Approximations for the Disjoint Paths Problem in High-Diameter Planar Networks

Jon Kleinberg* **Éva Tardos†**

Abstract

We consider the problem of connecting distinguished terminal pairs in a graph via edge-disjoint paths. This is a classical NP-complete problem for which no general approximation techniques are known; it has recently been brought into focus in papers discussing applications to admission control in high-speed networks and to routing in all-optical networks. In this paper we provide $O(\log n)$-approximation algorithms for two natural optimization versions of this problem for the class of nearly-Eulerian, uniformly high-diameter planar graphs, which includes two-dimensional meshes and other common planar interconnection networks. We give an $O(\log n)$-approximation to the maximum number of terminal pairs that can be simultaneously connected via edge-disjoint paths, and an $O(\log n)$-approximation to the minimum number of wavelengths needed to route a collection of terminal pairs in the "optical routing" model considered by Raghavan and Upfal, and others. The latter result improves on an $O(\log^2 n)$-approximation for the special case of the mesh obtained independently by Aumann and Rabani. For both problems the $O(\log n)$-approximation is a consequence of an $O(1)$-approximation for the special case when all terminal pairs are roughly the same distance apart. Our algorithms make use of a number of new techniques, including the construction of a "crossbar" structure in any nearly-Eulerian planar graph, and develops some connections with classical matroid algorithms.

1 Introduction

A basic problem that arises in large-scale communication networks is that of assigning paths to *connection requests.* Each connection request is a pair of physically separated nodes that wish to communicate over a path through the network; given a list of such requests, one wants to assign paths to as many as possible in such a way that no two paths "interfere" with each other. Thus, for a given list of requests, we can ask a number of natural questions. How many requests are simultaneously *realizable* using paths that are mutually edge-disjoint? How many *rounds of communication* are required to satisfy all requests, when all paths assigned in a single round must be edge-disjoint? These turn out to be classical NP-complete problems; previously

*E-mail: kleinber@theory.lcs.mit.edu. Laboratory for Computer Science, MIT, Cambridge MA 02139 USA. Supported by an ONR Graduate Fellowship and in part by the NSF PYI award of Éva Tardos. Part of this work was performed while visiting the School of ORIE at Cornell.

†E-mail: eva@cs.cornell.edu. School of ORIE, Cornell University, Ithaca NY 14853. Supported in part by a Packard Fellowship and an NSF PYI award.

known approximation techniques for these problems are limited either to very special graphs, or to "high-bandwidth" models in which a large number of paths can share a single edge.

The intractability of the disjoint paths problem does not appear to be simply a theoretical phenomenon. Awerbuch, Gawlick, et. al. [4] observe that much of the difficulty in establishing virtual circuits in large-scale communication networks comes from the lack of good heuristics for finding disjoint paths. In practice, admission control and routing for virtual circuits are typically performed using greedy algorithms, which perform badly on a number of very common interconnection patterns.

In this paper we develop approximation algorithms for both of the problems mentioned above for a fairly general class of planar graphs, illustrating along the way some algorithmic tools that appear to be of use in understanding the disjoint paths problem in other cases as well. In a subsequent paper [14] we give an $O(1)$-approximation algorithm and an on-line $O(\log n)$-approximation algorithm for the maximum disjoint paths problem in a related class of graphs. The improved algorithm relies on many of the techniques developed in this paper.

We can make these problems precise as follows. Given a graph $G = (V, E)$, each connection request is specified by a pair of terminals s_i and t_i, where $s_i, t_i \in V$. Let $\mathcal{T}$ be the set of all terminal pairs $(s_1, t_1), \ldots, (s_k, t_k)$. We say that $\mathcal{T}$ is *realizable* in G if there exist mutually edge-disjoint paths $P_1, \ldots, P_k$ such that P_i has endpoints s_i and t_i. Given G, k, and $\mathcal{T}$, determining whether $\mathcal{T}$ is realizable in G is one of Karp's original NP-complete problems [13]; it remains NP-complete even when the underlying graph G is the two-dimensional mesh [15].

In this paper we consider two natural optimization versions of this problem. In the *maximum disjoint paths problem* one wishes to find a maximum size subset of $\mathcal{T}$ that is realizable in G. The second of the questions raised in the opening paragraph can be described as the *minimum path coloring problem*; here we must assign to every terminal pair a path P_i and a "color" c_i so that no two paths that share an edge are assigned the same color. More succinctly, we are trying to find the fewest number of realizable subsets into which the set $\mathcal{T}$ of terminal pairs can be partitioned. Let us denote this minimum by $\chi(\mathcal{T})$. If $\mathcal{T}^*$ is a realizable set of maximum cardinality, then a c-approximation algorithm is one that always produces a realizable set of size at least $|\mathcal{T}^*|/c$. Analogously, a c-approximation for the path coloring problem partitions the set $\mathcal{T}$ into at most $c \cdot \chi(\mathcal{T})$ realizable subsets. All the approximation algorithms we discuss run in polynomial time.

Much of the previous work on this problem has considered the case in which each path consumes only a small fraction of the available bandwidth on an edge; this can be modeled by requiring $\Omega(\log n)$ parallel copies of each edge. Within this framework, the randomized rounding technique of Raghavan and Thompson [20, 19] gives good approximations for both problems. On-line algorithms for the maximum disjoint paths problem for this case are given in Awerbuch et al. [3].

In many applications, however, each communication path consumes a large fraction of the available bandwidth on a link; thus it makes sense to consider approximation algorithms for graphs without a large number of parallel edges. Previous approximation algorithms in this setting have only been developed for restricted types of graphs. For the maximum disjoint paths problem Garg, Vazirani, and Yannakakis [12] give a 2-approximation for trees with parallel edges (the maximization problem is NP-complete, though deciding realizability is easy). For the special case of the two-dimensional mesh, Awerbuch, Gawlick, Leighton, and Rabani [4] give an on-line $O(\log n \log\log n)$-approximation algorithm, and independently of our work Aumann and Rabani [2] give an (off-line) $O(\log n)$-approximation algorithm. Approximation algorithms for $\chi(\mathcal{T})$ have been considered in a number of previous papers concerned with routing in all-optical networks: Raghavan and Upfal [21] give a 3/2-approximation algorithm when the underlying graph is a tree, and a 2-approximation when the underlying graph is a cycle. Aumann and Rabani [2] give an $O(\log^2 n)$-approximation when the underlying graph is the two-dimensional mesh.

