratio granularity constraints, i.e., $x \in \mathcal{S}_K$, if and only if there exists a vector $\alpha \in \mathbb{R}^K$ satisfying:

1) $\alpha_k \ge 0$, $\alpha_k \in \mathbb{Z}$, for $k = 1, \ldots, K$.
2) $\sum_{k=1}^{K} \alpha_k = p$.
3) There exists some real number $\lambda \ge 0$ such that $x = \lambda \alpha$.

The vector α will be known as the (scaled) split ratio of x.



(a) $p = 4$



(b) $p = 5$

Fig. 2. The set $\mathcal{S}_2$ (shown as the bold lines above) under the cases $p = 4$ and $p = 5$.

Fig. 2 illustrates the set $\mathcal{S}_K$ for $K = 2$ and different choices of p. In general, the set $\mathcal{S}_K$ is the union of rays each of which passes through a point α satisfying the first and second conditions above, and $\mathcal{S}_K$ contains the corner $(1, 1, \ldots, 1)$ of $\mathcal{I}_K$ if and only if p is a multiple of K.

Denote the optimal value of the multipath problem (1) and the split ratio granularity problem (2) as opt_M and opt_G, respectively. Then the two questions proposed at the end of Section I can be parsed mathematically as following:

The first question asks how much the performance loss $\text{opt}_M - \text{opt}_G$ is after the split ratio granularity constraints are introduced. Because the objective function in NUM is separable for each user, the general idea of estimating the duality gap for nonconvex problems with separable objectives [10] can be applied here. First, observe that if the sending rate on some path is zero, then the split ratio of this path is also zero and thus automatically satisfies the granularity constraints. Therefore, given an optimal routing solution $(x^1, x^2, \ldots, x^N)$ to the multipath problem (1), it is natural to search for another solution $(\bar{x}^1, \bar{x}^2, \ldots, \bar{x}^N)$ with the same aggregate utility, while sending positive rates on paths as few as possible. Next, for each user i, find the rate vector $y^i \le \bar{x}^i$ maximizing $\|y^i\|_1$ among all rate vectors in $\mathcal{S}_{K^i}$ dominated by $\bar{x}^i$. The rounded solution $(y^1, y^2, \ldots, y^N)$ is feasible to the split ratio granularity problem (2). During the rounding step, if the total utility loss of all users

$$\sum_{i=1}^{N} U^i\left(\|\bar{x}^i\|_1\right) - \sum_{i=1}^{N} U^i\left(\|y^i\|_1\right)$$

is upper bounded by Δ_M, then

$$\begin{aligned} \text{opt}_G &\ge \sum_{i=1}^{N} U^i\left(\|y^i\|_1\right) \ge \sum_{i=1}^{N} U^i\left(\|\bar{x}^i\|_1\right) - \Delta_M \\ &= \sum_{i=1}^{N} U^i\left(\|x^i\|_1\right) - \Delta_M = \text{opt}_M - \Delta_M. \end{aligned} \tag{3}$$

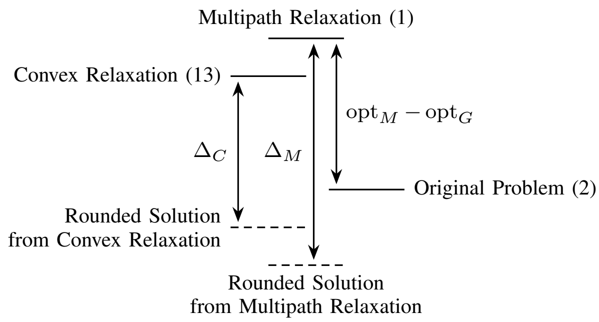


Fig. 3. The optimal values for the original split ratio granularity problem (2) and two relaxations (1) and (13) (represented by the solid lines). Δ_M or Δ_C is the upper bound of performance loss for the rounded solution (represented by the dash lines) obtained from the optimal solution to the corresponding relaxed problem.

We will devote Section III to calculating Δ_M.

The second question mentioned in Section I concerns finding an approximate solution to the split ratio granularity problem (2). By the inequality (3),

$$\mathrm{opt}_G \geq \sum_{i=1}^{N} U^i \left(\|y^i\|_1\right) \geq \mathrm{opt}_M - \Delta_M \geq \mathrm{opt}_G - \Delta_M,$$

so the rounded solution $(y^1, y^2, \ldots, y^N)$ has already been a suboptimal solution to (2) at most Δ_M away from the optimality. In Section IV, to obtain an approximate solution with better performance guarantee, we will tighten the multipath relaxation (1) by the convex relaxation (13), because in convex relaxation (13) the performance loss Δ_C during rounding can be smaller than Δ_M. The relationship among the optimal values of three optimization problems and the objective values of the rounded solutions from two relaxations is summarized in Fig. 3, which shows that Δ_M or Δ_C is an upper bound for not only the performance gap between the original problem and the relaxed problem but also the performance guarantee of how far the rounded solution can be away from the optimal value for the original problem.

III. Bounding the Performance Gap

Now we turn to the problem of establishing an upper bound for the performance gap between the multipath problem (1) and the split ratio granularity problem (2). Generally, the optimal solution to the multipath problem is not unique. As mentioned in Section II, we want to start with an optimal solution using very few number of paths such that it satisfies almost all the granularity constraints, then round it into a feasible solution to the split ratio granularity problem (2). In this section, we will present the upper bound Δ_M for the performance loss during rounding, which has already shown to be the performance gap by inequality (3).

The performance gap depends on the specific form of the utility function of each user. In this paper, we focus on two utility functions commonly encountered in practice, the linear utility $U^i(s) = s$ and the logarithmic utility $U^i(s) = \log s$, and in the following we will deal with them in parallel.

A. *Optimal Vertex Solution to Multipath Routing*

To find an optimal multipath solution as a good start for the rounding step, we invoke the following lemma, which is essentially the $N + L$ *property* that appears in the study of many other traffic engineering problems such as [6], [11].

Lemma 1: There exists an optimal solution to the multipath problem (1) sending positive rates on at most $N + L$ paths.

The optimal multipath solution satisfying the additional property in Lemma 1 will be called an *optimal vertex solution* later. For proof, see [6, Theorem 4]. In fact, the proof there actually demonstrates an explicit procedure to construct an optimal vertex solution from an arbitrary optimal multipath solution.

