Splittable Single Source-Sink Routing on CMP Grids: A Sublinear Number of Paths Suffice

Adrian Kosowski, Przemyslaw Uznanski

Splittable Single Source-Sink Routing on CMP Grids: A Sublinear Number of Paths Suffice

Adrian Kosowski, Przemysław Uznański
INRIA Bordeaux – Sud-Ouest
33400 Talence, France
Emails: {adrian.kosowski, przemyslaw.uznanski}@inria.fr

***Abstract*—We investigate the routing of communications in grid graphs, representing chip multiprocessors. We consider the k-MP and Max-MP models of Manhattan-path routing introduced by Benoit *et al.* (IPDPS 2012). In Max-MP, each of the communication requests can be split into arbitrarily many communication paths between the source and terminal node, while in k-MP, it can be split into at most k such paths, where $k \geq 1$ is a parameter of the model. The power cost associated with activating a communication link at a transmission speed of b bytes/second is proportional to b^{α}, for some constant exponent $\alpha > 2$.**

Our first result is to establish a trade-off between the power cost of single source-sink k-MP routing, and the value of the model parameter k. Our results imply that above the threshold value of $k = \Theta(n^{1/(\alpha-1)})$, the optimal solution to k-MP for grids of dimensions $\Theta(n)$ is the same as that of Max-MP, up to a constant factor. We also show that this threshold is tight for the natural single-request instance, i.e., it is necessary to split such a request into $\Theta(n^{1/(\alpha-1)})$ paths to achieve asymptotically-optimal cost. Our analysis relies on new algorithmic schemes for k-MP routing in grids.

Our second contribution is to provide efficient algorithms for solving k-MP. We point out that whereas the minimum-cost k-MP routing problem is NP-hard, there exist polynomial-time algorithms for k-MP routing which provides a constant approximation of the optimum cost. We close the paper by showing experimental evidence that for practical instances, our algorithm for k-MP achieves a power cost close to that of the optimal solution to Max-MP, starting from the above-discussed threshold value of k.

I. Introduction

In this paper we consider the design of single chip multiprocessors (CMP). The increase in the level of integration of CMPs creates demand for high-speed communication on-chip, which in turn increases the power consumption on CMP, and this trend is predicted to continue in the future [8]. A significant part of power is consumed by maintaining communications within the chip, and that makes efficient allocation of communication routes a very important issue [16].

We consider the problem of optimizing the power consumption cost of communication between two given cores. Our power consumption model assumes that if an edge is transmitting at rate f, the power cost of maintaining the frequency over an edge is proportional to f^{α} for a given constant $\alpha > 2$, identical for every edge. For purposes of optimization, we neglect the static part of power usage over edge, assuming that dynamic part (associated with transmission) is dominant for high communication rate.

Such a model of power consumption was studied in the context of splittable Manhattan-path routing by [5]. They introduced several routing schemes in an effort to minimize the total power cost, but observed that this may require the splitting of each communication request, and routing its fragments along a potentially very large number of communication paths. Splitting a request, taking care of the route for each part, and merging it at the target is very and imposes additional time and power overhead, so our goal in this work is to show how to limit splitting as much as possible, without excessively increasing communication cost. We address this problem by studying the family of k-MP routing problems, in which each request has an imposed bound $k \geq 1$ on the number of paths it can be split into.

A. Related work

Numerous studies concern the optimization of power cost in integrated chip designs, which take into account that both processors and communication buses may operate at variable frequency, determining the speed of computations or transmissions (cf. [9], [13], [15], [17]). The increase of power cost with the third power of workload in such designs is a well-established relation (cf. e.g. [17] [3], [7]). On CMP grids, links with dynamic frequency and/or voltage scaling are used ([14], [19]), and the dissipated power P on a link is related to the frequency f and voltage V on it by the following relation supported by both theory and experiments (cf. e.g. [2]):

$$P \sim f \cdot V^2$$

However, for most designs, an increase in operating frequency also results in an increase in voltage, roughly according tot the relation $V \sim f$ (cf. e.g. [19]), which gives the relation between power cost and transmission speed ([5], [6]):

$$P \sim f^3.$$

The study of power-aware routing with Manhattan paths was inaugurated in [5], who provided bounds on the power cost of source-sink routing, for requests which can be arbitrarily split, and supported them by experimental study. Our work considers the case in which each request can be split into at most k paths, bearing a relation to the so called k-splittable max-flow problem, which has been previously studied in the literature [4]. A discussion of different routing policies on a 2D grid, with latency and throughput as main measures, is provided in [18].

B. *Outline and results*

Our study concerns routing between a single source-sink pair of nodes using Manhattan paths on a grid CMP. Communication between these nodes is assumed to be static, i.e. constant in time, and the cost of a transmission along an edge is assumed to be proportional to a fixed power of the transmission rate. The considered model, power cost function, and rules of routing are formally presented in Section II.

In Section III, we briefly outline the theory of Manhattan-path routing with arbitrarily splittable requests (Max-MP). We provide an optimal convex programming formulation of the problem, leading to a routing schemed denoted as $\textsc{Opt}$, and recall the properties of the $\mathcal{C}$ routing scheme introduced in [5]. We also provide a convenient formulation of Manhattan routings in terms of transmission through nodes.

Our main results are given in Section IV. They concern the variant of the Manhattan routing problem in which each request can be satisfied by at most k communication paths, where k is a parameter of the model (k-MP). We study the value of the ratio of the cost of the optimal solution in this case, denoted $\textsc{Opt}_k$, to the cost of the routing scheme $\textsc{Opt}$ with arbitrarily splittable paths. We establish that in general, $\text{cost}(\textsc{Opt}_k)/\text{cost}(\textsc{Opt}) = O(1 + \frac{n}{k^{\alpha-1}})$, whereas for the special case of $d \geq 1$ identical requests of the same size, this ratio is given precisely as $\Theta(1 + \frac{n}{(kd)^{\alpha-1}})$. This means that for $k = o(n^{1/(\alpha-1)})$, the requirement that requests can be split into at most k paths impacts the cost of the routing scheme asymptotically, i.e., increases the cost by an unbounded factor for sufficiently large n. On the other hand, for k larger than the threshold value of $\Theta(n^{1/(\alpha-1)})$, the obtained k-splittable routings are within a constant factor of the optimal solution to Max-MP.

The proposed bounds are obtained through the analysis of three efficiently implementable algorithmic schemes for solving k-MP: $\mathcal{F}_k$ routing and $\mathcal{D}_k$ routing (for uniform requests), and $\mathcal{A}_k$ (for non-uniform requests). The latter two are shown to have a constant approximation ratio with respect to the cost of $\textsc{Opt}_k$ for all k, while the former converges to the cost of $\textsc{Opt}$ as k tends to infinity. The design of such approximate techniques results from the observation that $\textsc{Opt}_k$ is NP-hard, which we make in passing.

Finally, in Section V, we perform an experimental validation of the determined threshold value of $k = \Theta(n^{1/(\alpha-1)})$, showing the effect of smaller and larger values of k on the cost of the routing. We also experimentally compare the performance on $\mathcal{F}_k$ routing and $\mathcal{D}_k$ routing, studying their convergence to asymptotic behavior for increasing values of k and different values of the power cost exponent $\alpha \approx 3$.

Some closing remarks are provided in Section VI.

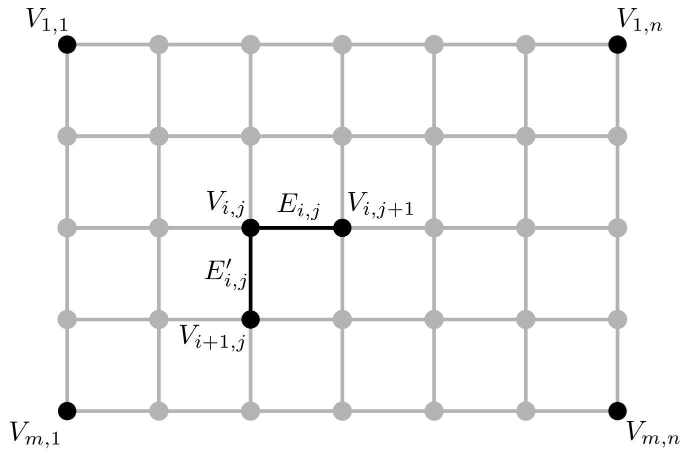

Figure 1. Notation for the CMP grid model

II. Framework

Our study is performed in exactly the same setting of single source-sink Manhattan routing as that introduced in [5]. In this section we recall the model assumptions and problem definitions.

A. *Platform and power consumption model*

Our platform is a modeled as a graph composed of a $m \times n$ uniform nodes $V_{i,j}$, with $1 \leq i \leq m$ and $1 \leq j \leq n$, arranged into a rectangular grid. Without the loss of generality, we assume that $m \geq n$. We will also assume for the purpose of analysis that the sides of the grid are of the same order of magnitude, i.e., $m = O(n)$. Nodes are connected by bidirectional edges. The horizontal edge $E_{i,j}$ connects $V_{i,j}$ and $V_{i,j+1}$ (for $1 \leq i \leq m$, $1 \leq j \leq n-1$), and the vertical edge $E'_{i,j}$ connects $V_{i,j}$ and $V_{i+1,j}$ ($1 \leq i \leq m-1$, $1 \leq j \leq n$), see Fig. 1 for an illustration.