Here we consider the class of *nearly-Eulerian, uniformly high-diameter planar graphs* (defined precisely below), which includes most common high-diameter planar interconnection networks such as the mesh and the hex. Indeed, in Section 2.1 we show that all "geometrically well-formed" graphs are uniformly high-diameter. The *nearly-Eulerian* condition requires that all nodes not on the outer face have even degree. This is not particularly limiting in practice, as it can be achieved simply by doubling every edge; it roughly corresponds to assuming that a communication path does not consume more than half of the bandwidth on a link. We note that such evenness conditions have proved crucial in much previous work on exactly solvable special cases of this problem (see the survey by Frank [11]).

Our first main result is an $O(\log n)$-approximation algorithm for the maximum disjoint paths problem, in nearly-Eulerian, uniformly high-diameter planar graphs. The second of our main results is an $O(\log n)$-approximation to $\chi(\mathcal{T})$ for the same class of graphs. When specialized to the two-dimensional mesh our algorithm is similar to that of Aumann and Rabani, but improves their bound by a factor of $\log n$. See Section 2.2 for more on the comparison of the algorithms. For both problems the $O(\log n)$-approximation is a consequence of an $O(1)$-approximation for the special case when all terminal pairs are roughly the same distance apart.

We feel that developing approximation algorithms for general classes of graphs such as this is important for a number of reasons. First of all, networks arising in the context of virtual circuit routing problems do not tend to have a structure as regular as that of the mesh, and so it is desirable to look for approaches that make as few assumptions as possible about the nature of the underlying graph. But perhaps more importantly, the disjoint paths problem is an area in which little work has been done on approximation algorithms, despite a great deal of early work devoted to exactly solvable special cases. Essentially no general approximation methods are known, even at a heuristic level. We feel that a contribution of this work is the development of a number of new techniques that appear to be interesting in their own right, as they provide tools for constructing disjoint paths in more general settings. These techniques are surveyed in Section 2.2.

Finally, a different line of work related to the construction of disjoint paths can be found in papers of Broder, Frieze, Peleg, Suen, and Upfal [5, 6, 18]. Here the underlying graph G is assumed to have strong expansion properties; in this case one can prove that any set of terminal pairs of at most a given size must be realizable in G. The goal then is to find the paths in (randomized) polynomial time. In this paper we deal only with planar graphs, which of course are not expanders.

In presenting the results of this paper we first focus on the maximum disjoint path problem, and then show how to adapt our algorithm to the optical routing problem. In Section 2 we define the class of graphs considered, and give a rough outline of the algorithms and the main techniques used. In Section 3 we discuss the two main technical tools that we need: the Okamura-Seymour theorem, and a theorem of Davies and McDiarmid regarding strongly base-orderable matroids. Section 4 presents the approximation algorithm for the disjoint paths problem, and in Section 5 we show how to adapt the approximation algorithm in Section 4 for the optical routing problem.

2 Outline

In this section we define the class of graphs considered in this paper and give an outline of our algorithm.

2.1 The class of graphs

Let the distance $d(u,v)$ between vertices u and v in G be the fewest number of edges in a u-v path; let $B_d(v)$ denote the set of all u with $d(v,u) \le d$. By a *plane graph*, we mean a planar graph with a specified plane embedding. We say that a simple plane graph is *uniformly high-diameter* with parameters α, β and L (sometimes written (α, β, L)-UHD) if

(i) for all v and d, $|B_d(v)| \le \beta d^2$, and

(ii) for all v, v', and d, $|B_{2d}(v)|/|B_d(v')| \le \alpha$, and

(iii) all internal faces have size at most L.

Expressed more simply, we require that (i) d-step neighborhoods of a vertex have at most quadratic size, and (ii) any two neighborhoods of about the same radius should have about the same size. Condition (iii) is a necessary technical assumption, since one can take any low-diameter graph and subdivide its edges to produce a high-diameter graph whose disjoint-paths structure is identical. Note that since G is a simple graph, the diameter condition implies that G has *maximum degree* $\Delta \le \beta - 1$, as $\Delta + 1 = \max_v |B_1(v)| \le \beta$. We can also handle graphs in this class with parallel edges. When each edge can appear with bounded multiplicity, our algorithm can be implemented without modification and still achieves an $O(\log n)$-approximation ratio.

We need a further assumption that the degree of every internal vertex (i.e. those not on the outer face) is even. This is a version of the *parity condition* that appears in a large number of previous results on exactly solvable special cases of the disjoint paths problem [11]. We will call a plane graph for which the degree of every internal vertex is even *nearly-Eulerian*.

It is easy to verify that the two-dimensional mesh and hex are nearly-Eulerian UHD plane graphs. In fact, UHD plane graphs encompass a fairly broad range of graphs, as the following geometric construction shows. Fix a simple polygon P, and a positive constant $\rho \ge 1$. We say that a graph G is *geometrically well-formed* with parameters ρ and P if G has a plane embedding in which the boundary of the outer face of G is equal to P, and for some positive constant r, every internal face contains a disc of radius r and is contained in a disc of radius ρr (so every internal face has geometrically about the same size). If we "fuse" the edges incident to degree-2 vertices of G (such vertices are irrelevant for the disjoint paths problem), it is not difficult to verify that for fixed ρ and P, there are constants α, β and L, so that every (ρ, P)-geometrically well-formed graph is an (α, β, L)-UHD plane graph. While this geometric class of graphs was the primary motivation for our definition of UHD graphs, it is easy to construct families of UHD plane graphs that are not geometrically well-formed.

2.2 Overview of the Algorithm

In Sections 4 and 5, we describe the disjoint paths algorithm and the optical routing algorithm in detail; here we give an overview of the main techniques used. The $O(\log n)$-approximation will be a consequence of obtaining a constant-factor approximation for the special case in which all terminal pairs are about the same distance apart: there is a d so that for every ℓ, $d(s_\ell, t_\ell)$ is between $d/2$ and d. Once we are in this special case, we fix a sufficiently small constant $\gamma < 1/2$ and construct overlapping "clusters" $B_{\gamma d}(v_i)$ for vertices v_i that we choose greedily until every vertex of G is contained in some cluster.

A single subproblem is now associated with a single pair (C_1, C_2) of clusters: we wish to route terminal pairs with one end (say s_ℓ) in C_1 and the other end (t_ℓ) in C_2. Using the uniformly high diameter property we can grow the clusters so that it is possible to solve a constant fraction of these subproblems without any two of them "interfering"; thus it is enough to obtain a constant-factor approximation for the subproblem associated with a fixed pair (C_1, C_2).