B. *The Throughput Loss Function*

The second step is to round the optimal vertex solution $(\bar{x}^1, \bar{x}^2, \ldots, \bar{x}^N)$ given by Lemma 1 into a feasible solution $(y^1, y^2, \ldots, y^N)$ to the split ratio granularity problem (2). If the rounded rate vector $y^i \in \mathcal{S}_{K^i}$ satisfies $y^i \leq \bar{x}^i$ for all $i = 1, \ldots, N$, the new solution is guaranteed not to violate the link capacity constraints. This illuminates us to consider the following optimization problem that finds the optimal rounding $y \in \mathcal{S}_K$ for a rate vector $x \in \mathcal{I}_K$ such that the throughput loss is minimized:

$$\begin{aligned} \min \quad & \|x\|_1 - \|y\|_1 \\ \mathrm{s.t.} \quad & y \leq x, \\ & y \in \mathcal{S}_K. \end{aligned} \tag{4}$$

For each rate vector $x \in \mathcal{I}_K$, define the *throughput loss function* $\phi_K(x)$ to be the optimal value of problem (4). In the next, we will give an explicit formula for the function $\phi_K(x)$. Using the equivalent definition for the set $\mathcal{S}_K$ given in Section II, i.e., write $y = \lambda\alpha$ where α is the split ratio of y, (4) can be rewritten as a problem over variables λ and α:

$$\begin{aligned} \min \quad & \|x\|_1 - p\lambda \\ \mathrm{s.t.} \quad & \lambda\alpha_k \leq x_k, \quad \forall k = 1, \ldots, K, \\ & \lambda \geq 0, \; \sum_{k=1}^{K} \alpha_k = p, \\ & \alpha_k \geq 0, \; \alpha_k \in \mathbb{Z}, \quad \forall k = 1, \ldots, K. \end{aligned} \tag{5}$$

Fix the split ratio α. To minimize the throughput loss, the largest possible λ should be to chosen, which is given by

$$\lambda^* = \min_{k=1,\ldots,K} \frac{x_k}{\alpha_k}.$$

Here we make the convention that $x_k/\alpha_k = +\infty$ when $\alpha_k = 0$. Plugin the formula for λ^* into (5), the optimal value

$$\phi_K(x) = \|x\|_1 - p \max_{\alpha} \min_{k=1,\ldots,K} \frac{x_k}{\alpha_k}, \tag{6}$$

where the maximization is taken over all the vectors $\alpha \geq 0$ with $\sum_{k=1}^{K} \alpha_k = p$ and $\alpha_k \in \mathbb{Z}$ for $k = 1, \ldots, K$.

In Section IV, when we state the approximation algorithm for solving the split ratio granularity problem (2), the optimal rounding for the optimal vertex solution needs to be actually calculated out, and we will discuss how to efficiently find the optimal rounding for any rate vector at that time.

The throughput loss function $\phi_K(x)$ will play an important role in our analysis of the performance gap. Here we list some simple properties of $\phi_K(x)$:

1) $\phi_K(x)$ is continuous on the domain $\mathcal{I}_K$.
2) If $x \in \mathcal{I}_K$ and $\lambda x \in \mathcal{I}_K$, $\phi_K(\lambda x) = \lambda\phi_K(x)$.
3) For any $x \in \mathcal{I}_K$, let $\tilde{x} \in \mathcal{I}_{\tilde{K}}$ be the vector containing the positive components in x, then $\phi_{\tilde{K}}(\tilde{x}) = \phi_K(x)$.

Define

$$\rho_K = \max_{x \in \mathcal{I}_K} \phi_K(x)$$

to be the maximum throughput loss. As a convention, $\rho_0 = 0$.

The throughput loss function $\phi_2(x)$ for the case $p = 4$ is illustrated in Fig. 4(a), and the actual optimal rounding for a rate vector in this case is described in Fig. 4(b). The local maximizers for $\phi_2(x)$ are

$$(1/4, 1), (1, 1/4), (2/3, 1), (1, 2/3).$$

The later two are global maximizers and $\rho_2 = 1/3$. The situations in higher dimensions are similar but the number of local maximizers is much larger (see Fig. 4(c)).

Assume the utility function is $U^i(s) = s$. By Lemma 1, there exists an optimal vertex solution $(\bar{x}^1, \bar{x}^2, \ldots, \bar{x}^N)$ to the multipath problem (1) using at most $N + L$ paths in total. Let $\tilde{K}^i$ be the number of nonzero components in the vector $\bar{x}^i$. Then $0 \le \tilde{K}^i \le K^i$, $\sum_{i=1}^N \tilde{K}^i \le N + L$, and the throughput loss for user i during the rounding step is bounded by $\rho_{\tilde{K}^i}$. Since $\tilde{K}^i$ are unknown, we have to enumerate all the possible values of $\tilde{K}^i$ and find the worst case in order to bound the total throughput loss, which leads to the following optimization problem:

$$\begin{aligned} \max \quad & \sum_{i=1}^N \rho_{\tilde{K}^i} \\ \text{s.t.} \quad & \sum_{i=1}^N \tilde{K}^i \le N + L, \\ & 0 \le \tilde{K}^i \le K^i,\ \tilde{K}^i \in \mathbb{Z}, \quad \forall i = 1, \ldots, N. \end{aligned} \tag{7}$$

Let Δ_M denote its optimal value. By the previous analysis, Δ_M will be an upper bound for the performance loss during rounding and thus also be an upper bound for the performance gap $\mathrm{opt}_M - \mathrm{opt}_G$.

The case for the logarithmic utility function $U^i(s) = \log s$ is similar. Here if the original rate vector of user i is x^i and the rounded rate vector is y^i, then

$$U^i(\|x^i\|_1) - U^i(\|y^i\|_1) = \log \frac{\|x^i\|_1}{\|y^i\|_1}.$$

Accordingly, consider the maximum relative throughput loss

$$\rho_K^R = \max_{x \in \mathcal{I}_K, x \ne 0} \frac{\phi_K(x)}{\|x\|_1}.$$

By the same analysis, we can obtain the performance gap for the logarithmic utility case by replacing $\rho_{\tilde{K}^i}$ with $-\log(1 - \rho_{\tilde{K}^i}^R)$ in (7) (of course, in this case $\tilde{K}^i$ can never be zero).

Note that when the maximum capacity of links $\|c\|_\infty \ne 1$, in the linear utility case the upper bound for the performance gap will become $\Delta_M \|c\|_\infty$. However, in the logarithmic utility case, the upper bound will not be affected.

To compute the upper bound Δ_M for the performance gap, we need to solve the optimization problem (7), in which the numbers ρ_K for the linear utility case or ρ_K^R for the logarithmic utility case are still to be determined. Although the throughput loss function $\phi_K(x)$ and the relative throughput loss function $\phi_K(x)/\|x\|_1$ are not concave, it turns out that we can still write down exact formulas for the maximum values ρ_K and ρ_K^R of these functions. Because the maximizers for the throughput loss function can be different from those for the relative one, in the subsequent two subsections we will treat the two cases separately.