The power consumed on each edge is closely related to the amount of data sent through this edge in a unit of time. To simplify analysis of the model, we discard constant factors, and (following [5]) set the cost of transmission at rate x as $C(x) = x^{\alpha}$, where $\alpha > 2$ is an absolute constant of the model (it is reasonable to assume $\alpha \approx 3$). In our model, we discard the static part of power leakage in our model, especially since static part could disturb convexity of the cost function, and the importance of static leakage is predicted to be less important in future designs (cf. [5]).

B. Communication and routing rules

The study of routing with Manhattan-type paths (of shortest length) is motivated by practical concerns, in particular, the need to minimize communication latency, and to confine communications between nearby processors to a local area of the grid. For the purpose of the study of single source-sink communications, it is assumed that the source and target are placed in the opposite corners of the grid; for communications between a different pair of nodes, considerations can be restricted to the respective rectangular sub-grid.

The *routing* R of a single communication request of size s is a weighted set of paths, $\{(w_1, p_1), \ldots, (w_k, p_k)\}$, where each path p_i starts at the same source vertex $V_{1,1}$, and ends at the same target vertex $V_{m,n}$ in the opposite corner of the grid. The real-valued weights w_i satisfy $w_i \geq 0$ and $\sum_i w_i = s$. This definition of a routing extends naturally to a set of $d \geq 1$ requests, which maybe *uniform* (with identical request size $s = K/d$), or *non-uniform* (with possibly distinct request sizes $s_1, \ldots, s_d$).

Given a routing R, we define $R(e)$ as the size of the transmission going through an edge e, i.e.: $R(e) = \sum_{i:\, e \in p_i} w_i$. (We will adapt this notation accordingly for routings denoted by letters different from R.)

The *routing policy* is expressed through the upper bound k on the splitability of each request:

- In *Single-Path Manhattan Routing (*1-MP*)*, communication for each request may take place along one path only ($k = 1$).
- In *k-Path Manhattan Routing (*k-MP*)*, communication for each request can be split into any number of $k' \leq k$ (partially overlapping) source-sink paths, where k is a parameter of the model.
- In *Max-Paths Manhattan Routing (*Max-MP*)*, the number of paths allowed for each request is unbounded ($k = +\infty$).

C. Problem definition

For a given routing policy with parameter k and power coefficient α, we define our optimization problem as follows: *Given a $m \times n$ grid and a set of requests of sizes $(s_1, \ldots, s_d)$, find a routing R of this set of requests minimizing the total power cost of transmission through all the edges of the grid, expressed by the cost function:*

$$\operatorname{cost}(R) = \sum_{i=1}^{m} \sum_{j=1}^{n-1} R(E_{i,j})^{\alpha} + \sum_{i=1}^{m-1} \sum_{j=1}^{n} R\left(E'_{i,j}\right)^{\alpha}$$

We remark that in a real-world scenario, the size of a request should be measured in a fixed unit of time, i.e., as the throughput required to process the request.

III. Preliminaries: The Max-MP Problem

In this section we present an algorithmic characterization of the routing problem with arbitrarily splittable requests. We start by providing a formal convex programming formulation of the optimal routing strategy, OPT. We also outline a different routing strategy, denoted $\mathcal{C}$. Although the procedure for $\mathcal{C}$ leads to identical routings to that used in [5], we provide a different description of the algorithm, to enable its easier adaptation to the case of k-MP routing.

A. Optimal solution to Max-MP routing

When considering Max-MP routing, since all requests share endpoints, we can merge the considered set of requests of sizes $\{s_1, \ldots, s_d\}$ into one request of size $K = \sum_i s_i$. Now, finding the optimal solution to Max-MP can be formulated as a convex programming instance given in Algorithm 1.

Algorithm 1 OPT routing {for Max-MP}

Input: A set of arbitrarily splittable requests of total size K in a $m \times n$ grid.
Solve:

$$\text{Minimize } \left(\sum_{i,j} \mathrm{OPT}(E_{i,j})^{\alpha} + \sum_{i,j} \mathrm{OPT}\left(E'_{i,j}\right)^{\alpha} \right)$$

Subject to:

$$\mathrm{OPT}(E_{i,j}) \geq 0$$
$$\mathrm{OPT}\left(E'_{i,j}\right) \geq 0$$
$$\mathrm{OPT}\left(E_{i,j-1}\right) + \mathrm{OPT}\left(E'_{i-1,j}\right) = \mathrm{OPT}(E_{i,j}) + \mathrm{OPT}\left(E'_{i,j}\right)$$
$$\mathrm{OPT}(E_{1,1}) + \mathrm{OPT}\left(E'_{1,1}\right) = K$$

Output: values of $\mathrm{OPT}(E_{i,j}), \mathrm{OPT}\left(E'_{i,j}\right)$

Applications of convex programming algorithms lead to polynomial-time schemes providing an arbitrarily good approximation of OPT (cf. e.g. [1], [11] for a discussion of convex programming in the context of routing schemes).

We remark on the following lower bound on the size of OPT. Consider any Max-MP routing which transmits requests of total size S. The edges adjacent to node $V_{1,1}$, i.e., $\{E_{1,1}, E'_{1,1}\}$, have to transmit K in total. It follows that:

$$\operatorname{cost}(\mathrm{OPT}) \geq \left(\frac{K}{2}\right)^{\alpha} = \Theta(K^{\alpha}). \tag{1}$$

Remarkably, as shown in [5], this lower bound is tight regardless of the size of the grid, since it can be achieved using the routing scheme $\mathcal{C}$ described later in this section.

B. Techniques: edge and vertex costs for grid diagonals

In order to describe the routing schemes proposed in this paper, we will attempt to perform "load balancing" of paths with respect to transmission through vertices rather than edges. Hence, in a similar fashion to the notation $R(e)$ for an edge e, we define $R(v)$ as the total transmission size going through a vertex v in routing R.

We introduce the notion of the l-th *vertex diagonal*, denoted as DV_l ($1 \leq l \leq n+m-1$) by splitting the set

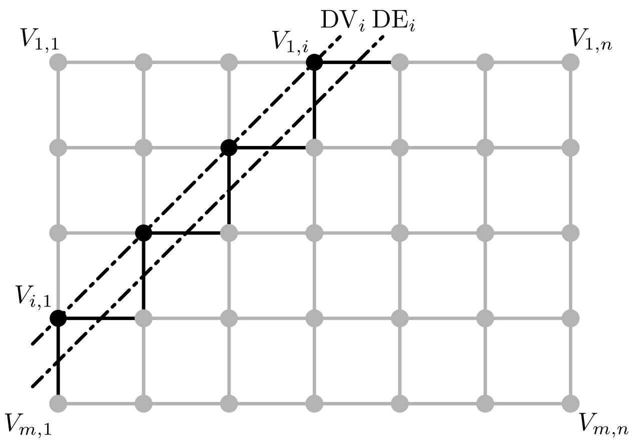

Figure 2. Vertex diagonal DV_i and edge diagonal DE_i

of vertices according to their distance from the source, as follows (see Fig. 2 for an illustration):

$$V_{i,j} \in \mathrm{DV}_l, \text{ if } i+j = l+1$$

Likewise, by the l-th *edge diagonal*, denoted DE_l ($1 \le l \le n+m-2$), we mean the set of edges connecting vertices from DV_l and DV_{l+1}:

$$E_{i,j}, E'_{i,j} \in \mathrm{DE}_l, \text{ if } i+j = l+1$$

It is worth noting that:

$$|\mathrm{DV}_l| = \begin{cases} l & \text{if } 1 \le l \le n \\ n & \text{if } n \le l \le m \\ n+m-l & \text{if } m \le l \le n+m-1 \end{cases} \tag{2}$$

$$|\mathrm{DE}_l| = \begin{cases} 2 \cdot l & \text{if } 1 \le l < n \\ 2 \cdot n - 1 & \text{if } n \le l < m \\ 2 \cdot (n+m-1-l) & \text{if } m \le l < n+m-1 \end{cases} \tag{3}$$

We start by observing that the values of $R(v)$ uniquely determine the values of $R(e)$. This property will allow us to design routing schemes simply by setting $R(v)$ for all nodes.

Lemma 3.1: For arbitrary $R \in \mathrm{MP}$, we can calculate values of $R(e)$ from values of $R(v)$.

Proof: We will look at DE_j for various values of j. We denote $\mathrm{DV}_j = \{v_1, v_2, \dots\}$ and $\mathrm{DV}_{j+1} = \{u_1, u_2, \dots\}$, with vertices reverse-ordered by first coordinate. Also, the edges in $\mathrm{DE}_j = \{e_1, \dots\} \cup \{e'_1, \dots\}$ are assumed to be ordered respectively horizontal and vertical edges.

1) Let $1 \ge j < n$, and $1 \le i \le j$. If we expand the sum of i leftmost vertices from DV_j, we get:

$$R(v_1) + \ldots + R(v_i) =$$
$$= \left(R\left(e'_1\right) + R(e_1)\right) + \ldots + \left(R\left(e'_i\right) + R(e_i)\right)$$

Expanding the other sum:

$$R(u_1) + \ldots + R(u_i) =$$
$$= R\left(e'_1\right) + \left(R(e_1) + R\left(e'_2\right)\right) + \ldots + \left(R(e_{i-1}) + R\left(e'_i\right)\right)$$

So we obtain:

$$R(e_i) = (R(v_1) + \ldots + R(v_i)) - (R(u_1) - \ldots - R(u_i))$$

which directly gives us the formula for the weight of the i-th horizontal edge of DE_j.
In a similar fashion we establish:

$$R\left(e'_i\right) = (R(u_1) + \ldots + R(u_i)) - (R(v_1) + \ldots + R(v_{i-1}))$$

which gives the formula for the weight of the i-th vertical edge of DE_j.