Our approximation algorithm for such a subproblem is based loosely the following three-step process: first we route a subset of the terminals $s_\ell \in C_1$ to the boundary of C_1, then we route the corresponding subset of the terminals $t_\ell \in C_2$ to the boundary of C_2, and finally we connect the corresponding endpoints of these paths via edge-disjoint paths from the boundary of C_1 to the boundary of C_2. There are a number of problems in getting such an approach to work: there is the "coordination" problem of choosing a large set of terminal pairs so that both s_ℓ *and* t_ℓ can be routed to their respective cluster boundaries; and we have to ensure that the pairs routed to the respective boundaries can be connected with each other. At this high level, the above outline is analogous to the outline of the algorithm of Aumann and Rabani [2] for the special case of the mesh. For our general class of graphs, a number of additional difficulties arise, for which we develop the following techniques.

First consider the problem of routing paths from the boundary of C_1 to the boundary of C_2. We define a *crossbar* to be a collection of edge-disjoint paths from C_1 to C_2, such that each pair of paths in the collection meets at some vertex; we call the endpoints of these paths the *crossbar ports.* A crossbar has the very useful property that given any bijection from the crossbar ports on C_1 to the crossbar ports on C_2, the pairs of ports can be partitioned into three groups each of which can be routed using edge-disjoint paths. Notice that the $n \times 2n$ mesh has a crossbar of size $n-1$ connecting the two smaller sides. Aumann and Rabani [2] use this natural mesh-crossbar for routing. We use a theorem of Okamura and Seymour [17] to prove that in any nearly-Eulerian planar graph, if C_1 and C_2 are sufficiently far apart, then there is crossbar from C_1 to C_2 of size at least half the maximum number of edge-disjoint C_1-C_2 paths.

Now consider the "coordination problem" of making sure that s_ℓ is routed to the boundary if and only if its partner t_ℓ is routed. For the maximum disjoint paths problem Aumann and Rabani [2] use a maximum flow computation to select a maximum size subset of the terminals that can be simultaneously routed to the boundary of C_1 and C_2. For the coloring problem (optical routing) they use the greedy set-covering approach of repeatedly selecting large realizable sets; this costs an extra $O(\log n)$ factor in the quality of approximation. We handle the coordination problem by exploiting a certain matroidal structure of disjoint paths and thereby save the $\log n$ cost of the greedy set-cover algorithm.

Our algorithms will route the selected set of terminals in C_1 and C_2 to the crossbar ports, and use the crossbar to connect up the paths. This brings into focus a final difficulty that arises in general UHD plane graphs: since not all vertices on the boundary of C_1 and C_2 will generally be crossbar ports, we have to argue that the optimal routing cannot gain a lot by using paths that cross the boundary of C_1 and C_2 at vertices other than our particular crossbar ports. (Note that in the mesh, one can take almost the entire boundary to be the set of crossbar ports, and this problem does not arise.) To make sure that our solution is close to optimal, we show how to modify each cluster C_i using a procedure we call *ε-linking.* A cluster with an ε-linked boundary has the property that for any subsets S and S' of the boundary of C_i with $|S| = |S'|$, if a subset U of the terminals s_ℓ can be routed to S then a constant fraction of these terminals can also be routed to S'. Thus, the placement of crossbar ports on the boundary of an ε-linked cluster will not affect the quality of approximation by more than a constant factor.

In a subsequent paper [14] we give an $O(1)$-approximation algorithm, and an on-line $O(\log n)$-approximation algorithm, for the maximum disjoint paths problem in the related class of *densely embedded, nearly-Eulerian* graphs. Densely embedded graphs are defined by a property stronger the UHD property used in this paper, but the stronger property is only required to hold locally on $O(\log n)$-size neighborhoods. The improved algorithm relies on two of the techniques developed in this paper: crossbars, and clusters with ε-linked boundaries.

3 Preliminaries

In this section we discuss two known results from combinatorial optimization that we will use for our algorithms. The first is a theorem of Okamura and Seymour [17] concerning an exactly solvable special case of the disjoint paths problem; the second gives us some useful tools from matroid theory.

3.1 An exactly solvable special case

A large amount of work has been done on identifying special cases of the disjoint paths problem that are solvable in polynomial time. Much of this previous work has dealt with the case in which the underlying graph G is planar and satisfies a certain crucial *parity condition*: if we form an *augmented* graph by adding to G an edge from s_i to t_i for each terminal pair, then the parity condition requires that the augmented graph be Eulerian. It is interesting that very little is known about the existence of polynomial-time algorithms in cases not satisfying this parity condition; some variants become NP-complete when the condition is lifted [16]. In a relatively early paper, Okamura and Seymour [17] gave a polynomial-time algorithm for the case in which G is planar and satisfies the parity condition, and all terminals lie on a single face of G. The algorithm is obtained along with a proof that the following *cut condition* is sufficient for realizability: one cannot remove j edges and separate more than j terminal pairs in G. A linear time algorithm for this problem has been recently obtained by Wagner and Weihe [22]. We will use an extension of the Okamura–Seymour theorem due to Frank [10] to build the "crossbar" mentioned above. For other results on polynomially solvable special cases of the realizability problem see the survey by Frank [11].

3.2 Matroidal tools

The connection between paths and matroids that we use here stems from the following construction (see e.g. [23] for definitions and basic facts related to matroids). Let G be a graph, and S and T two subsets of the vertices. Call a set $S' \subset S$ *independent* if there are edge-disjoint paths that connect the vertices in S' to different vertices in T. Then it is not difficult to show (see e.g. [23]) that these sets form the independent sets of a matroid $M_G(S,T)$ over the ground set S; matroids arising by this construction are called *gammoids*.

We make use of Edmonds' matroid intersection theorem [9]. For a matroid M, let $\rho(U)$ denote the rank of a subset U of the ground set; this is the size of the largest independent set contained in U. Let M_1 and M_2 be two matroids over the same ground set S with ρ_1 and ρ_2 the corresponding rank functions; the matroid intersection theorem implies that if for every $S' \subset S$ we have $\min(\rho_1(S'), \rho_2(S')) \geq \frac{1}{k}|S'|$, then S contains a set of cardinality at least $\frac{1}{k}|S|$ which is independent in both matroids.