C. Calculating ρ_K^R for Logarithmic Utility

We start with the easier case of logarithmic utility, which requires us to calculate the maximum relative throughput loss ρ_K^R.

Theorem 2: For every $x \in \mathcal{I}_K$ with $x \ne 0$, the relative throughput loss

$$\frac{\phi_K(x)}{\|x\|_1} \le \frac{K-1}{p+K-1}.$$

Proof: Since the function $\phi_K(x)/\|x\|_1$ is continuous, we only need to prove this inequality for every x in a dense subset of $\mathcal{I}_K$. Let $\mathcal{I}'_K$ be the set containing all nonzero $x \in \mathcal{I}_K$ with the following property: For any pair of paths $1 \le k, l \le K$, $k \ne l$ and any pair of integers $1 \le a, b \le p$, $x_k/a \ne x_l/b$. $\mathcal{I}'_K$ is a dense subset of $\mathcal{I}_K$ since it is a finite intersection of open dense subsets, so it is sufficient to prove the inequality on $\mathcal{I}'_K$.

For any $x \in \mathcal{I}'_K$, let α be the split ratio of the optimal rounding for x, i.e., α is the one attaining the maximum in (6). Without loss of generality, we also assume that

$$\min_{k=1,\ldots,K} \frac{x_k}{\alpha_k} = \frac{x_1}{\alpha_1}.$$

In this case, $\alpha_1 > 0$, which implies that $x_1 > 0$ and $\alpha_l < p$ for all other $l = 2, \ldots, K$. For these l, we are going to prove by contradiction that

$$\frac{x_1}{\alpha_1} \ge \frac{x_l}{\alpha_l + 1}. \tag{8}$$

If the above inequality does not hold for a particular path l, define a new split ratio β by setting

$$\beta_k = \begin{cases} \alpha_1 - 1, & \text{if } k = 1, \\ \alpha_l + 1, & \text{if } k = l, \\ \alpha_k, & \text{otherwise.} \end{cases}$$

By our assumption that x_1/α_1 is the minimum,

$$\frac{x_k}{\beta_k} = \frac{x_k}{\alpha_k} \ge \frac{x_1}{\alpha_1},\ k \ne 1, l.$$

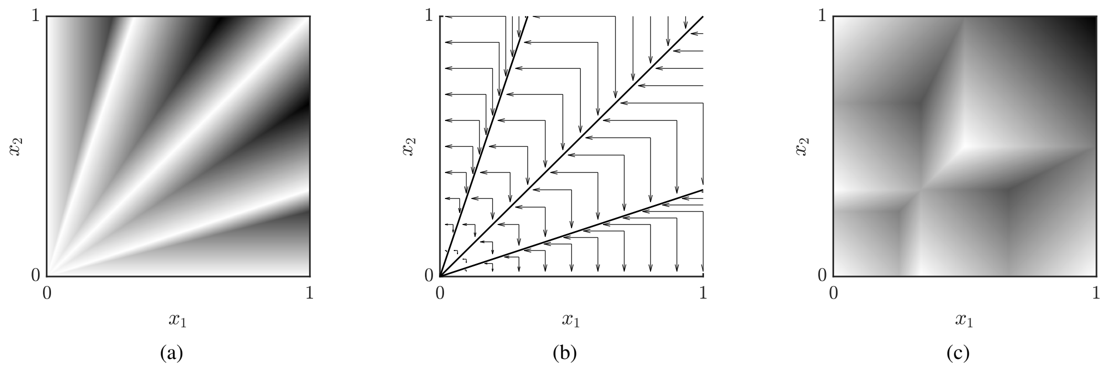


Fig. 4. (a) The throughput loss function $\phi_2(x)$ for the case $p = 4$. The function value is close to zero (white area) on points near the set $\mathcal{S}_2$ and becomes large (dark area) on points far away from the set $\mathcal{S}_2$. (b) The optimal rounding for a rate vector in $\mathcal{I}_2$. To find the optimal rounding, start with that vector and move according to the arrow until a point in $\mathcal{S}_2$ (represented by the solid lines) is reached. (c) The value of function $\phi_3(x)$ on the plane $x_3 = 1$. Darker area indicates larger function value.

Furthermore, the above inequality is strict because $x \in \mathcal{I}'_K$. In addition,

$$\frac{x_1}{\beta_1} = \frac{x_1}{\alpha_1 - 1} > \frac{x_1}{\alpha_1}, \quad \frac{x_l}{\beta_l} = \frac{x_l}{\alpha_l + 1} > \frac{x_1}{\alpha_1}.$$

Therefore,

$$\min_{k=1,\ldots,K} \frac{x_k}{\beta_k} > \frac{x_1}{\alpha_1} = \min_{k=1,\ldots,K} \frac{x_k}{\alpha_k},$$

which contradicts with the fact that α is the optimal split ratio for the rounding of x.

Combining all the inequalities (8) for $l = 2, \ldots, K$, we have

$$\frac{x_1}{\alpha_1} \geq \frac{x_1 + \sum_{l=2}^{K} x_l}{\alpha_1 + \sum_{l=2}^{K}(\alpha_l + 1)} = \frac{\|x\|_1}{p + K - 1}$$

and

$$\frac{\phi_K(x)}{\|x\|_1} = \frac{1}{\|x\|_1}\left(\|x\|_1 - p \min_{k=1,\ldots,K} \frac{x_k}{\alpha_k}\right) = 1 - \frac{p x_1}{\|x\|_1 \alpha_1}$$
$$\leq 1 - \frac{p}{p + K - 1} = \frac{K - 1}{p + K - 1},$$

which is the desired inequality. ■

Theorem 2 shows that

$$\rho_K^R \leq \frac{K-1}{p+K-1}.$$

In the next subsection, we will see that this bound is in fact tight as an easy corollary of Lemma 3 and Lemma 4. In other words,

$$\rho_K^R = \frac{K-1}{p+K-1}.$$

D. Calculating ρ_K for Linear Utility

Now we calculate the maximum throughput loss ρ_K for the linear utility. To find the maximum value ρ_K for the nonconcave function $\phi_K(x)$, we first establish a type of first-order condition that all local maximizers of $\phi_K(x)$ will satisfy. Despite the fact that the number of local maximizers can be exponentially large, we are able to classify the local maximizers into a few categories, and in each category we can find the one with the maximum throughput loss. Roughly speaking, the following definition provides the necessary condition for a rate vector to be a local maximizer:

Definition 1: A rate vector $x \in \mathcal{I}_K$ is said to be *integral* if there exists a vector $\beta \in \mathbb{R}^K$ satisfying:

1) $\beta_k \geq 0$, $\beta_k \in \mathbb{Z}$, for $k = 1, \ldots, K$.
2) $p \leq \sum_{k=1}^{K} \beta_k \leq p + \|x\|_0 - 1$. Here $\|x\|_0$ is the number of nonzero components in x.
3) There exists some real number $\lambda \geq 0$ such that $x = \lambda\beta$.