2) For $n \le j < m$, we can perform similar reasoning:

$$R(e_i) = (R(v_1) + \ldots + R(v_i)) - (R(u_1) + \ldots + R(u_{i-1}))$$
$$R\left(e'_i\right) = (R(u_1) + \ldots + R(u_i)) - (R(v_1) + \ldots + R(v_i)) = 0$$

3) For $m \le j < n+m$:

$$R(e_i) = (R(v_1) + \ldots + R(v_i)) - (R(u_1) + \ldots + R(u_{i-1}))$$
$$R\left(e'_i\right) = (R(u_1) + \ldots + R(u_i)) - (R(v_1) + \ldots + R(v_i))$$

■

Next, we remark on the close connections between the costs measured in terms of nodes and edges. For any Manhattan routing R we have:

$$R\left(E_{i,j-1}\right) + R\left(E'_{i-1,j}\right) = R\left(V_{i,j}\right) = R\left(E_{i,j}\right) + R\left(E'_{i,j}\right)$$

Since the cost function $f(x) = x^\alpha$ is convex, we get:

$$R\left(V_{i,j}\right)^\alpha \ge R\left(E_{i,j}\right)^\alpha + R\left(E'_{i,j}\right)^\alpha$$

which leads to useful inequalities between the costs related to vertices and edges:

$$\sum_{v \in \mathrm{DV}_l} R(v)^\alpha \ge \sum_{e \in DE_l} R(e)^\alpha \tag{4}$$
$$\sum_{v \in \mathrm{DV}_l} R(v)^\alpha \ge \sum_{e \in DE_{l-1}} R(e)^\alpha . \tag{5}$$

C. Routing scheme $\mathcal{C}$ for Max-MP

We define the routing scheme $\mathcal{C}$ for Max-MP by putting a limit on the transmission going through vertices. Since each diagonal of vertices has a total transmission of exactly K, we choose to set an equal value of transmission for all vertices in the layer:

$$\forall_{v \in \mathrm{DV}_j} \mathcal{C}(v) = \frac{K}{\left|\mathrm{DV}_j\right|}.$$

Taking into account equation (2), this gives us the following value of transmission for vertices of the grid:

$$\mathcal{C}\left(V_{i,j}\right) = \begin{cases} \frac{1}{i+j-1} & \text{if } 1 \le i+j-1 \le n \\ \frac{1}{n} & \text{if } n \le i+j-1 \le m \\ \frac{1}{n+m+1-i-j} & \text{otherwise} \end{cases} \tag{6}$$

To verify that this routing is well-defined, we can calculate transfers over each edge, applying the construction from the proof of Lemma 3.1. The obtained routing scheme is presented as Algorithm 2, while examples of transfers obtained using this algorithm are shown in Fig. 3.

Algorithm 2 $\mathcal{C}$ routing {for Max-MP, cf. [5]}

Input: A set of arbitrarily splittable requests of total size K in a $m \times n$ grid.

Solution: For each diagonal DE_j of the grid, $1 \leq j < n+m$, set the flow on its successive horizontal edges e_i and vertical edges e'_i, $1 \leq i \leq j$, as follows:

- If $1 \leq j < n$, set:

$$\mathcal{C}(e_i) := K\frac{i}{j} - K\frac{i}{j+1}$$
$$\mathcal{C}(e'_i) := K\frac{i}{j+1} - K\frac{i-1}{j}$$

- If $n \leq j < m$, set:

$$\mathcal{C}(e_i) := K\frac{i}{n} - K\frac{i-1}{n} = \frac{K}{n}$$
$$\mathcal{C}(e'_i) := 0$$

- If $m \leq j < n+m$, set:

$$\mathcal{C}(e_i) := K\frac{i}{n+m-j} - K\frac{i-1}{n+m-j-1}$$
$$\mathcal{C}(e'_i) := K\frac{i}{n+m-j-1} - K\frac{i}{n+m-j}$$

It corresponds precisely to the algorithm studied in [5], who showed that $\mathcal{C}$ admits a constant approximation ratio for Max-MP.

Theorem 3.2 ([5]): $\mathrm{cost}(\mathcal{C}) = \Theta(\mathrm{cost}(\textsc{Opt})) = \Theta(K^\alpha)$.

We remark that the formulation of Algorithm 2 does not specify the mapping from the transfer on edges to paths used in the routing. It is remarked in [5] that there are $\binom{n+m-2}{n-1} = \Theta(2^n)$ different paths from source to sink, hence the number of paths used in a $\mathcal{C}$ routing could be exponentially large.

To obtain a better bound, we recall the following folklore observation: for a given graph $G = (V,E)$ and flow f on G, f can be represented as the union of at most $|E|$ weighted paths. (Indeed, it is enough to repeat following process: choose path p on f linking source to sink, and subtract it with maximal possible weight from the flow f. In this way, f loses one arc at each step. So, after at most $|E|$ steps, f is empty.) It follows that both $\textsc{Opt}$ routing and $\mathcal{C}$ routing provide a solution to Max-MP which uses $O(nm)$ splits per request.

In the next section, we will prove that it is possible to preserve a constant approximation ratio of the optimal cost, while using a much smaller number of splits, sublinear in the dimensions of the grid.

IV. Schemes for k-MP Routing

In this section, we present three schemes for solving the k-Path Manhattan Routing problem (k-MP). The first two, denoted $\mathcal{F}_k$ and $\mathcal{D}_k$, are designed for uniform sets of requests. As the bound k on the number of allowed paths per request tends to infinity, these approaches will be shown to converge to the performance of schemes $\textsc{Opt}$ and $\mathcal{C}$ for Max-MP, respectively. The third scheme, denoted $\mathcal{A}_k$, is an extension of $\mathcal{D}_k$ which also works for non-uniform sets of requests.

A. 1-MP *routing with uniform requests*

We start by considering the 1-MP routing policy, meaning that no splitting of requests is allowed. We can treat this problem as a discrete version of a continuous Max-MP problem. First, we will consider uniform requests (of equal sizes); without loss of generality, we can assume that the input consists of d requests of size 1, each.

This considered problem can be solved by the flow-based $\mathcal{F}_1$ routing approach presented in Algorithm 3. The obtained solution is optimal, i.e., for uniform instances, we have $\mathrm{cost}(\textsc{Opt}_1) = \mathrm{cost}(\mathcal{F}_1)$. Moreover, using a classical min-cost

Algorithm 3 $\mathcal{F}_1$ routing {for uniform 1-MP}

Input: A set of d unsplittable requests of size $s = 1$ in a $m \times n$ grid.

Solution:

1) Construct a multigraph G' such that $V(G') = V(G)$.
2) For every directed edge $e \in E(G)$, add d weighted directed edges to G', having the same endpoints as e, and weights given as:
$$1^\alpha, 2^\alpha - 1^\alpha, \ldots, d^\alpha - (d-1)^\alpha$$
3) Return the min-cost flow of size d in G', using the two opposite corners of the grid as the source and sink.

flow algorithm, a $\mathcal{F}_1$ routing can be found in polynomial time with respect to parameters n, m, and d.

We will now provide asymptotic bounds on the size of the (optimal) solution to the uniform 1-MP problem. We obtain the lower bound by combining the lower bound for problem Max-MP (formula (6) with $K = d$), with an additional factor resulting from the discrete nature of 1-MP.

Lemma 4.1: For every $R \in$ uniform 1-MP:

$$\mathrm{cost}(R) = \Omega(d^\alpha) + \Omega(nd)$$

Proof: The bound $\Omega(d^\alpha)$ holds since we cannot get a better solution than that for Max-MP with the same set of requests. Since each message of size 1 creates a cost at least $n+m-2$, by the convexity of the cost function, we get a total cost of $\Omega(nd)$ for d of them. ■

To provide a complementary upper bound, we do not analyze the optimal $\mathcal{F}_1$ routing scheme, but instead propose

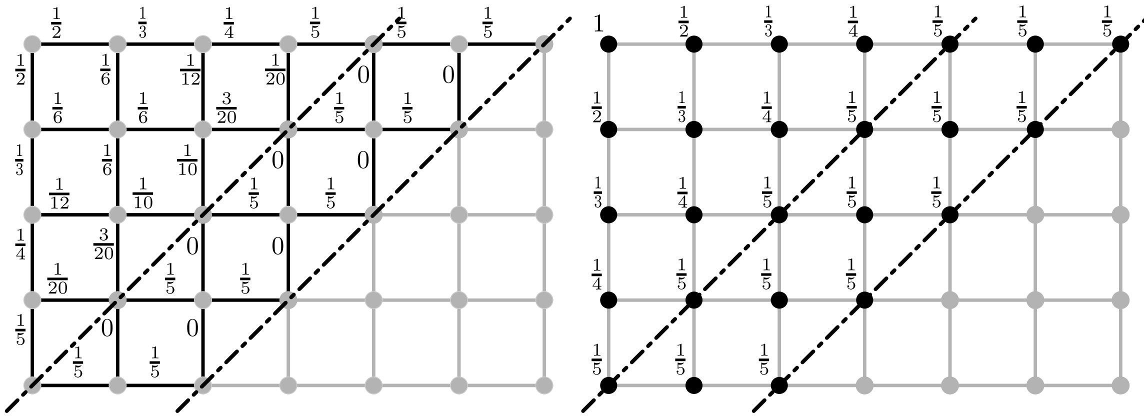

Figure 3. Example of $\mathcal{C}$ routing on a 5×7 grid, with highlighted values of transfer on edges (on the left) and vertices (on the right)

an approximation scheme called $\mathcal{D}_1$ routing, which proves to be easier to analyze.