For the optical routing problem we need a stronger version of the above fact: instead of finding one large set that is independent in both M_1 and M_2, we need to cover the ground set by a minimum number of common independent sets. Gammoids are *strongly base-orderable* matroids (see e.g. [23]); this means that if B_1 and B_2 are two bases of a gammoid, then there is a bijection $\psi : B_1 \rightarrow B_2$ such that for any $X \subset B_1$, the sets $(B_1 - X) \cup \psi(X)$ and $(B_2 - \psi(X)) \cup X$ are both bases. Davies and McDiarmid [7] proved that if M_1 and M_2 are two strongly base-orderable matroids over the same ground set S, then S can be covered by k sets each of which is independent in both M_1 and M_2 if and only if S can be covered by k independent sets of M_1 and by k independent sets of M_2. Moreover, such a covering can be found in polynomial time. Combined with Edmonds' matroid covering theorem [8] the Davies-McDiarmid theorem implies that if for every $S' \subset S$ we have $\min(\rho_1(S'), \rho_2(S')) \geq \frac{1}{k}|S'|$, then S can be covered by k sets each of which is independent in both M_1 and M_2.

4 The Disjoint Paths Approximation

Fix $d \geq 2$, and let $\mathcal{T}^d$ denote the set of all terminal pairs (s_i, t_i) for which $d/2 \leq d(s_i, t_i) \leq d$. The maximum disjoint paths problem with input G and $\mathcal{T}^d$ will be called the subproblem associated with distance d. In this section we obtain a constant-factor approximation for such subproblems. This will give an $O(\log n)$-approximation for the original problem, by first solving the $O(\log n)$ subproblems associated with $2, 4, 8, \ldots, 2^{\lceil \log \mathrm{diam}(G) \rceil}$ and then only routing pairs in the subproblem in which we find the largest realizable subset.

Some additional notation will be useful: if $S \subset V$, then $G[S]$ denotes the subgraph of G induced by the vertices of S; $\delta(S)$ denotes the set of edges with exactly one endpoint in S; $\pi(S)$ denotes the set of vertices of S incident to an edge in $\delta(S)$; and $S^o = S - \pi(S)$. Observe that removing $\pi(S)$ from S disconnects it from the rest of the graph, and $\pi(B_d(v))$ consists of vertices at exactly distance d from v.

4.1 Decomposing the Graph

The first step of the algorithm will be to decompose the graph into a collection of *clusters* $\{C_i\}$; we will then solve a subproblem for each pair of clusters (C_i, C_j).

The decomposition is done by the following straightforward procedure. Fix a small constant $\gamma > 0$. In the interest of making the presentation easier to follow, we have not tried to optimize the constants here; let us use $\gamma = (75L\beta^2)^{-1}$. We pick a vertex $v_i \in V$, add $C_i = B_{\gamma d}(v_i)$ to our collection of clusters, and label the vertices in C_i "covered." We then iterate, choosing an uncovered vertex and growing another cluster, until no vertices remain uncovered. Note that the clusters can overlap, but that the "roots" v_i of the clusters are at least γd apart from one another. Thus we have the following lemma.

Lemma 4.1 *The sets $B_{\frac{1}{2}\gamma d}(v_i)$ are pairwise disjoint, where $\{v_i\}$ is the set of roots of the clusters.*

4.2 ε-Linking the Boundaries

Now consider a pair of clusters (C_i, C_j); let $\mathcal{T}_{ij}$ denote the collection of all terminal pairs in $\mathcal{T}^d$ with one terminal in C_i and the other in C_j. We work toward obtaining a constant-factor approximation to the maximum realizable subset of $\mathcal{T}_{ij}$. First we refine the clusters $\{C_i\}$ by growing around each an "augmented cluster" C_i' with ε-linked boundary components. Appropriately linked boundaries will guarantee that disjoint paths can be routed to the crossbar ports essentially as easily as to any other vertices on the boundary of the cluster.

Before we can precisely define ε-linkage we need some additional notation. Let G be a graph; as is standard, if $U, W \subseteq V(G)$ we define a U-W flow to be a collection of edge-disjoint paths, each of which has one endpoint in U and the other in W. Let $f(U, W)$ denote the maximum value of a U-W flow. A related notion is that of a *simple* U-W flow, which we define to be a U-W flow in which the endpoints of the flow paths are all distinct. Let $f_s(U, W)$ denote the maximum value of a simple U-W flow. (Both $f(U, W)$ and $f_s(U, W)$ can be computed by a single-source, single-sink maximum flow computation: we compute a maximum flow from an additional vertex u' to an additional vertex w', such that graph edges have capacity 1, and u' (resp. w') is connected to each vertex in U (resp. W) via an edge of infinite capacity in case of $f(U, W)$ and capacity 1 in case of $f_s(U, W)$.) We say that U is *ε-linked* to W if for every $U' \subseteq U$,

$$f_s(U', W) \geq \varepsilon |U'|.$$

Abusing terminology somewhat, let us say that *a single set $S \subseteq V$* is *ε-linked* if it satisfies the following guarantee: if U and W are subsets of S with $|U| \leq |W|$, then $f_s(U, W) \geq \varepsilon |U|$.

The next lemma shows that ε-linkage is a useful notion, because it allows us to re-route flow from one part of G to another while preserving a constant fraction of it. We will use this property in arguing that by routing the paths through the crossbar ports we do not lose more than a constant fraction of the flow.

Lemma 4.2 *Let $U, W \subseteq V(G)$ with U ε-linked to W, and $S \subseteq V(G)$. Then $f_s(S, W) \geq \frac{\varepsilon}{1+\varepsilon} f_s(S, U)$.*

Proof. We use the max-flow min-cut theorem stating that the max-flow value $f_s(S, X)$ for a set X equals to the value of the minimum S-X cut. In this setting, an arbitrary S-X cut can be obtained by deleting some vertices in S and X, and some edges of G. We will show that for any S-W cut, there is an S-U cut of at most $1 + \frac{1}{\varepsilon}$ times its value; this implies the statement of the lemma.

Let Y be an S-W cut, and let U' be the set of vertices of U that are reachable from S after the removal of Y. We construct an S-U cut Y' by simply adding U' to Y.

By our assumption, there is a collection of edge-disjoint paths $\mathcal{P}$ of cardinality at least $\varepsilon|U'|$ from U' into W. Y must contain at least one edge (or the endpoint in W) from each member of $\mathcal{P}$, for otherwise it would not separate S from W — if Y were to miss $P \in \mathcal{P}$ with endpoint $u' \in U'$, then we could construct an S-W path missing Y by concatenating an S-u' path avoiding Y with the path P.