To be specific, we will also use the terminology Γ-*integral* for integral rate vectors with $\sum_{k=1}^{K} \beta_k = \Gamma$.

Clearly, a rate vector $x \in \mathcal{I}_K$ is p-integral if and only if $x \in \mathcal{S}_K$ and $x \neq 0$. In Fig. 4(a), the local maximizer $(1/4, 1)$ of $\phi_2(x)$ is integral and the corresponding vector $\beta = (1, 4)$. The global maximizer $(2/3, 1)$ is also integral for $\beta = (2, 3)$.

Here are the major steps to compute ρ_K:

1) Prove that any maximizer of the function $\phi_K(x)$ must be integral. Hence we only need to figure out the maximum throughput loss over all integral rate vectors.
2) Find the relative throughput loss for integral rate vectors. We will see that all rate vectors that are Γ-integral have the same relative throughput loss, which is $1 - p/\Gamma$.
3) Show that the maximum throughput among all Γ-integral rate vectors is given by

$$\frac{\Gamma}{\lceil \Gamma/K \rceil}.$$

Finally, we can conclude that

$$\rho_K = \max_{\Gamma=p,\ldots,p+K-1} \left(1 - \frac{p}{\Gamma}\right) \frac{\Gamma}{\lceil \Gamma/K \rceil}$$
$$= \max_{\Gamma=p,\ldots,p+K-1} \frac{\Gamma - p}{\lceil \Gamma/K \rceil}.$$

In the following, we will first address the last two steps in Lemma 3 and Lemma 4. The first step needs more elaboration and will be dealt with thereafter.

Lemma 3: If a rate vector $x \in \mathcal{I}_K$ is Γ-integral, then its relative throughput loss

$$\frac{\phi_K(x)}{\|x\|_1} = 1 - \frac{p}{\Gamma}.$$

Proof: Let $x = \lambda\beta$ as required by Definition 1. Since $\Gamma = \sum_{k=1}^{K} \beta_k \geq p$, we can find a split ratio vector α such that

$$0 \leq \alpha_k \leq \beta_k,\ \alpha_k \in \mathbb{Z},\ k = 1, \dots, K,$$

and $\sum_{k=1}^{K} \alpha_k = p$. Obviously, $\lambda\alpha$ is a feasible solution to the optimal rounding problem (4) for x, so

$$\frac{\phi_K(x)}{\|x\|_1} \leq 1 - \frac{\sum_{k=1}^{K} \lambda\alpha_k}{\sum_{k=1}^{K} \lambda\beta_k} = 1 - \frac{p}{\Gamma}.$$

On the other hand, now assume that another split ratio α gives the optimal rounding for x. Suppose $\alpha_k < \beta_k$ for all $k = 1, \dots, K$ with $x_k > 0$. Since α_k and β_k are integers, $\alpha_k + 1 \leq \beta_k$ for these k. Noticing that $x_k > 0$ if and only if $\beta_k > 0$, and $\alpha_k > 0$ implies $x_k > 0$, we have

$$p + \|x\|_0 = \sum_{k:x_k>0} (\alpha_k + 1) \leq \sum_{k:x_k>0} \beta_k,$$

which contradicts with the requirement on β. Thus there must exist some path l such that $\alpha_l \geq \beta_l > 0$, which implies

$$\min_{k=1,\dots,K} \frac{x_k}{\alpha_k} \leq \frac{x_l}{\alpha_l} \leq \frac{x_l}{\beta_l} = \lambda$$

and

$$\begin{aligned}\frac{\phi_K(x)}{\|x\|_1} &= \frac{1}{\|x\|_1}\left(\|x\|_1 - p \min_{k=1,\dots,K} \frac{x_k}{\alpha_k}\right) \\ &\geq 1 - \frac{p\lambda}{\sum_{k=1}^{K} \lambda\beta_k} = 1 - \frac{p}{\Gamma}.\end{aligned}$$

The desired result is followed from combining the two directions we have already shown. ■

Lemma 4: For integer Γ with $p \leq \Gamma \leq p + K - 1$, there exists a rate vector in $\mathcal{I}_K$ that is Γ-integral, and the maximum throughput among all Γ-integral rate vectors is

$$\frac{\Gamma}{\lceil \Gamma/K \rceil}.$$

Proof: First, we construct a Γ-integral rate vector with the given throughput. Note that we can choose integers $\beta_k \in \{\lfloor \Gamma/K \rfloor, \lceil \Gamma/K \rceil\}$ for $k = 1, \dots, K$ such that $\sum_{k=1}^{K} \beta_k = \Gamma$. Let

$$\lambda = \frac{1}{\lceil \Gamma/K \rceil}$$

and $x = \lambda\beta$. Then $x_k \leq 1$ for all k, so $x \in \mathcal{I}_K$. If $\Gamma \geq K$, all $x_k > 0$ and thus

$$\Gamma \leq p + K - 1 = p + \|x\|_0 - 1.$$

If $\Gamma < K$, then $\|x\|_0 = \Gamma$ and we also have $\Gamma \leq p + \|x\|_0 - 1$. Therefore, x is Γ-integral and its throughput

$$\|x\|_1 = \lambda\Gamma = \frac{\Gamma}{\lceil \Gamma/K \rceil}.$$

Consider an arbitrary rate vector $x \in \mathcal{I}_K$ that is Γ-integral with corresponding vector β. Because $\sum_{k=1}^{K} \beta_k = \Gamma$, there must exist some path l such that $\beta_l \geq \Gamma/K$. Since β_l is an integer, $\beta_l \geq \lceil \Gamma/K \rceil > 0$, which implies

$$\|x\|_1 = \lambda\Gamma = \frac{x_l}{\beta_l}\Gamma \leq \frac{x_l}{\lceil \Gamma/K \rceil}\Gamma \leq \frac{\Gamma}{\lceil \Gamma/K \rceil},$$

showing that the given throughput is indeed maximal. ■

The remaining task is to prove that any maximizer of the function $\phi_K(x)$ is integral. However, there is a technical difficulty which can be demonstrated by the following counterexample: Assume $x \in \mathcal{I}_3$ is a rate vector and $x_1 \geq x_2 \geq x_3$. In the case of $p = 2$, since the vector $(x_2, x_2, 0)$ satisfies the split ratio granularity constraints,

$$\phi_3(x) \leq x_1 - x_2 + x_3 \leq x_1 \leq 1$$

and thus $\rho_3 \leq 1$. On the other hand, we can directly check that for all rate vectors y of the form $(1, t, t)$ where $1/2 \leq t \leq 1$, $\phi_3(y) = 1$. Thus the function $\phi_3(x)$ has infinite maximizers and almost all of them are not integral. Fortunately, our goal is not finding all the maximizers of $\phi_K(x)$ but computing the maximum of $\phi_K(x)$, and the plan presented at the beginning of this section still works as long as one of the maximizers is integral. The integrality of the maximizer $(1, 1, 1)$ dominating all other maximizers enlightens the additional requirement in Lemma 5.