We design the $\mathcal{D}_1$ routing through a discretization of the construction of $\mathcal{C}$ routing proposed in the previous section for Max-MP.

Similarly to equation (6), we will place limits on the size of the transfer going through vertices. Consider the vertex diagonal DV_p with $1 \le p \le n+m-1$, and let $i = \left|\mathrm{DV}_p\right|$. Suppose that the vertices of DV_p are ordered by decreasing first coordinate, as $\mathrm{DV}_p = \{v_1, \ldots, v_i\}$. Then, for $1 \le j \le i$, we successively set $\mathcal{D}_1(v_j)$ so that the following equality holds at each step:

$$\mathcal{D}_1(v_1) + \ldots + \mathcal{D}_1(v_j) = \left\lfloor d \cdot \frac{j}{i} \right\rfloor$$

This is achieved by setting:

$$\mathcal{D}_1(v_j) = \left\lfloor d \cdot \frac{j}{i} \right\rfloor - \left\lfloor d \cdot \frac{j-1}{i} \right\rfloor \tag{7}$$

To verify the correctness of this construction, we once again deduce transfer values over vertical and horizontal edges from values over vertices, using Lemma 3.1. A formal implementation of $\mathcal{D}_1$ routing is provided in Algorithm 4, whereas an exemplary comparison of the vertex and edge transfers for $\mathcal{C}$ routing and $\mathcal{D}_1$ routing is shown in Fig. 4.

We start the analysis of the cost of $\mathcal{D}_1$ routing with the following lemma.

Lemma 4.2: Let DV be an arbitrary vertex diagonal, and let $|\mathrm{DV}| = i$. Then:

$$\sum_{v \in DV} \mathcal{D}_1(v)^\alpha = \begin{cases} i\left(\left(\frac{d}{i}\right)^\alpha + O\left(\left(\frac{d}{i}\right)^{\alpha-2}\right)\right), & \text{for } i < d \\ d, & \text{for } i \ge d. \end{cases}$$

Proof: By equation (7) we have:

$$\sum_{v\in \mathrm{DV}} \mathcal{D}_1(v)^\alpha = \sum_{j=1}^{i} \left(\left\lfloor d \cdot \frac{j}{i} \right\rfloor - \left\lfloor d \cdot \frac{j-1}{i} \right\rfloor\right)^\alpha =$$

$$= \left\lfloor \frac{d}{i} \right\rfloor^\alpha \cdot (i - (d \bmod i)) + \left(\left\lfloor \frac{d}{i} \right\rfloor + 1\right)^\alpha \cdot (d \bmod i)$$

We consider two cases:

- ($i < d$) Substituting $x = \frac{d}{i}$, $\lambda = \frac{d}{i} - \left\lfloor \frac{d}{i} \right\rfloor = \frac{d \bmod i}{i}$:
$$\sum_{v\in \mathrm{DV}} \mathcal{D}_1(v)^\alpha = (x-\lambda)^\alpha (1-\lambda)\, i + (x+1-\lambda)^\alpha \lambda i =$$
(using Taylor's series theorem, for some $x - \lambda \le u_1 \le x \le u_2 \le x+1-\lambda$)
$$= \left(x^\alpha - \lambda \alpha x^{\alpha-1} + \frac{\lambda^2}{2}\alpha(\alpha-1)u_1^{\alpha-2}\right)(1-\lambda)\, i +$$

Algorithm 4 $\mathcal{D}_1$ routing {for uniform 1-MP}

Input: A set of d unsplittable requests of size $s = 1$ in a $m \times n$ grid.

Solution: For each diagonal DE_j of the grid, $1 \le j < n+m$, set the flow on its successive horizontal edges e_i and vertical edges e'_i, $1 \le i \le j$, as follows:

- If $1 \le j < n$, set:
$$\mathcal{D}_1(e_i) = \left\lfloor d\frac{i}{j} \right\rfloor - \left\lfloor d\frac{i}{j+1} \right\rfloor$$
$$\mathcal{D}_1(e'_i) = \left\lfloor d\frac{i}{j+1} \right\rfloor - \left\lfloor d\frac{i-1}{j} \right\rfloor$$
- If $n \le j < m$, set:
$$\mathcal{D}_1(e_i) = \left\lfloor d\frac{i}{n} \right\rfloor - \left\lfloor d\frac{i-1}{n} \right\rfloor$$
$$\mathcal{D}_1(e'_i) = 0$$
- If $m \le j < n+m$, set:
$$\mathcal{D}_1(e_i) = \left\lfloor d\frac{i}{n+m-j} \right\rfloor - \left\lfloor d\frac{i-1}{n+m-j-1} \right\rfloor$$
$$\mathcal{D}_1(e'_i) = \left\lfloor d\frac{i}{n+m-j-1} \right\rfloor - \left\lfloor d\frac{i}{n+m-j} \right\rfloor$$

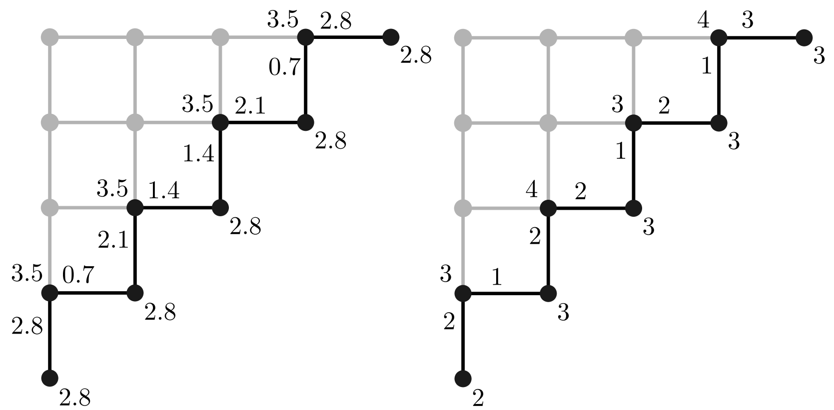

Figure 4. Comparison of transfer values over one diagonal for a $\mathcal{C}$ routing with $K = 14$ (on the left) and a $\mathcal{D}_1$ routing with $d = 14$ (on the right)

$$+\left(x^{\alpha}+(1-\lambda)\,\alpha x^{\alpha-1}+\frac{(1-\lambda)^2}{2}\alpha(\alpha-1)\,u_2^{\alpha-2}\right)\lambda i \le$$

(since $u_1 \le x$ and $u_2 \le x+1-\lambda \le 2x$)

$$\le ix^{\alpha}+i\alpha(\alpha-1)\left(\frac{\lambda^2}{2}(1-\lambda)+\frac{(1-\lambda)^2}{2}\lambda 2^{\alpha}\right)x^{\alpha-2}=$$

$$= i\left(x^{\alpha}+O\left(x^{\alpha-2}\right)\right)=i\left(\left(\frac{k}{i}\right)^{\alpha}+O\left(\left(\frac{d}{i}\right)^{\alpha-2}\right)\right).$$

- ($i \ge d$) The sum simplifies to:

$$\sum_{v\in \mathrm{DV}} \mathcal{D}_1(v)^{\alpha} = 0\cdot(i-(d \bmod i))+1\cdot(d \bmod i)=d.$$

■

We are now ready to show that the cost of a $\mathcal{D}_1$ routing is asymptotically the best possible for 1-MP.

Theorem 4.3: For a uniform set of d requests (with total size $K = d$), $\mathrm{cost}(\mathcal{D}_1) = \Theta(d^{\alpha}) + \Theta(nd)$.