Thus $|Y| \geq |\mathcal{P}| \geq \varepsilon|U'|$ and hence

$$|Y'| = |Y| + |U'| \leq (1 + \frac{1}{\varepsilon})|Y|. \quad \blacksquare$$

Before we give our procedure to ε-link the boundary of a cluster C_i we need a little more notation. Let S denote a set of vertices. If C is a connected subset of $G - S$, then C clearly belongs to a single connected component of $G - S$; we use $\Gamma(S, C)$ to denote this component. The set of vertices in $\pi(S)$ which have a neighbor in $\Gamma(S, C)$ will be called the *segment of* $\pi(S)$ *bordering* C and denoted $\sigma(S, C)$. Consider a segment $\sigma(S, C)$. By contracting all of $\Gamma(S, C)$ to a vertex we see that the vertices in $\sigma(S, C)$ all belong to a single face of $G[S]$; thus the vertices of $\sigma(S, C)$ inherit the natural (cyclic) ordering $\preceq$ around this face. Call a subset of $\sigma(S, C)$ *contiguous* if it forms an interval with respect to $\preceq$.

The following lemma provides a procedure to "slice off" parts of a cluster boundary that are not ε-linked.

Lemma 4.3 *Assume that all internal faces of $G[S]$ that are at most $L\varepsilon|\sigma(S, C)|$ away from any node on the outer face have size at most L. If $\sigma(S, C)$ is not ε-linked then there exists a contiguous subset $T \subset \sigma(S, C)$ and a path P' in S, such that $|P'| \leq L\varepsilon|T|$ and the removal of P' from S disconnects T from the rest of $\sigma(S, C)$. Furthermore, if $\sigma(S, C)$ is not ε-linked, then such a set T and path P' can be found in polynomial time.*

Proof. Observe that in order for $\sigma(S, C)$ to be ε-linked, it is enough to require that for $U, W \subseteq \sigma(S, C)$, with U and W non-overlapping in $\preceq$ and $|U| \leq |W|$, one has $f_s(U, W) \geq 2\varepsilon|U|$. For then if $U, W \subseteq \sigma(S, C)$ are overlapping sets, we can first obtain non-overlapping subsets U'

and W' of half the respective sizes of U and W. (Start from somewhere in $\sigma(S, C)$ and walk around clockwise until you've seen half of one of U or W; say U. Let the vertices of U seen so far be U' and the vertices of W not seen so far be W'.)

Assume there are non-overlapping subsets U and W of $\sigma(S, C)$ with $|U| \leq |W|$ and $f_s(U, W) < 2\varepsilon|U|$; hence there is a U-W cut Y of value less than $2\varepsilon|U|$. Recall that such a cut consists of vertices of U and W, as well as edges of $G[S]$. If we let $U' = U - Y$ and $W' = W - Y$, and $Y' = Y - (U \cup W)$, then U' and W' are separated by the edge cut Y', which has size less than $2\varepsilon|U'|$. Let F be the set of all internal faces of $G[S]$ touched by edges in Y'. First we argue that all these faces must have size at most L. The set of faces F and edges Y' form a path in the planar dual graph with endpoints equal to the outer face. By contradiction assume that one of these faces has size bigger than L, and consider the part of this path from the outer face to the first such large face. By going around the faces in this path the "short way" (each face has size at most L), we get a path in $G[V(F)]$ of length at most $\frac{1}{2}L|F| \leq L\varepsilon|U'| \leq L\varepsilon|\sigma(S, C)|$ from the outer face to the large face, contradicting the assumption.

Consider again the path in the planar dual graph through the faces F and edges Y'. By going around the faces in this path the "short way", we get a path P in $G[V(F)]$ of length at most $|P| \leq \frac{1}{2}L|F| \leq L\varepsilon|U'|$ with endpoints on the outer face of $G[S]$ whose removal disconnects U' from W'. Furthermore, there is some subpath P' of P with only its endpoints on the outer face of $G[S]$, which disconnects some set $T \subset \sigma(S, C_j)$ from the rest of $\sigma(S, C_j)$ and satisfies $|P'| \leq L\varepsilon|T|$ as required by the lemma. To find such a set T and path P' in polynomial time we have to find the minimum cut value for every contiguous subset T of $\sigma(S, C)$. ∎

We are working on the subproblem defined by the pair (C_i, C_j). We want C_i and C_j to have the property that $\sigma(C_i, C_j)$ and $\sigma(C_j, C_i)$ are ε-linked for some constant $\varepsilon > 0$. There is no reason why the clusters should have this property as constructed; thus we grow "augmented clusters" C_i' and C_j' which are not much larger and do have this property. Again making no attempt at optimizing constants we use $\varepsilon = (50\beta L)^{-1}$.

Theorem 4.4 *There exists a set C_i' such that*

(i) $C_i \subset C_i' \subset B_{6\gamma d}(v_i)$,

(ii) the segment $\sigma(C_i', C_j)$ is ε-linked, and

(iii) $|\sigma(C_i', C_j)| \leq 25\beta\gamma d$.

Proof. To construct C_i', we begin by continuing the radial growth process by which C_i was created. We choose a distance s between $2\gamma d$ and $3\gamma d$ such that $|\pi(B_s(v_i))| \leq 25\beta\gamma d$; this is possible since otherwise we would have $|B_{3\gamma d}(v_i)| > (25\beta\gamma d)(\gamma d) > \beta(3\gamma d)^2$.

Consider the plane subgraph $G[B_s(v_i)]$ of G. By contracting all of $\Gamma(B_s(v_i), C_j)$ to a single node we see that the segment $\sigma(B_s(v_i), C_j)$ belongs to a single face of $G[B_s(v_i)]$. Other faces of $G[B_s(v_i)]$ correspond to other segments of $\pi(B_s(v_i))$, or to faces of G.

Assume for now that expect for the face containing $\sigma(B_s(v_i), C_j)$ all other faces of $G[B_s(v_i)]$ are of lize at most L. In this case we now have a set $S = B_s(v_i)$ which will ultimately contain C_i'; the remainder of the process only decreases S and the size of $\pi(S)$ by pulling $\pi(S)$ back towards C_i. This will establish part (iii) and the second half of part (i).

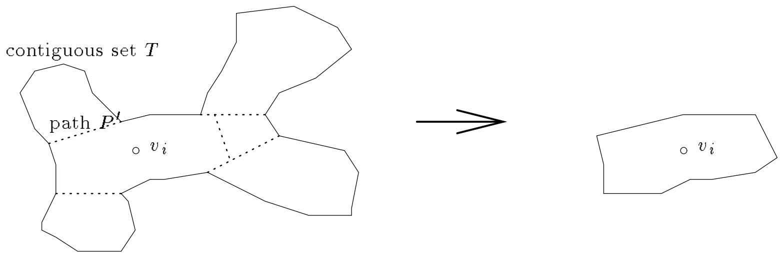

Figure 1: Linking the boundary

We produce the final cluster by using Lemma 4.3 for $G[S]$ with the face containing $\sigma(S, C_j)$ as the outside face, to iteratively slice off parts of the boundary of S that are not ε-linked. The difficult part of the proof is to show the first part of (i), i.e., to guarantee that the slicing off process terminates before any of the cluster disappears.