Lemma 5: If a rate vector $x \in \mathcal{I}_K$ satisfies:

1) x maximizes the function $\phi_K(x)$;
2) For any $z \in \mathcal{I}_K$, if $\phi_K(z) = \phi_K(x)$ and $z \geq x$, then $z = x$,

then x is integral.

Proof: Repeating what we have done in the proof of Theorem 2, assume $x \neq 0$, α is the split ratio of the optimal rounding for x, and

$$\min_{k=1,\dots,K} \frac{x_k}{\alpha_k} = \frac{x_1}{\alpha_1}.$$

Define W to be the set of indices k such that there exists some integer $\beta_k > 0$ with $x_k/\beta_k = x_1/\alpha_1$. Then for any $l \notin W$ and any split ratio α' that also gives the optimal rounding for x, i.e.,

$$\min_{k=1,\dots,K} \frac{x_k}{\alpha'_k} = \min_{k=1,\dots,K} \frac{x_k}{\alpha_k} = \frac{x_1}{\alpha_1},$$

it must be true that

$$\frac{x_1}{\alpha_1} = \min_{k=1,\dots,K} \frac{x_k}{\alpha'_k} < \frac{x_l}{\alpha'_l}. \tag{9}$$

We want to prove that W contains all the paths. Assume the existence of a path $l \notin W$. If $x_l < 1$, consider another rate vector $z = (x_1, \dots, x_l + \delta, \dots, x_K)$ with sufficiently small $\delta > 0$. The optimal split ratio for the rounding of z must be one of the above α', and by (9) $\min x_k/\alpha'_k$ is not attained by path l, so

$$\min_{k=1,\dots,K} \frac{z_k}{\alpha'_k} = \min_{k=1,\dots,K} \frac{x_k}{\alpha'_k} = \frac{x_1}{\alpha_1}.$$

The throughput loss for the rate vector z is

$$\phi_K(z) = \|x\|_1 + \delta - \frac{px_1}{\alpha_1} > \phi_K(x),$$

which contradicts with the assumption that x is the maximizer.

In the case of $x_l = 1$, we have $x_k < 1$ for $k \in W$, otherwise l would belong to W. Construct a rate vector z defined by

$$z_k = \begin{cases} \delta x_k, & \text{if } k \in W, \\ x_k, & \text{if } k \notin W, \end{cases}$$

for some δ that is sufficiently close to 1. Similar to the previous case, the optimal split ratio for the rounding of z must be one of the α' giving the optimal rounding for x. By (9), the path minimizing x_k/α'_k is in W, which also minimizes z_k/α'_k because the difference between z and x is sufficiently small. Thus

$$\min_{k=1,\dots,K} \frac{z_k}{\alpha'_k} = \delta \min_{k=1,\dots,K} \frac{x_k}{\alpha'_k} = \frac{\delta x_1}{\alpha_1}.$$

The throughput loss for the rate vector z is

$$\phi_K(z) = \|x\|_1 + (\delta - 1) \sum_{k \in W} x_k - \frac{\delta x_1}{\alpha_1},$$

and

$$\phi_K(z) - \phi_K(x) = (\delta - 1) \left(\sum_{k \in W} x_k - \frac{x_1}{\alpha_1} \right).$$

By choosing appropriate δ, we can let either $\phi_K(z) > \phi_K(x)$ or $\phi_K(z) = \phi_K(x)$ but $z_k \geq x_k$ for all $k = 1, \dots, K$ with the inequality being strict for $k = 1 \in W$, which contradicts with either of the two conditions on x.

In the above, we have proved that the set W contains all the paths. For the integers β_k we have found, $x_k/\beta_k = x_1/\alpha_1$ so $x = \lambda\beta$ for $\lambda = x_1/\alpha_1$ and $\|x\|_0 = K$. It remains to prove that $p \leq \sum_{k=1}^K \beta_k \leq p + K - 1$. First, since

$$\frac{x_k}{\beta_k} = \frac{x_1}{\alpha_1} \leq \frac{x_k}{\alpha_k},\ k = 1, \dots, K,$$

we have $\beta_k \geq \alpha_k$ and $\sum_{k=1}^K \beta_k \geq \sum_{k=1}^K \alpha_k = p$. Moreover, the relative throughput loss of x is

$$\frac{\phi_K(x)}{\|x\|_1} = 1 - \frac{p x_1/\alpha_1}{\sum_{k=1}^K \lambda\beta_k} = 1 - \frac{p}{\sum_{k=1}^K \beta_k}.$$

Theorem 2 tells us

$$\frac{\phi_K(x)}{\|x\|_1} \leq \frac{K-1}{p+K-1},$$

which implies that $\sum_{k=1}^K \beta_k \leq p + K - 1$. ∎

Now we can prove the main result in this section:

Theorem 6: The maximum throughput loss for a rate vector in $\mathcal{I}_K$ is

$$\rho_K = \max_{\Gamma = p,\dots,p+K-1} \frac{\Gamma - p}{\lceil \Gamma/K \rceil}. \tag{10}$$

Proof: Since $\phi_K(x)$ is continuous, the set

$$\{z \in \mathcal{I}_K | \phi_K(z) = \rho_K\}$$

is nonempty, closed and bounded. We can find a rate vector x maximizing the l_1 norm on the above set, which satisfies the two conditions in Lemma 5 and thus being integral. Now we can obtain the desired result by following the plan given at the beginning of this subsection. ∎

The formula (10) for ρ_K can be further simplified. Actually, only two possibilities of Γ have to be checked. In the range $\Gamma = p, \dots, p+K-1$, $\lceil \Gamma/K \rceil$ can only increase by at most 1. Let $0 \leq \omega \leq K-1$ be the number that $(p+\omega)/K = \lceil p/K \rceil$, then the maximum in (10) is attained by either $\Gamma = p + \omega$ or $\Gamma = p + K - 1$, so

$$\rho_K = \max \left\{ \frac{\omega}{\lceil p/K \rceil}, \frac{K-1}{\lceil (p+K-1)/K \rceil} \right\}. \tag{11}$$

IV. Routing Optimization with Split Ratio Granularity Constraints

Section III discussed how to estimate the performance gap between the NUM problem with split ratio granularity constraints (2) and its multipath relaxation (1). As we mentioned in Section II, the method proving the upper bound for the performance gap can also be used to find a near-optimal routing solution under split ratio granularity constraints. The main steps can be summarized as following:

1) Solve the multipath problem (1) to find the optimal multipath solution $(x^1, x^2, \dots, x^N)$.
2) Obtain the optimal vertex solution $(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^N)$ using Lemma 1.
3) For each user $i = 1, \dots, N$, solve the optimal rounding problem (4) for $\bar{x}^i$ to get a rounded rate vector y^i.