Proof: The lower bound follows from Lemma 4.1. Now, to put upper bound on cost of the $\mathcal{D}_1$ routing scheme, we will consider two cases:

- ($d \ge n$) Then, using Lemma 4.2 and inequalities (4-5), we have:

$$\mathrm{cost}(\mathcal{D}_1)=\sum_{i=1}^{n+m-2}\sum_{e\in \mathrm{DE}_i}\mathcal{D}_1(e)^{\alpha}=$$

$$=\sum_{i=n}^{m-1}\sum_{e\in \mathrm{DE}_i}\mathcal{D}_1(e)^{\alpha}+2\sum_{i=1}^{n-1}\sum_{e\in \mathrm{DE}_i}\mathcal{D}_1(e)^{\alpha}\le$$

$$\le(m-n)\sum_{v\in \mathrm{DV}_n}\mathcal{D}_1(v)^{\alpha}+2\sum_{i=2}^{n}\sum_{v\in \mathrm{DV}_i}\mathcal{D}_1(v)^{\alpha}=$$

$$=(m-n)\,n\left(\left(\frac{d}{n}\right)^{\alpha}+O\left(\left(\frac{d}{n}\right)^{\alpha-2}\right)\right)+$$

$$+2\sum_{i=2}^{n} i\left(\left(\frac{d}{i}\right)^{\alpha}+O\left(\left(\frac{d}{i}\right)^{\alpha-2}\right)\right).$$

However, we have:

$$(m-n)\cdot n\cdot\left(\frac{d}{n}\right)^{\alpha}+2\sum_{i=2}^{n} i\left(\frac{d}{i}\right)^{\alpha}=$$

$$=(m-n)\frac{d^{\alpha}}{n^{\alpha-1}}+2\sum_{i=2}^{n}\frac{d^{\alpha}}{i^{\alpha-1}}\le$$

$$\le(m-n)\frac{d^{\alpha}}{n^{\alpha-1}}+2\int_1^n\frac{d^{\alpha}}{x^{\alpha-1}}dx=$$

$$=d^{\alpha}\cdot\left(\frac{m-n}{n}\frac{1}{n^{\alpha-2}}+\frac{2}{(\alpha-2)}(1-\frac{1}{n^{\alpha-2}})\right)=\Theta(d^{\alpha}).$$

It follows that:

$$\mathrm{cost}(\mathcal{D}_1)\le\Theta(d^{\alpha})+(m-n)\,n\cdot O\left(\left(\frac{d}{n}\right)^{\alpha-2}\right)+$$

$$+O\left(\sum_{i=2}^{n} i\left(\frac{d}{i}\right)^{\alpha-2}\right)=\Theta(d^{\alpha})+O\left(d^{\alpha-2}n^{4-\alpha}\right)+$$

$$+O\left(n\max\left(d^{\alpha-2},d^{\alpha-2}n^{3-\alpha}\right)\right)=\Theta(d^{\alpha})+$$

$$+O\left(d^{\alpha-2}n^{4-\alpha}\right)+O\left(d^{\alpha-2}n\right)+O\left(d^{\alpha-2}n^{4-\alpha}\right)=\Theta(d^{\alpha}),$$

where the last step holds since $1 \le n \le d$ and $\alpha > 2$.

- ($d < n$) In this case, we have:

$$\mathrm{cost}(\mathcal{D}_1)=\sum_{i=1}^{n+m-2}\sum_{e\in \mathrm{DE}_i}\mathcal{D}_1(e)^{\alpha}=$$

$$=\sum_{i=d}^{n+m-d+1}\sum_{e\in \mathrm{DE}_i}\mathcal{D}_1(e)^{\alpha}+2\sum_{i=1}^{d-1}\sum_{e\in \mathrm{DE}_i}\mathcal{D}_1(e)^{\alpha}\le$$

$$=(n+m-2d+2)\,d+2\sum_{i=2}^{d}\sum_{v\in DV_i}\mathcal{D}_1(v)^{\alpha}=$$

$$=O(nd)+2\sum_{i=2}^{d} i\left(\left(\frac{d}{i}\right)^{\alpha}+O\left(\left(\frac{d}{i}\right)^{\alpha-2}\right)\right)=$$

$$=O(nd)+\Theta(d^{\alpha})+O\left(d\max\left(d^{\alpha-2},d\right)\right)=$$

$$=O(nd)+\Theta(d^{\alpha}).$$

■

We remark that from Lemma 4.1 and the above Theorem it follows that for a uniform set of d requests (with $K = d$), $\text{cost}(\text{OPT}_1) = \text{cost}(\mathcal{F}_1) = \Theta(d^\alpha) + \Theta(nd)$.

B. k-MP *routing with uniform requests*

We now proceed to extend our results from the previous section to the case of k-MP uniform routing. We will consider sets of d requests of total size K, i.e., of size K/d each. A natural generalization of $\mathcal{D}_1$ routing, called $\mathcal{D}_k$ routing, is presented in Algorithm 5. Since in a $\mathcal{D}_k$ routing, we split the transmission of each request equally along its k paths, the cost of such a routing is the same as that of a $\mathcal{D}_1$ routing on the extended set of kd requests of size $\frac{K}{kd}$ each. Hence, the following result follows directly from Theorem 4.3 by a scaling argument.

Algorithm 5 $\mathcal{D}_k$ routing {for uniform k-MP}

Input: A set of d k-splittable requests, of size K/d each, in a $m \times n$ grid.
Solution: Split each of the requests into k smaller ones, each of size $\frac{K}{kd}$. Return the $\mathcal{D}_1$ routing of this new set of requests.

Theorem 4.4:

$$\text{cost}(\mathcal{D}_k) = \Theta(K^\alpha) + \Theta\left(K^\alpha \frac{n}{(kd)^{\alpha-1}}\right).$$

It turns out that although $\mathcal{D}_k$ performs only splits into paths of equal weight, one cannot achieve a better asymptotic result by using unequal splits.

Theorem 4.5: For any $R \in$ uniform k-MP :

$$\text{cost}(R) = \Omega\left(K^\alpha + K^\alpha \frac{n}{(kd)^{\alpha-1}}\right).$$

Proof: We have already established that for Max-MP, $\text{cost}(R)$ is in $\Omega(K^\alpha)$, which also holds for k-MP. For the second part of the lower bound, we use convexity of the cost function and the fact that on any diagonal, the stream of the routing of the d requests can be split into at most kd distinct paths:

$$\sum_{e \in \text{DE}} R(e)^\alpha \geq (kd)\left(\frac{K}{kd}\right)^\alpha$$

Thus:

$$\text{cost}(R) = \sum_{i=1}^{n+m-2} \sum_{e \in \text{DE}_i} R(e)^\alpha \geq$$

$$\geq (n+m-2)(kd)\left(\frac{K}{kd}\right)^\alpha = \Theta\left(K^\alpha \frac{n}{(kd)^{\alpha-1}}\right).$$

■

The two above theorems imply that:

$$\text{cost}(\text{OPT}_k) = \Theta(K^\alpha) + \Theta\left(K^\alpha \frac{n}{(kd)^{\alpha-1}}\right), \tag{8}$$

where OPT_k denotes the optimal cost solution to the considered (uniform) set of requests for k-MP.

Comparing this result with Theorem 3.2 for Max-MP routing, we obtain our main result: the threshold value of k necessary for achieving asymptotically optimal cost.

Theorem 4.6: A routing policy with a split limit of $k = \Theta(\frac{1}{d} \cdot n^{\frac{1}{\alpha-1}})$ is necessary and sufficient to achieve $\text{cost}(\text{OPT}_k) = \Theta(\text{cost}(\text{OPT}))$.

Proof: To prove sufficiency, we set $k = \Omega(\frac{1}{d} \cdot n^{\frac{1}{\alpha-1}})$ and by (8) we obtain:

$$\text{cost}(\text{OPT}_k) = \Theta(K^\alpha) + \Theta\left(K^\alpha \frac{n}{(kd)^{\alpha-1}}\right) =$$

$$= \Theta(K^\alpha) + \Theta\left(K^\alpha \frac{n}{\Omega(n)}\right) =$$

$$= \Theta(K^\alpha) + o(K^\alpha) = \Theta(K^\alpha).$$

To prove necessity, we observe that with $k = o(\frac{1}{d} \cdot n^{\frac{1}{\alpha-1}})$, the lower bound on cost becomes:

$$\text{cost}(\text{OPT}_k) = \Omega(K^\alpha) + \Omega\left(K^\alpha \frac{n}{o(n)}\right) = \omega(K^\alpha).$$

■

We end this subsection with a remark on the asymptotic behavior of the considered routing schemes for uniform instances, when $k \to +\infty$. By Theorem 4.4, we immediately have

Theorem 4.7: For fixed dimensions of the grid:

$$\lim_{k \to +\infty} \text{cost}(\mathcal{D}_k) = \text{cost}(\mathcal{C}).$$

Since, in general $\text{cost}(\mathcal{C}) > \text{cost}(\text{OPT})$, it is interesting to consider different routing schemes for k-MP with improved limit behavior. A natural candidate is $\mathcal{F}_k$ routing, obtained by a natural generalization of $\mathcal{F}_1$ routing, as shown in Algorithm 6.

Algorithm 6 $\mathcal{F}_k$ routing {for uniform k-MP}

Input: A set of d k-splittable requests, of size K/d each, in a $m \times n$ grid.
Solution:

1) Split each of the requests into k smaller ones, each of size $\frac{K}{kd}$.
2) Return the $\mathcal{F}_1$ routing of the new set of requests.

Theorem 4.8: For fixed dimensions of the grid:

$$\lim_{k \to +\infty} \text{cost}(\mathcal{F}_k) = \text{cost}(\text{OPT}).$$

Proof: Denote by OPT' an assignment of transfer sizes to edges of G obtained by rounding up the value of the transfer size in an OPT routing to the nearest integer multiple of $\frac{K}{kd}$, i.e., for all edges e, $\mathrm{OPT}'(e) = \frac{K}{kd}\lceil \frac{kd}{K}\mathrm{OPT}(e)\rceil$. By the properties of the min-cost flow used in the design of $\mathcal{F}_k$, the assignment OPT' majorizes the cost of the $\mathcal{F}_k$ routing for the considered instance. Thus, we have:

$$\mathrm{cost}(\mathcal{F}_k) \le \mathrm{cost}(\mathrm{OPT}') \le \mathrm{cost}(\mathrm{OPT}) + O\left(n \cdot \frac{K}{kd}\right).$$

Taking the limit $k \to +\infty$, the claim follows. ■

C. k-MP *routing for non-uniform requests*

We close our considerations with a discussion of the general (non-uniform) case, when no assumptions are made about the sizes of the routed request. We remark that the considered problem is computationally hard.