We apply Lemma 4.3 repeatedly; if we find a set T and path P' such that $|P'| \leq L\varepsilon|T|$, and removing the vertices of P' from S disconnects T from the rest of $\pi(S)$, then we delete the T-side of this cut from S. The "updated" cluster has P' as part of its boundary. We then iterate on the new cluster, finding a contiguous set T as in Lemma 4.3 and slicing it off; Lemma 4.3 implies that if this slicing off process terminates then the boundary S is ε-linked.

We need to show that all the vertices on the new boundary of S are "close" to the boundary of $B_s(v_i)$ (at most γd away), and therefore "far" from C_i. This will imply the first part of (i). In the first iteration of the slicing off process this follows from the definition, since the new boundary is connected to the boundary of $B_s(v_i)$ by the path P' of length at most $L\varepsilon|T|$, and $|T| \leq |\sigma(S, C_j)| \leq 25\beta\gamma d$. We now show by induction on the number of iterations of slicing off that all vertices on the boundary of the final cluster S are at most γd away from the boundary of $B_s(v_i)$, and hence that $S \supseteq C_i$.

We divide the slicing off process into *phases*. As long as portions of the original boundary $\pi(B_s(v_i))$ remain on the outer face of S, we will say that we are in the *first phase*; other phases will be defined later. At any given point in the first phase, $\pi(S)$ will consist of alternating intervals $P_0, \sigma_0, \ldots, P_r, \sigma_r$, where $\sigma_\ell \subset \pi(B_s(v_i))$ and the interval $\sigma'_{\ell+1}$ of $\pi(B_s(v_i))$ lying between σ_ℓ and $\sigma_{\ell+1}$ (with indices understood mod r) has been sliced off by the new vertices $P_{\ell+1}$. See Figure 1, which shows a cluster after five iterations of slicing off.

Let $h = L\varepsilon$. In each iteration, some subpath of the boundary is being replaced by a new path that is $h < 1$ times as long. By induction on the number of iterations in the first phase, one can thus verify that $|P_\ell| \leq h|\sigma'_\ell|$ for all ℓ.

This finishes the proof in case the iterations come to an end before the end of the first phase. Otherwise, consider the iteration on which the first phase comes to an end. In this iteration, the path P has both endpoints on P_1, which then covers all the remaining boundary $\pi(B_s(v_i))$. Again it is not hard to show that $|P_1| \leq h|\pi(B_s(v_i))|$; moreover, any vertex on the new boundary P_1 can reach the old boundary by a path of length at most $h|\pi(B_s(v_i))|$.

Each subsequent phase now proceeds exactly like the previous phase, except that it begins with an initial set whose boundary is at most h times the length of the boundary at the start of the previous phase. Thus the phases will terminate with a set C_i' each of whose boundary vertices can reach the original boundary by a path of length at most

$$\begin{aligned} |\pi(B_s(v_i))| \cdot \sum_{\ell=1}^{\infty} h^\ell &< 2h|\pi(B_s(v_i))| \\ &\leq \frac{1}{25\beta} \cdot 25\beta\gamma d \\ &= \gamma d. \end{aligned}$$

Hence the final set C_i' contains C_i. This finishes the proof in the special case when all internal faces of $B_s(v_i)$ have size at most L.

Notice, the outer face of $G[S]$ might contain some part of the outer face of G. It is important that in this case the length of the paths used for slicing off, and the progress of the iterations, are measured by $|\sigma(S, C_j)|$; *i.e.*, we do not count parts of S that lie on the outer face of G.

Now consider the case in which $G[B_s(v_i)]$ has large faces other than the one containing $\sigma(B_s(v_i), C_j)$. Again we want to apply Lemma 4.3 to $G[B_s(v_i)]$ with the face containing $\sigma(B_s(v_i), C_j)$ as the outer face. The idea is that we modify $B_s(v_i)$ to get a set S so that initially $\sigma(S, C_j)$ is at a distance of at least γd from any vertex belonging to any other large face of $G[S]$, and we show that the hypotheses of Lemma 4.3 will be satisfied throughout the slicing off process. First consider the case in which the outer face of G is one of these extra large faces in $G[B_s(v_i)]$, and the distance from $\sigma(B_s(v_i), C_j)$ to the outer face is less than γd. In this case, we continue growing $B_s(v_i)$ radially (i.e., increasing s) until $\sigma(B_s(v_i), C_j)$ is at least γd away from the outer face, but $s \leq 5\gamma d$ and $\sigma(B_s(v_i), C_j)$ has size small enough to satisfy (iii). This is possible since $s = 4\gamma d$ is large enough to guarantee that the outside face of G is far from $\sigma(B_s(v_i), C_j)$; and for some s between $4\gamma d$ and $5\gamma d$ the set $\sigma(B_s(v_i), C_j)$ must satisfy (iii), as otherwise we would have $|B_{5\gamma d}(v_i)| > (25\beta\gamma d)(\gamma d) = \beta(5\gamma d)^2$. If $G[B_s(v_i)]$ has large faces close to $\sigma(B_s(v_i), C_j)$ defined by other segments of $\pi(B_s(v_i))$, then we grow $B_s(v_i)$ a distance of γd into every component of $G - B_s(v_i)$ except for $\Gamma(B_s(v_i), C_j)$, and let S be the resulting set.

We apply a slicing off process to $\sigma(S, C_j)$ analogous to the case described above, to get the extended cluster C_i' whose boundary segment $\sigma(S, C_j)$ is ε-linked. We show by induction on the number of slicing off steps that Lemma 4.3 continues to apply. The size of the outer face of $G[S]$ throughout phase ℓ is bounded by $h^{\ell-1}(25\beta\gamma d)$. We claim that $\sigma(S, C_j)$ is distance of at least $\varepsilon L h^{\ell-1}(25\beta\gamma d) = h^{\ell-1}\gamma d/2$ from any other large face of $G[S]$ throughout phase ℓ. During phase ℓ any node in $\sigma(S, C_j)$ is at distance at most

$$|\sigma(B_s(v_i), C_j)| \cdot \sum_{k=1}^{\ell} h^k \leq (25\beta\gamma d) \cdot \sum_{k=1}^{\ell} h^k$$

from $\sigma((B_s(v_i), C_j)$. Hence $\sigma(S, C_j)$ is at this point a distance of at least $\gamma d - (25\beta\gamma d) \cdot \sum_{k=1}^{\ell} h^k \geq h^{\ell-1}\gamma d/2$ from any vertex belonging to any other large face of $G[S]$. This proves that Lemma 4.3 applies throughout the process. Now the theorem follows by the same argument as used in the case above, when all faces of $G[S]$ were small. ■

4.3 Building a Crossbar

We now prove that every nearly-Eulerian high-diameter plane graph is rich in crossbar structures. If $G = (V, E)$ is a graph, and U and W are disjoint subsets of V, then a *U-W crossbar* is a set of edge disjoint paths from U to W (i.e., a not necessarily simple U-W flow) for which each pair of paths meets in at least one vertex. The *value* of a crossbar is the number of paths it contains. A *simple crossbar* is a crossbar defined by a simple flow.