The resulted routing solution $(y^1, y^2, \dots, y^N)$ will be a suboptimal solution to the split ratio granularity problem (2) within the bound of Δ_M (Δ_M is the optimal value of (7)).

The first and second steps above only involve convex optimizations, but we have not demonstrated how to efficiently find the optimal rounding for a rate vector in the last step. In the first part of this section, we will propose an algorithm for solving the optimal rounding problem (4). Furthermore, the multipath relaxation (1) can be quite loose in some cases, so in the second part we will introduce a tighter relaxation to the original problem (2) called convex relaxation, which will lead to an approximation algorithm with better performance guarantee. Finally, a numerical example is employed to demonstrate the effectiveness of the methods.

A. Finding the Optimal Rounding for a Rate Vector

A naive approach to optimally round a rate vector is to enumerate all the possible split ratio vectors and find the one attaining the maximum in (6). This approach is not computationally feasible because it requires searching for exponential number of cases. In the next, we will give an algorithm (Algorithm 1) which can find the optimal rounding by checking only Kp cases in the worst situation.

Given a nonzero rate vector $x \in \mathcal{I}_K$, Algorithm 1 loops through $l = 1, \dots, K$ with $x_l > 0$. For each path l, it finds the smallest integer α_l with

$$\sum_{k=1}^K \left\lfloor \frac{x_k \alpha_l}{x_l} \right\rfloor \geq p. \tag{12}$$

Note that (12) implies

$$\alpha_l = \frac{x_l}{\|x\|_1} \sum_{k=1}^K \frac{x_k \alpha_l}{x_l} \geq \frac{p x_l}{\|x\|_1},$$

so the algorithm starts to find the desired α_l at $\lceil p x_l / \|x\|_1 \rceil$. After the minimal α_l satisfying (12) is found, we can directly construct a split ratio vector α by choosing the given α_l and the other integers

$$0 \le \alpha_k \le \frac{x_k \alpha_l}{x_l},\ k = 1, \ldots, K,\ k \ne l$$

such that $\sum_{k=1}^{K} \alpha_k = p$. As a result, the rate vector y produced at the end of each iteration is feasible to the optimal rounding problem (4).

To see why Algorithm 1 is capable of finding the optimal solution to (4), assume β is the split ratio of the optimal rounding for x. Then there exists some path l such that $\beta_l > 0$, $x_l > 0$ and

$$\frac{x_l}{\beta_l} \le \frac{x_k}{\beta_k},\ k = 1, \ldots, K,$$

so $x_k \beta_l / x_l \ge \beta_k$ and

$$\sum_{k=1}^{K} \left\lfloor \frac{x_k \beta_l}{x_l} \right\rfloor \ge \sum_{k=1}^{K} \beta_k = p.$$

Therefore, β_l also satisfies the inequality (12) and $\beta_l \ge \alpha_l$ where α_l is the integer found in the algorithm. If y is the rate vector in the iteration for path l, then

$$\|y\|_1 = \frac{p x_l}{\alpha_l} \ge \frac{p x_l}{\beta_l} = p \min_{k=1,\ldots,K} \frac{x_k}{\beta_k}.$$

Comparing with (6), we know that y is at least as good as the optimal rounding.

Algorithm 1 (Finding the Optimal Rounding):

Input: A nonzero rate vector $x \in \mathcal{I}_K$.

Output: A rate vector $y \in \mathcal{S}_K$ that is the optimal rounding for x.

```
best ← 0
for l ← 1 to K do
    if x_l > 0 then
        α_l ← ⌈p x_l / ‖x‖_1⌉
        while Σ_{k=1}^{K} ⌊x_k α_l / x_l⌋ < p do
            α_l ← α_l + 1
        end while
        Choose nonnegative integers α_k for k = 1, ..., K,
k ≠ l such that α_k ≤ x_k α_l / x_l and Σ_{k=1}^{K} α_k = p.
        if p x_l / α_l > best then
            best ← p x_l / α_l
            Set y_k ← α_k x_l / α_l for k = 1, ..., K.
        end if
    end if
end for
return y
```

B. Convex Relaxation

As shown in (7), the performance guarantee for our approximation algorithm depends on the maximum throughput loss ρ_K or maximum relative throughput loss ρ_K^R. In some situation, these numbers can be large and consequently the multipath relaxation (1) is very loose. For example, for a user with K paths, where K is much larger than p, the set $\mathcal{I}_K$ contains the vector of all ones, but all rate vectors in $\mathcal{S}_K$ can only use up to p paths. As a result, the maximum throughput loss $\rho_K \ge K - p$.

To design an approximation algorithm with better performance guarantee, we introduce a tighter relaxation by choosing a set $\mathcal{T}_K$ satisfying $\mathcal{S}_K \subseteq \mathcal{T}_K \subseteq \mathcal{I}_K$ to replace the original set $\mathcal{S}_K$ in (2):

$$\begin{aligned} \max \quad & \sum_{i=1}^{N} U^i \left(\|x^i\|_1 \right) \\ \text{s.t.} \quad & \sum_{i=1}^{N} R^i x^i \le c, \\ & x^i \in \mathcal{T}_{K^i}, \quad \forall i = 1, \ldots, N. \end{aligned} \tag{13}$$

To be a tractable optimization problem, the set $\mathcal{T}_K$ above has to be convex. The most natural choice for $\mathcal{T}_K$ is the convex hull of $\mathcal{S}_K$. However, there is a huge difficulty for this choice, which is illustrated as follows: The region shown in Fig. 5(b) is the convex hull of the set $\mathcal{S}_3$ (Fig. 5(a)) for $p = 5$. Because each vertex of the unit cube using more than one path is not in the convex hull, it has to be excluded by one linear constraint, and totally we need to add four extra constraints to get the convex hull of $\mathcal{S}_3$. By generalizing the idea here, for any integer K, we can construct examples in which the description of the convex hull of $\mathcal{S}_K$ requires exponential number of constraints, so the relaxation (13) may still be a hard problem if the convex hull is chosen as $\mathcal{T}_K$.