Theorem 4.9: The following decision version of non-uniform 1-MP routing is NP-complete: "Given $(n, m, K = (K_1, \dots, K_i), C, \alpha)$, decide if it is possible to perform 1-MP routing with cost $\le C$."

Proof: The proof proceeds by reduction from PARTITION PROBLEM [10]: "Given a set S of integers, decide if it is possible to partition S into subsets S_1 and S_2 such that $\sum S_1 = \sum S_2$."

For such an instance, we select the instance of 1-MP routing as follows: $n = 2$, $m = 2$, $K = S$ and $C = 4\left(\frac{1}{2}\sum_{s\in S} s\right)^{\alpha}$.

Observe that by identifying the sets of requests routed along the 2 different paths of the grid with S_1 and S_2, we have:

$$\mathrm{cost}(S_1, S_2) = 2\left(\sum_{s\in S_1} s\right)^{\alpha} + 2\left(\sum_{s\in S_2} s\right)^{\alpha}$$

Because of convexity of $f(x) = x^{\alpha}$, we obtain:

$$\mathrm{cost}(S_1, S_2) \ge 4\left(\frac{1}{2}\sum_{s\in S} s\right)^{\alpha} = C$$

and equality in the bound is achieved only if $\sum_{s\in S_1} s = \sum_{s\in S_2}$. ■

Despite the hardness of the studied problem, a careful modification of the $\mathcal{D}_k$ routing scheme, called an $\mathcal{A}_k$ routing (Algorithm 7) provides us with a good tool for routing non-uniform requests on the grid. (Note that applying $\mathcal{D}_k$ routing naively to a set of non-uniform requests could lead to excessively large additional cost.)

Theorem 4.10: $\mathcal{A}_k$ finds a solution to k-MP nonuniform routing whose cost is within a constant factor of the optimum.

Proof: It is enough to prove that for each edge diagonal DE, the cost induced by $\mathcal{A}_k$ on this diagonal is bounded by a constant with relation to the possible cost of the optimal solution $\mathrm{OPT}_k \in k$-MP on this diagonal.

We keep the notation $c_i = |S_i|$, and recall that DV denotes the vertex diagonal adjacent to DE. Also, let $p = |DV|$, and

---

Algorithm 7 $\mathcal{A}_k$ routing {for non-uniform k-MP}

Input: A set of d k-splittable requests, of given sizes $S = (s_1, s_2, \dots, s_d)$ (with $1 = s_1 \le s_2 \le \dots \le s_d$), in a $m \times n$ grid.
Solution:

1) Partition the set of request sizes into the union of disjoint subsets, $S = S_0 \cup S_1 \cup \dots$, such that $\forall_{s\in S_i}\ 2^i \le s < 2^{i+1}$.
2) For all non-empty sets S_i:
   - Find a $\mathcal{D}_k$ routing for the uniform instance consisting of $|S_i|$ requests of size 2^{i+1} each.
   - For all $1 \le j \le |S_i|$, route the j-th input request belonging to S_i using the paths assigned to the j-th request in the corresponding $\mathcal{D}_k$ routing.

---

let $\mathcal{D}[S_i]$ refer to the discrete $\mathcal{D}_1$ routing scheme applied inside S_i. We have:

$$\mathrm{cost}(\mathrm{DE}, \mathcal{A}_k) = \sum_{e\in \mathrm{DE}} \mathcal{A}_k(e)^{\alpha} \le \sum_{v\in \mathrm{DV}} \mathcal{A}_k(v)^{\alpha} \le$$

(taking into account that for $s \in S_i$ we have $s < 2 \cdot 2^i$)

$$\le \sum_{v\in \mathrm{DV}} \left(\sum_i \mathcal{D}[S_i](v) \cdot \frac{2\cdot 2^i}{k}\right)^{\alpha} =$$

$$= 2^{\alpha} \sum_{v\in \mathrm{DV}} \left(\sum_{i:\ c_i\cdot k \le p} \left(\mathcal{D}[c_i](v) \cdot \frac{2^i}{k}\right) + \right.$$

$$\left. + \sum_{i:\ c_i\cdot k > p} \left(\mathcal{D}[c_i](v) \cdot \frac{2^i}{k}\right)\right)^{\alpha} =$$

(taking into account that $\forall_{a,b\ge 0}\,(a+b)^{\alpha} \le 2^{\alpha-1}(a^{\alpha} + b^{\alpha})$)

$$\le 2^{2\alpha-1} \sum_{v\in \mathrm{DV}} \left(\left(\sum_{i:\ c_i\cdot k\le p} \mathcal{D}[c_i](v) \cdot \frac{2^i}{k}\right)^{\alpha} + \right.$$

$$\left. + \left(\sum_{i:\ c_i\cdot k > p} \mathcal{D}[c_i](v) \cdot \frac{2^i}{k}\right)^{\alpha}\right).$$

Observe that $\mathcal{D}[c_i](v) \le \lceil \frac{c_i \cdot k}{p} \rceil$.
Consequently, for $c_i \cdot k \le p$, we have $\mathcal{D}[c_i](v) \in \{0, 1\}$, and since $2^d + 2^{d-1} + \dots < 2 \cdot 2^d$:

$$\sum_{v\in \mathrm{DV}} \left(\sum_{i:\ c_i\cdot k\le p} \mathcal{D}[c_i](v) \cdot \frac{2^i}{k}\right)^{\alpha} \le$$

$$\le \sum_{v\in \mathrm{DV}} \left(2 \cdot \max_{i:\ c_i k \le p} \left\{\mathcal{D}[c_i](v) \cdot \frac{2^i}{k}\right\}\right)^{\alpha} \le$$

$$\le 2^{\alpha} \sum_{v\in \mathrm{DV}} \sum_{i:\ c_i\cdot k\le p} \left(\mathcal{D}[c_i](v) \cdot \frac{2^i}{k}\right)^{\alpha} = 2^{\alpha} \sum_{i:\ c_i\cdot k\le p} k \cdot c_i \left(\frac{2^i}{k}\right)^{\alpha}$$

For $c_i k > p$, we have $\mathcal{D}[c_i](v) \leq \lceil \frac{c_i \cdot k}{p} \rceil \leq 2\frac{c_i k}{p}$:

$$\sum_{v\in DV}\left(\sum_{i:\ c_i\cdot k>p}\mathcal{D}[c_i](v)\cdot\frac{2^i}{k}\right)^\alpha \leq \sum_{v\in DV}\left(\sum_{i:\ c_i k>p}2\cdot\frac{c_i\cdot k}{p}\cdot\frac{2^i}{k}\right)^\alpha =$$

$$= 2^\alpha\cdot\left(\frac{1}{p}\sum_{i:\ c_i k>p}c_i\cdot 2^i\right)^\alpha\cdot p$$

Adding the two sums, we can write in general:

$$\mathrm{cost}(DE,\mathcal{A}_k) \leq 2^{3\alpha-1}\cdot\left(\sum_{i:\ c_i k\leq p}c_i\left(\frac{2^i}{k}\right)^\alpha + p\cdot\left(\frac{1}{p}\sum_{i:\ c_i k>p}c_i\cdot 2^i\right)^\alpha\right)$$

Now we proceed to lower-bound the cost induced on diagonal DE by routing OPT_k. By $\mathrm{OPT}_k[S_i]$, we will denote OPT_k restricted to requests from S_i, and by $\mathrm{OPT}_k[s]$ we will understand OPT_k restricted to request s. We have:

$$\mathrm{cost}(DE,\mathrm{OPT}_k) = \sum_{e\in DE}\mathrm{OPT}_k(e)^\alpha \geq \frac{1}{2^{\alpha-1}}\sum_{v\in DV}\mathrm{OPT}_k(v)^\alpha =$$

$$= \frac{1}{2^{\alpha-1}}\sum_{v\in DV}\left(\sum_{i:\ c_i k\leq p}\mathrm{OPT}_k[S_i](v) + \sum_{i:\ c_i k>p}\mathrm{OPT}_k[S_i](v)\right)^\alpha \geq$$

$$\geq \frac{1}{2^{\alpha-1}}\sum_{v\in DV}\left(\left(\sum_{i:\ c_i k\leq p}\mathrm{OPT}_k[S_i](v)\right)^\alpha + \left(\sum_{i:\ c_i k>p}\mathrm{OPT}_k[S_i](v)\right)^\alpha\right)$$

We put bounds on both parts of sum:

$$\sum_{v\in DV}\left(\sum_{i:\ c_i k\leq p}\mathrm{OPT}_k[S_i](v)\right)^\alpha = \sum_{v\in DV}\left(\sum_{i:\ c_i k\leq p}\sum_{s\in S_i}\mathrm{OPT}_k[s](v)\right)^\alpha \geq$$

$$\geq \sum_{i:\ c_i k\leq p}\sum_{s\in S_i}\sum_{v\in DV}\mathrm{OPT}_k[s](v)^\alpha \geq$$

(observing that $\sum_{v\in DV}\mathrm{OPT}_k[s](v) = s$, and s can be split into at most k parts)

$$\geq \sum_{i:\ c_i k\leq p}\sum_{s\in S_i}k\cdot\left(\frac{s}{k}\right)^\alpha \geq \sum_{i:\ c_i k\leq p}c_i k\left(\frac{2^i}{k}\right)^\alpha$$