The following lemma shows why crossbars are useful.

Lemma 4.5 *Let $\mathcal{F}$ be a simple U-W crossbar of value $k = |U|$ and $\psi : \{1, \ldots, k\} \to \{1, \ldots, k\}$ a bijection. Then the collection of terminal pairs $\{(u_i, w_{\psi(i)})\}$ can be partitioned into three sets, $\mathcal{T}_1$, $\mathcal{T}_2$, and $\mathcal{T}_3$, each of which can be routed via edge-disjoint paths.*

Proof. We route the pair $(u_i, w_{\psi(i)})$ on a path P_i constructed as follows: begin by following the u_i-w_i flow path, and switch to the $u_{\psi(i)}$-$w_{\psi(i)}$ flow path at their first intersection. Now observe that the path P_i only shares edges with $P_{\psi(i)}$ and $P_{\psi^{-1}(i)}$; we can therefore 3-color the collection of paths so that no two paths of the same color share an edge. ■

The two-dimensional mesh is the canonical example of a graph with crossbars; we now show that crossbars are present in any plane graph with small internal face size and even internal degree. Let $G = (V, E)$ be a nearly-Eulerian plane graph with internal face size at most L. Let A, B be two connected subsets of V, let $A' = \pi(A)$, $B' = \pi(B)$, $G' = G[V - A^o - B^o]$, and $f = f_{G'}(A', B')$.

Theorem 4.6 *Let A, B, A', B', G' and f defined above. Assume that G is a nearly-Eulerian plane graph and has maximum internal face size at most L. Then if $d(A', B') > L(f + 3)$, there is an A'-B' crossbar in G' of value at least $\frac{1}{2}f$.*

Proof. We begin with an A'-B' flow in G' of value f; suppose that it consists of edge-disjoint paths $P_1, \ldots, P_f$, with P_i joining $u_i \in A'$ and $w_i \in B'$ (these endpoints need not be all distinct). If P_i and P_j cross at some vertex, we can form new u_i-w_j and u_j-w_i paths with one fewer crossing; in this way we eventually obtain a collection of paths which do not cross. Relabeling, we can assume again that P_i is a u_i-w_i path.

Assume for now that the relative position of A and B is as depicted in Figure 2 — all internal faces of both A and B are internal faces of G. If we contract A^o and B^o in G to single vertices, the neighbors of these vertices will be A' and B' respectively. Thus the vertices of A' (resp. B') all lie on a common face φ_1 (resp. φ_2) of G'. Since the P_i do not cross, the cyclic order of the vertices $\{u_i\}$ is the reverse of the cyclic order of the $\{w_i\}$.

Let φ_0 denote the outer face. Since the paths P_i do not cross, we can find a polygonal curve $\mathcal{C}_1$ in the plane which meets G' only at vertices, such that its endpoints lie on φ_0 and φ_1, and no P_i crosses $\mathcal{C}_1$. Let Q_1 be the set of vertices of G that meet $\mathcal{C}_1$. We "cut open" the graph G' along $\mathcal{C}_1$, splitting the vertices in Q_1. We repeat this process for the faces φ_0 and φ_2, using a curve $\mathcal{C}_2$ meeting vertices Q_2, and thereby obtain a graph G'_0 whose outer face φ'_0 consists of

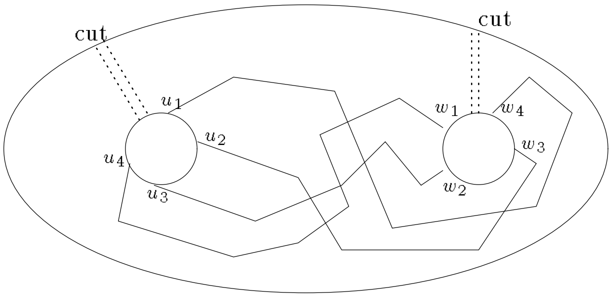

Figure 2: The crossbar construction

$\varphi_0 \cup \varphi_1 \cup \varphi_2$, as well as the vertices on faces crossed by the curves $\mathcal{C}$ and $\mathcal{C}'$, with the vertices in Q_1 and Q_2 appearing twice. A crucial point is that there is still an A'-B' flow of value f in G_0'. Say the order of $A' \cup B'$ on the outer face of G_0' is $u_1, \ldots, u_f, w_f, \ldots, w_1$.

To find the A'-B' crossbar, we set up an edge-disjoint paths problem in G_0' with terminal pairs $\mathcal{T}'$ consisting of (u_i, w_{f+1-i}) for $i = 1, \ldots, f$. If we are able to route $f' \leq f$ pairs, then since each pair of paths in the routing must meet at some vertex, this provides a crossbar of value f'. Thus, the claim below will imply the theorem.

Finally we have to consider the case when either A or B have internal faces that are not internal faces of G (either the outer face, or faces left by other components of $G - A$ or $G - B$). We apply the above argument to the connected component of $G - (A^o \cup B^o)$ containing the boundaries $\sigma(A, B)$ and $\sigma(B, A)$. If the outer face of G does not touch this component, then consider a plane drawing of this graph with any face, other than the hole left by A^o or B^o, as the outer face. ∎

Claim 4.7 *There are edge-disjoint paths in G_0' connecting at least $\frac{1}{2}f$ of the terminal pairs in $\mathcal{T}'$.*

Proof. For $S \subset V(G_0')$, let $D(S)$ denote the set of terminals in S whose corresponding source or sink does not also belong to S. Recall that the *cut condition*, which is clearly necessary for the existence of the edge-disjoint paths, requires that $|\delta(S)| \geq |D(S)|$ for all S. Let us verify that the cut condition holds in the present case, for the entire terminal set $\mathcal{T}'$.