Instead, we use the following set

$$\mathcal{T}_K = \{x \in \mathcal{I}_K \,|\, \|x\|_1 \le C_K\}$$

to approximate the convex hull of $\mathcal{S}_K$ (see Fig. 5(c)). Here C_K is the maximum throughput over all rate vectors in $\mathcal{S}_K$. Because a nonzero rate vector is in $\mathcal{S}_K$ if and only if it is p-integral, by Lemma 4 we immediately know that

$$C_K = \frac{p}{\lceil p/K \rceil}.$$

Using this particular choice of $\mathcal{T}_K$ in (13), we obtain a new relaxation to the original problem (2), which is referred to as the *convex relaxation*. If p is a multiple of K, then the set $\mathcal{S}_K$ contains the vector of all ones and $C_K = K$, so there is no difference between the convex relaxation and the multipath relaxation. Otherwise, we can get a better approximation algorithm by using convex relaxation instead of multipath relaxation.

In the case of linear utility, following the similar procedure for multipath relaxation (Lemma 1 still holds for the convex relaxation (13)), we can obtain the bound Δ_C of performance guarantee for this new approximation algorithm by solving (7) except that ρ_K is replaced by

$$\rho_K^C = \max_{x \in \mathcal{T}_K} \phi_K(x).$$

To obtain the exact bound, we need to calculate the numbers ρ_K^C. There are two simple cases in which ρ_K^C can be easily calculated. First, if there exists a rate vector $x \in \mathcal{T}_K$

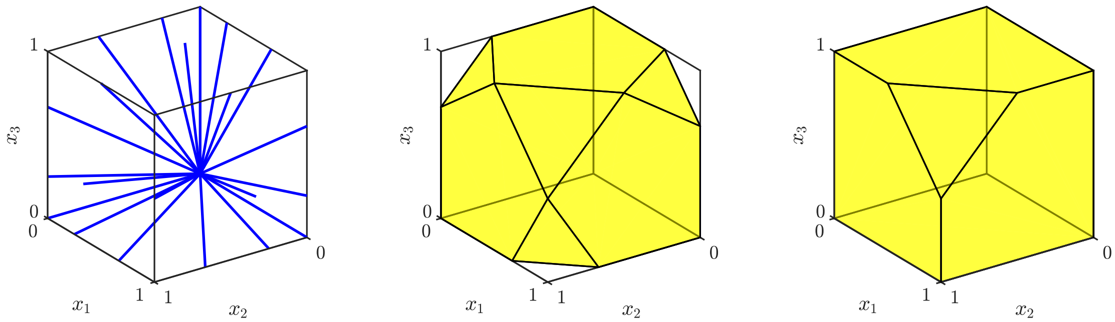


Fig. 5. (a) The set $\mathcal{S}_3$ for $p = 5$ is represented by the solid lines. (b) The convex hull of the set $\mathcal{S}_3$. Note that there is a distinct linear constraint corresponding to each vertex of the cube using more than one path. (c) The set $\mathcal{T}_3$ used in the convex relaxation (13). It preserves the linear constraint in (b) that cuts off the largest part from the cube but removes the other three linear constraints that play minor roles.

achieving the maximum throughput and the maximum relative throughput loss simultaneously, that is $\|x\|_1 = C_K$ and

$$\frac{\phi_K(x)}{\|x\|_1} = \rho_K^R = \frac{K-1}{p+K-1},$$

then we can directly conclude that

$$\rho_K^C = \phi_K(x) = \frac{K-1}{p+K-1} C_K.$$

Second, if there is a rate vector in $\mathcal{T}_K$ whose throughput loss attains the maximum throughput loss ρ_K in the set $\mathcal{I}_K$, then obviously $\rho_K^C = \rho_K$. In fact, the following theorem shows that either of the above two cases will occur.

Theorem 7: There exists a rate vector $x \in \mathcal{I}_K$ whose relative throughput loss is

$$\frac{\phi_K(x)}{\|x\|_1} = \frac{K-1}{p+K-1},$$

making one of the following two statements holds: (i) $\|x\|_1 \geq C_K$; (ii) $\phi_K(x) = \rho_K$.

Proof: The result is obvious for $K = 1$. If $K > 1$, by Lemma 4, there exists an integral rate vector x such that

$$\|x\|_1 = \frac{p+K-1}{\lceil (p+K-1)/K \rceil},$$

and by Lemma 3 its relative throughput loss is

$$\frac{\phi_K(x)}{\|x\|_1} = \frac{K-1}{p+K-1}.$$

If statement (i) does not hold,

$$\frac{p+K-1}{\lceil (p+K-1)/K \rceil} = \|x\|_1 < C_K = \frac{p}{\lceil p/K \rceil}.$$

Let $\lceil p/K \rceil = \Theta$. The only possibility that the above inequality holds is $\lceil (p+K-1)/K \rceil = \Theta + 1$, so this inequality can be rewritten as

$$\frac{p+K-1}{\Theta+1} < \frac{p}{\Theta},$$

which implies that

$$(K-1)\Theta < p. \tag{14}$$

If statement (ii) does not hold either, i.e.,

$$\phi_K(x) = \frac{K-1}{\lceil (p+K-1)/K \rceil} < \rho_K,$$

we will end up with a contradiction. In (11), the maximization is not attained by $\Gamma = p + K - 1$, and thus

$$\rho_K = \frac{\omega}{\Theta} > \frac{K-1}{\Theta+1}, \tag{15}$$

where ω is the integer such that $(p+\omega)/K = \Theta$. Note that $\omega < K - 1$, so

$$\frac{(K-1)\Theta}{\Theta+1} < \omega \leq K-2,$$

which implies

$$\frac{K-2}{K-1} > \frac{\Theta}{\Theta+1},$$

and thus $K - 1 > \Theta + 1$.

On the other hand, we can substitute $\omega = K\Theta - p$ into (15), which gives

$$\frac{K-1}{\Theta+1} < \frac{K\Theta - p}{\Theta} = K - \frac{p}{\Theta} < 1,$$

where the last inequality is from (14). Now we have $K - 1 < \Theta + 1$, which leads to a contradiction. ■

Theorem 7 offers the following approach to calculate ρ_K^C: First, we find the integral rate vector x with the maximum relative throughput loss as described in the proof of Lemma 4. If $\|x\|_1 \geq C_K$, then scale down x to make $\|x\|_1 = C_K$ without changing the relative throughput loss. In this case,

$$\rho_K^C = \frac{K-1}{p+K-1} C_K.$$

If $\|x\|_1 < C_K$, then $x \in \mathcal{T}_K$ and Theorem 7 guarantees that $\phi_K(x) = \rho_K = \rho_K^C$.