In the second part, we obtain:

$$\sum_{v\in DV}\left(\sum_{i:\ c_i k>p}\mathrm{OPT}_k[S_i](v)\right)^\alpha \geq$$

$$\geq p\cdot\left(\frac{1}{p}\sum_{v\in DV}\sum_{i:\ c_i k>p}\mathrm{OPT}_k[S_i](v)\right)^\alpha =$$

$$= p\cdot\left(\frac{1}{p}\sum_{i:\ c_i k>p}\sum_{s\in S_i}s\right)^\alpha \geq p\cdot\left(\frac{1}{p}\sum_{i:\ c_i k>p}c_i\cdot 2^i\right)^\alpha$$

Merging both results, we have:

$$\mathrm{cost}(DE,\mathrm{OPT}_k) \geq$$

$$\geq \frac{1}{2^{\alpha-1}}\left(\sum_{i:\ c_i k\leq p}c_i k\left(\frac{2^i}{k}\right)^\alpha + p\cdot\left(\frac{1}{p}\sum_{i:\ c_i k>p}c_i\cdot 2^i\right)^\alpha\right).$$

Combining the lower-bound on the cost of OPT_k and the upper bound on the cost of $\mathcal{A}_k$, we finally have:

$$\mathrm{cost}(DE,\mathcal{A}_k) \leq 2^{4\alpha-2}\mathrm{cost}(DE,\mathrm{OPT}_k),$$

which proves that $\mathcal{A}_k$ is a $(2^{4\alpha-2})$-approximation algorithm for non-uniform k-MP. ■

We end this section with similar threshold theorems as for the uniform case, obtaining bounds on value of k necessary and enough for achieving cost $= \Theta(K^\alpha)$. However, in this case the lower and upper bounds do not match, since the lower bound depends on the structure of the set of requests.

Theorem 4.11: For $k = o\left(\frac{1}{d}\cdot n^{\frac{1}{\alpha-1}}\right)$:

$$\mathrm{cost}(\mathrm{OPT}_k) = \omega(\mathrm{cost}(\mathrm{OPT})).$$

Proof: The optimal-cost routing for a set of d requests of total size K, each of which can be split into k paths, cannot be better than the optimal-cost routing on a single request of size K, which can be split into kd parts. The latter routing problem belongs to uniform $(k\cdot d)$-MP and so, taking into account formula (8), we can write:

$$\mathrm{cost}(\mathrm{OPT}_k) = \Omega\left(K^\alpha + K^\alpha\frac{n}{(kd)^{\alpha-1}}\right) = \Omega\left(K^\alpha + K^\alpha\frac{n}{o(n)}\right) =$$

$$= \omega(K^\alpha) = \omega(\mathrm{cost}(\mathrm{OPT}))$$

■

Theorem 4.12: For $k = \Omega\left(n^{\frac{1}{\alpha-1}}\right)$:

$$\mathrm{cost}(\mathrm{OPT}_k) = \Theta(\mathrm{cost}(\mathrm{OPT})).$$

Proof: Only the upper bound needs to be shown. Observe that the cost of an optimal k-MP routing cannot decrease if we replace a pair of requests of size s_1, s_2 by a single request of size $s_1 + s_2$. By iterating the argument, we can upper-bound the value of OPT_k for a given instance of d requests of total size K by the value of OPT_k for an instance consisting of a single request of size K. Once again, the claim follows by an application of formula (8). ■

V. Experimental results

In this section we discuss the results of experimental evaluation of the algorithms presented in the previous section. We analyze the effect of n, k and α on the efficiency of solutions found for k-MP routing of instances with uniform (identical-size) requests. Throughout the section, we choose the number of requests as $d = 1$ (for uniform instances, other values result only in a scaling factor for k in k-MP, and do not affect Max-MP).

We focus on the approximation ratio, looking at the cost of the routing obtained using the two schemes designed for uniform MPk ($\mathcal{D}_k$, $\mathcal{F}_k$), relative to the cost of the optimal solution OPT to Max-MP, which is treated as the reference solution. In some graphs, we also provide the cost of the sub-optimal Max-MP routing $\mathcal{C}$ as an additional reference.

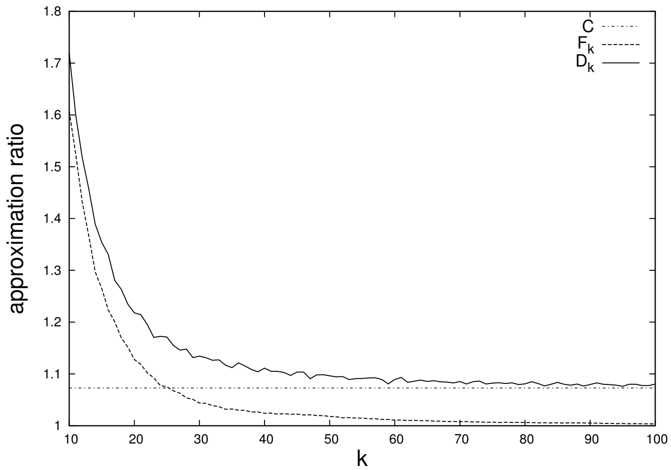

Figure 5. Effect of k on the cost of $\mathcal{F}_k$ and $\mathcal{D}_k$ routing for a 30×30 grid

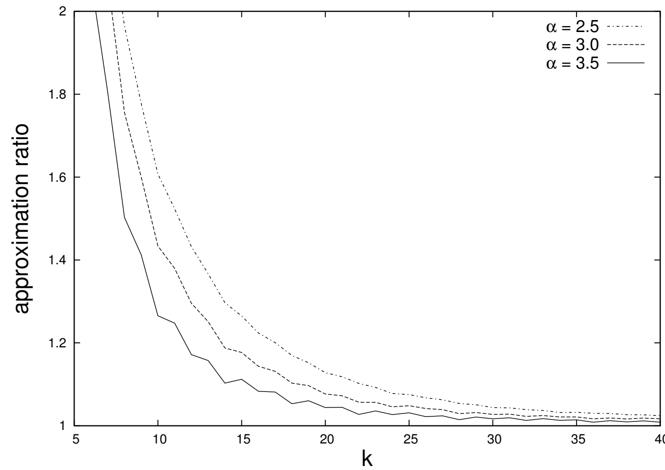

Figure 7. Effect of α on $\text{cost}(\mathcal{F}_k)/\text{cost}(\text{OPT})$, for a 30×30 grid

We recall that the value of cost of the optimal solution to Max-MP, $\text{cost}(\text{OPT}_k)$, is bounded from below by $\text{cost}(\text{OPT})$, and from above by both $\text{cost}(\mathcal{D}_k)$ and $\text{cost}(\mathcal{F}_k)$.

The implementation and tests were performed within a software package, written by the authors for this purpose in GNU C++. The min-cost flow subroutines were implemented using the standard cycle-canceling method [12] with some optimizations for faster performance. The presented results of the tests are deterministic and fully reproducible, independent of the test environment and the details of the implementation of the flow algorithms.

A. Impact of k on the routing cost

We begin by studying the approximation ratio of algorithms $\mathcal{F}_k$ and $\mathcal{D}_k$ for increasing values of k, the allowed number of splits of each requests. In the first plot (Fig. 5), we fix the dimensions of the grid $n, m = 30$, model power cost exponent $\alpha = 2.5$, plotting the values of $\text{cost}(\mathcal{F}_k)/\text{cost}(\text{OPT})$ and $\text{cost}(\mathcal{D}_k)/\text{cost}(\text{OPT})$ for k in the range $k \in [10, 100]$. For reference, we also provide the approximation ratio of $\mathcal{C}$ routing for the studied instance.

We observe that, as predicted by theory (Theorems 4.7 and 4.8), $\lim_{k\to+\infty} \text{cost}(\mathcal{F}_k) = \text{cost}(\text{OPT})$ and $\lim_{k\to+\infty} \text{cost}(\mathcal{D}_k) = \text{cost}(\mathcal{C})$, and the respective costs converge to their limits quickly, reaching a point 10% over the respective limit already for $k < n$. In general, the convergence need not be monotone for either of the approximation algorithms, since partitioning a request into $k+1$ equally-weighted paths may give worse results than partitioning it into k equally-weighted algorithms.

In our second plot (Fig. 6), we present more compelling evidence of the relation $k = \Theta(n^{1/(\alpha-1)})$ for the threshold split value resulting in asymptotically-optimal cost, derived theoretically as Theorem 4.6. Once again, we choose model parameter $\alpha = 2.5$. In the experiment, we consider square grids of increasing size in the range $n = m \in [10, 120]$, testing three different relations between n and k ($k = \lfloor 2n^{1/2} \rfloor$, $k = \lfloor \frac{3}{2} n^{2/3} \rfloor$, $k = n$). For each of these relations, we plot the approximation ratios $\text{cost}(\mathcal{F}_k)/\text{cost}(\text{OPT})$ and $\text{cost}(\mathcal{D}_k)/\text{cost}(\text{OPT})$. Based on the plot, we can presume that:

- For the relation $k \sim n^{1/2}$, we have in the limit: $\text{cost}(\mathcal{F}_k)/\text{cost}(\text{OPT}) \to +\infty$, $\text{cost}(\mathcal{D}_k)/\text{cost}(\text{OPT}) \to +\infty$.
- For the relation $k \sim n^{2/3}$, we have in the limit: $\text{cost}(\mathcal{F}_k)/\text{cost}(\text{OPT}) \to \mathit{const}$, $\text{cost}(\mathcal{D}_k)/\text{cost}(\text{OPT}) \to \mathit{const}$.
- For the relation $k \sim n$, we have in the limit: $\text{cost}(\mathcal{F}_k)/\text{cost}(\text{OPT}) \to 1$, $\text{cost}(\mathcal{D}_k)/\text{cost}(\text{OPT}) \to \mathit{const}$.

We remark that the relation $k \sim n^{2/3}$ precisely corresponds to the threshold exponent $1/(\alpha-1) = 2/3$ for the considered value of α. Thus, the limit behavior of all the algorithms is consistent with the theory derived in the previous section. We note that, whereas the cost achieved by both $\mathcal{F}_k$ routing and $\mathcal{D}_k$ routing is highly satisfactory, the performance of $\mathcal{F}_k$ routing proves to be superior to $\mathcal{D}_k$ in all of the performed tests. (This observation does not directly result from the theory.)

B. Effect of power exponent α

In our final plot, we show the effect of the power exponent α (which is a constant of the model) on the required threshold value of split parameter k. Specifically, in Fig. 7, we show the rate of convergence of the approximation ratio $\text{cost}(\mathcal{F}_k)/\text{cost}(\text{OPT})$ to 1 in a grid of dimensions $n = m = 30$ for three different values of the power exponent, $\alpha \in \{2.5, 3, 3.5\}$. We can observe that the convergence is faster for larger values of α. This is consistent with the theoretical threshold, $k = \Theta(n^{1/(\alpha-1)})$, whose growth rate decreases with the increase of α.

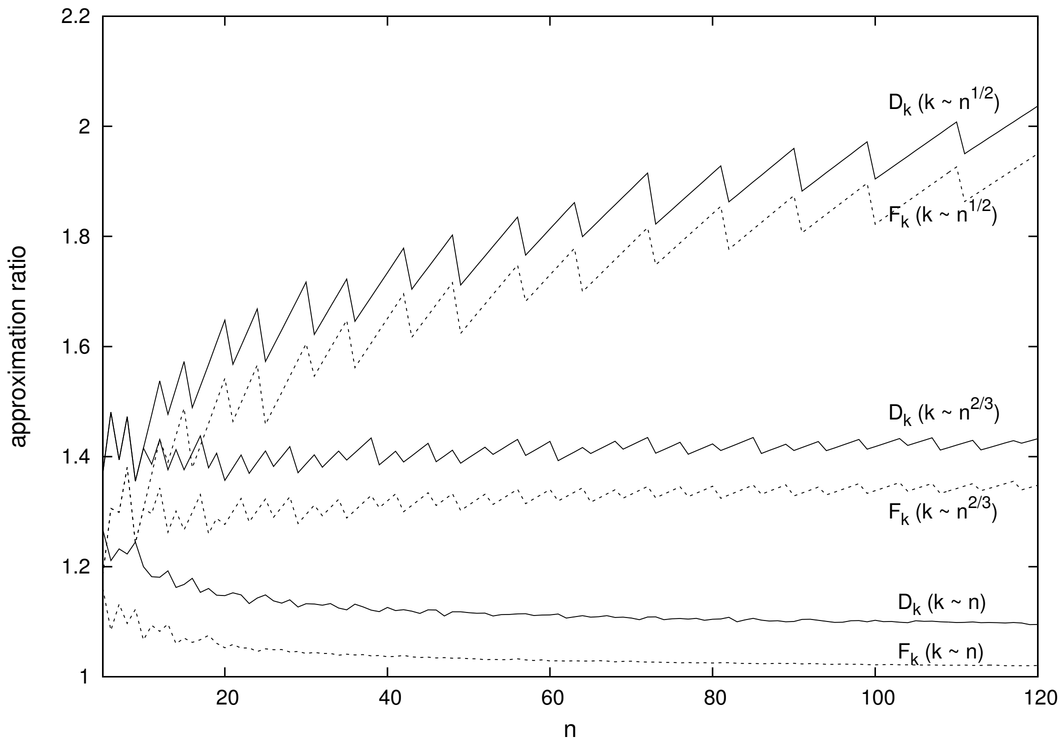

Figure 6. Approximation ratio for $\mathcal{F}_k$ and $\mathcal{D}_k$ routing with split parameter $k \sim n^{\beta}$, for β greater, equal and smaller than $1/(\alpha-1)$. $[\alpha = 2.5]$

VI. Conclusions

The contribution of our study is twofold. On the one hand, we advance the theory of splitting of requests in Manhattan routing on the grid, and point out that in practice, only a relatively small number of splits per request will be beneficial from a power-cost perspective. On the other hand, we propose efficient approximation schemes for such a k-path routing problem. Experimental evidence corroborates the theoretical results, showing that the designed algorithms lead to routings with a cost which is, in practice, highly superior to that resulting from our theoretical bounds.

In future work, it would be beneficial to improve the constant bounds on the approximation ratios of the algorithms, establishing more tightly their dependence on the power exponent α. Another promising direction of study would extend our results to routing requests between multiple sources and targets on the grid.

References

[1] R. K. Ahuja, D. S. Hochbaum, and J. B. Orlin. Solving the convex cost integer dual network flow problem. In G. Cornuéjols, R. E. Burkard, and G. J. Woeginger, editors, *IPCO*, volume 1610 of *Lecture Notes in Computer Science*, pages 31–44. Springer, 1999.

[2] A. Andrei, M. T. Schmitz, P. Eles, Z. Peng, and B. M. Al-Hashimi. Simultaneous communication and processor voltage scaling for dynamic and leakage energy reduction in time-constrained systems. In *International Conference on Computer-Aided Design*, pages 361–367. IEEE, 2004. Event Dates: Nov 2005.

[3] H. Aydin and Q. Yang. Energy-aware partitioning for multiprocessor real-time systems. In *IPDPS*, page 113. IEEE Computer Society, 2003.

[4] G. Baier, E. Köhler, and M. Skutella. The k-splittable flow problem. *Algorithmica*, 42(3-4):231–248, 2005.

[5] A. Benoit, R. G. Melhem, P. Renaud-Goud, and Y. Robert. Power-aware manhattan routing on chip multiprocessors. In *IPDPS*, pages 189–200. IEEE Computer Society, 2012.

[6] A. Benoit, P. Renaud-Goud, and Y. Robert. On the performance of greedy algorithms for power consumption minimization. In G. R. Gao and Y.-C. Tseng, editors, *ICPP*, pages 454–463. IEEE, 2011.

[7] J.-J. Chen and T.-W. Kuo. Multiprocessor energy-efficient scheduling for real-time tasks with different power characteristics. In *ICPP*, pages 13–20. IEEE Computer Society, 2005.

[8] R. D. G. Blake and T. Mudge. A survey of multicore processors. *Signal Processing Magazine*, page 26(6):26 –37, 2009.

[9] T. Ishihara and H. Yasuura. Voltage scheduling problem for dynamically variable voltage processors. In *Low Power Electronics and Design, 1998. Proceedings. 1998 International Symposium on*, pages 197–202, Aug. 1998.

[10] N. Karmarker and R. M. Karp. The differencing method of set partitioning. Technical Report UCB/CSD-83-113, EECS Department, University of California, Berkeley, 1983.

[11] A. V. Karzanov and S. T. McCormick. Polynomial methods for separable convex optimization in unimodular linear spaces with applications. *SIAM J. Comput.*, 26(4):1245–1275, 1997.

[12] M. Klein. A primal method for minimal cost flows. *Management Sciences*, 14:205–220, 1967.

[13] P. Langen and B. Juurlink. Leakage-aware multiprocessor scheduling. *Journal of Signal Processing Systems*, 57:73–88, 2009.

[14] S. E. Lee and N. Bagherzadeh. A variable frequency link for a power-aware network-on-chip (noc). *Integration*, 42(4):479–485, 2009.

[15] R. Mishra, N. Rastogi, D. Zhu, D. Mossé, and R. Melhem. Energy aware scheduling for distributed real-time systems. In *In International Parallel and Distributed Processing Symposium*, page 21, 2003.

[16] J. D. Owens, W. J. Dally, R. Ho, D. J. Jayasimha, S. W. Keckler, and L.-S. Peh. Research challenges for on-chip interconnection networks. *IEEE Micro*, 27:96–108, 2007.

[17] K. Pruhs, R. van Stee, and P. Uthaisombut. Speed scaling of tasks with precedence constraints. In T. Erlebach and G. Persiano, editors, *WAOA*, volume 3879 of *Lecture Notes in Computer Science*, pages 307–319. Springer, 2005.

[18] D. Seo, A. Ali, W.-T. Lim, N. Rafique, and M. Thottethodi. Near-optimal worst-case throughput routing for two-dimensional mesh networks. In *Proceedings of the 32nd annual international symposium on Computer Architecture*, ISCA '05, pages 432–443, Washington, DC, USA, 2005. IEEE Computer Society.

[19] L. Shang, L.-S. Peh, and N. K. Jha. Dynamic voltage scaling with links for power optimization of interconnection networks. In *HPCA*, pages 91–102. IEEE Computer Society, 2003.