Finally, we notice that the maximum relative throughput loss in the convex relaxation

$$\rho_K^{C,R} = \max_{x \in \mathcal{T}_K, x \neq 0} \frac{\phi_K(x)}{\|x\|_1}$$

is the same as that in the multipath relaxation, i.e., $\rho_K^{C,R} = \rho_K^R$. This is because any rate vector can always be scaled to satisfy

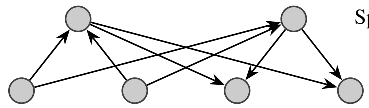


Fig. 6. An example of the leaf-spine topology with two source nodes, two destination nodes and two spine nodes.

TABLE I
NUMERICAL RESULT FOR THE LEAF-SPINE TOPOLOGY

M	K	p	Multipath Relaxation		Convex Relaxation	
			Optimal Value	Rounded Solution	Optimal Value	Rounded Solution
3	10	3	30	9	27	16.5
3	10	15	30	22.5	30	26.3
3	15	3	45	9	27	27
3	15	20	45	30	45	42

the additional constraint $\|x\|_1 \leq C_K$ in the convex relaxation without changing the relative throughput loss. As a result, for the logarithmic utility case, the new algorithm cannot provide tighter performance guarantee compared with the one from multipath relaxation.

C. Numerical Example

In the above discussion, we have given a worst-case analysis for the performance guarantee of the approximation algorithm. However, in many cases the performance of our algorithm can be much better. In this part, we will apply the algorithm to a concrete example.

We consider the common leaf-spine topology (Fig. 6) in data-center networks. In the leaf layer, assume that we have M source nodes and M destination nodes. There is a traffic flow from each source node to each destination node, so the number of users (source-destination pairs) is $N = M^2$. In the spine layer, suppose there are K spine nodes. Each spine node has links connecting to all the source and destination nodes, and all links have the unit capacity. As a result, each user has K available paths from its source to its destination, each of which corresponds to a spine node. Previous works such as CONGA [12] focused on how to implement multipath routing in this topology, but they did not consider the restrictions on how routers split traffic. Now we will apply our algorithm to the split ratio granularity problem (2) for the leaf-spine topology and linear utility.

Table I lists the optimal values of multipath relaxation (1), convex relaxation (13) and the objective values of solutions obtained from the rounding of the optimal solutions to the two relaxations for different combinations of M, K and p. It is easy to see that the optimal value opt_M of the multipath relaxation (1) is $L/2$, where $L = 2MK$ is the number of links in the network. On the other hand, although the accurate optimal value opt_G of the original problem (2) is not provided due to computational complexity, we know that it must be between the optimal value of the convex relaxation and the value of rounded solution from convex relaxation.

From Table I we can see in the case that p is large, the optimal values of the multipath relaxation and convex relaxation are the same. In terms of the optimal value, the convex relaxation is not tighter than the multipath relaxation. However, the performance of rounded solution from convex relaxation is still better than the one from multipath relaxation. When p is small, the benefit of using convex relaxation is even more evident. In addition, when p is large or K is large, the performance loss during the rounding of the optimal solution from convex relaxation becomes smaller, and we can expect that the solution we found is closer to the actual optimal solution to the original split ratio granularity problem (2).

V. CONCLUSION

Routing optimization with split ratio granularity constraints is investigated. Normally, if the split ratio granularity gets finer (p increases), or if there are fewer available paths for each user (K decreases), the performance loss due to inflexibility in traffic split tends to be smaller. In this paper, we obtain explicit upper bounds to quantitatively characterize the performance loss, which certifies these intuitions.

Based on convex relaxation of the original problem, we further develop an efficient approximation algorithm to compute a good routing solution that satisfies the split ratio granularity constraints. We also quantify the difference between our solution and the solution to the original problem, which can be viewed as another layer of cost of not being able to split arbitrarily.

REFERENCES

[1] R. G. Gallager, “A minimum delay routing algorithm using distributed computation,” *IEEE Transactions on Communications*, vol. 25, no. 1, pp. 73–85, Jan. 1977.

[2] N. Michael and A. Tang, “HALO: Hop-by-hop adaptive link-state optimal routing,” *IEEE/ACM Transactions on Networking*, vol. 23, no. 6, pp. 1862–1875, Dec. 2015.

[3] Y. Bi, C. W. Tan, and A. Tang, “Network utility maximization with path cardinality constraints,” in *IEEE INFOCOM 2016*, Apr. 2016.

[4] K. Németh, A. Kőrösi, and G. Rétvári, “Optimal OSPF traffic engineering using legacy equal cost multipath load balancing,” in *2013 IFIP Networking Conference*, May 2013.

[5] Y. Lee, Y. Seok, Y. Choi, and C. Kim, “A constrained multipath traffic engineering scheme for MPLS networks,” in *IEEE ICC 2002*, vol. 4, Apr. 2002, pp. 2431–2436.

[6] M. Wang, C. W. Tan, W. Xu, and A. Tang, “Cost of not splitting in routing: Characterization and estimation,” *IEEE/ACM Transactions on Networking*, vol. 19, no. 6, pp. 1849–1859, Dec. 2011.

[7] B. Fortz and M. Thorup, “Increasing Internet capacity using local search,” *Computational Optimization and Applications*, vol. 29, no. 1, pp. 13–48, Oct. 2004.

[8] M. Chiesa, G. Kindler, and M. Schapira, “Traffic engineering with equal-cost-multipath: An algorithmic perspective,” *IEEE/ACM Transactions on Networking*, vol. 25, no. 2, pp. 779–792, Apr. 2017.

[9] S. H. Low, “A duality model of TCP and queue management algorithms,” *IEEE/ACM Transactions on Networking*, vol. 11, no. 4, pp. 525–536, Aug. 2003.

[10] J. P. Aubin and I. Ekeland, “Estimates of the duality gap in nonconvex optimization,” *Mathematics of Operations Research*, vol. 1, no. 3, pp. 225–245, Aug. 1976.

[11] X. Liu, S. Mohanraj, M. Pióro, and D. Medhi, “Multipath routing from a traffic engineering perspective: How beneficial is it?” in *IEEE ICNP 2014*, Oct. 2014, pp. 143–154.

[12] M. Alizadeh *et al.*, “CONGA: Distributed congestion-aware load balancing for datacenters,” *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 4, pp. 503–514, Oct. 2014.