By a standard argument, it is enough to consider sets S with S and $V - S$ both connected, and for which $D(S) \neq \phi$. First note that if there were some set S with $|D(S)| > |\delta(S)|$ for which $D(S) \subseteq A'$ (or equivalently $D(S) \subseteq B'$), then this would contradict the fact that there is an A'-B' flow of value f. Now suppose there is a connected set S for which $D(S)$ meets both A' and B'. Contracting $V - S$ to a single vertex, we see that the vertices of $\sigma(S, V - S)$ all lie on the outer face φ_0'' of $G_0'[S]$. Let P denote the simple path $V[\varphi_0''] - V[\varphi_0']$ on the outer face of $G_0'[S]$ (i.e. P consists of those vertices newly added to the outer face). Note that $\sigma(S, V - S) \subseteq P$, and since the endpoints of P share faces with members of A' and B' respectively, we have

$|P| \geq d(A', B') - 2L > L(f+1)$. If we walk along P from one endpoint to the other, we must encounter a vertex of $\sigma(S, V - S)$ at least every L steps; otherwise there would be an internal face of G'_0 and hence of G containing too many vertices. But each such boundary vertex we encounter adds at least 1 to $|\delta(S)|$; since $|P| > L(f+1)$, we have $|\delta(S)| > f+1 \geq |D(S)|$. Note this proves $|\delta(S)| > |D(S)|$ for every set S such that $D(S)$ meets both A' and B'.

Observe that G'_0 is nearly-Eulerian, and all terminals lie on its outer face. In such a situation, an extension of the Okamura–Seymour theorem due to Frank [10] says that the following *strict cut condition* is sufficient for realizability: $|\delta(S)| > |D(S)|$ for all $S \neq \phi$. We have already verified the (non-strict) cut condition, and hence consider this as follows. Call a set S *tight* if $|\delta(S)| = |D(S)|$; our goal is to remove fewer than half the terminal pairs so that there will be no non-empty tight sets.

We noted above that no set S for which $D(S)$ meets both A' and B' can be tight. So each non-empty tight set S in G'_0 can be labeled either "A'-tight" or "B'-tight," depending on whether $D(S)$ meets A' or B'. Let S and S' be two arbitrary A'-tight sets. Then using the submodularity of δ and the fact that there can be no demand between S and S', it is easy to show that $S \cap S'$ and $S \cup S'$ are both tight as well. (Note that $S \cap S'$ may be empty.) This implies the following: the inclusionwise-minimal A'-tight sets are all disjoint. The analogous statement holds for the B'-tight sets. Moreover, we may assume that all minimal tight sets contain at least two terminals; if a tight set S contains only u_i then $|\delta(S)| = 1$, and so we can "slide" u_i across the single edge in $\delta(S)$, producing an equivalent problem with fewer nodes.

Let $S_1, \ldots, S_p$ denote the minimal A'-tight sets, and $T_1, \ldots, T_q$ denote the minimal B'-tight sets. Consider the bipartite graph whose nodes correspond to these sets, with two nodes joined by an edge if there is a pair of terminals with ends in the corresponding two sets. We need to prove that we can select at least half of the edges of this bipartite graph without selecting all of the edges incident to any node. Selecting the maximum number of such edges is a maximum degree-constrained subgraph problem, and is equivalent to a matching problem on a larger bipartite graph. Using the fact that the maximum size of a matching in a bipartite graph is equal to the minimum size of a vertex cover, we get that the maximum number of edges we can select is equal to the minimum sum of the degree bounds on a set of vertices that covers all edges. Since all vertex degrees are at least two, this latter sum is at least half of the edges.

The set of terminal pairs selected by the bipartite matching algorithm defines a routing problem with no minimal tight sets, and hence with no non-empty tight sets at all. Thus we now have a set of at least $\frac{1}{2}f$ terminal pairs for which the strict cut condition holds; by Frank's Theorem, there exist edge-disjoint paths connecting these pairs. ■

In order to use Theorem 4.6 in our approximation algorithm we need an A'-B' crossbar that does not use edges far away from A' and B'. Let G'' be the graph induced only on vertices of $V - A^o - B^o$ that are within a distance $2d(A', B')$ of $[A' \cup B']$.

Corollary 4.8 *Let A, B, A', B', G'' and f defined above. Assume that G is a nearly-Eulerian plane graph and has maximum internal face size at most L. Then if $d(A', B') > L(f+3)$, there is an A'-B' crossbar in G'' of value at least $\frac{1}{2}f$.*

Proof. Since $d(A', B') > L(f+3)$, and the internal face size of G'' is at most L, an argument

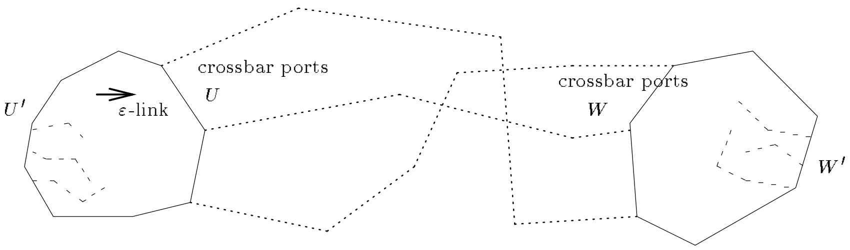

Figure 3: Approximating a subproblem

analogous to one used to verify the cut condition in the above Claim shows that any A'-B' cut in G'' which is not also an A'-B' cut in G' must have value at least f; and hence $f = f_{G'}(A', B') = f_{G''}(A', B')$. Now Theorem 4.6 can be applied in the graph $G'' \cup (A^o \cup B^o)$, yielding the corollary. ■

Furthermore, we are interested in constructing simple crossbars.

Corollary 4.9 *Let A, B, A', B', G'' and f be as defined above. If $d(A, B) > L(f + 3)$, there is a simple A'-B' crossbar in G'' of value at least $f/2\Delta$.*

Proof. Viewing the terminal pairs in the crossbar given by Corollary 4.8 as the edges of a bipartite graph, we can color the crossbar paths with Δ colors so that no two paths that share an endpoint have the same color. The largest color class gives the desired simple crossbar. ■

4.4 Routing the Terminal Pairs

We now use the techniques developed in the previous sections to obtain a constant-factor approximation for the subproblem associated with clusters C_i and C_j, and subsequently for the subproblem associated with distance d. As before, let $\mathcal{T}_{ij}$ denote the set of terminal pairs with one end in C_i and the other in C_j; for each such pair (s_ℓ, t_ℓ), say that $s_\ell \in C_i$ and $t_\ell \in C_j$. Let $\mathcal{T}_i$ denote the set of $s_\ell \in C_i$ and $\mathcal{T}_j$ the set of $t_\ell \in C_j$. Our approximation algorithm for a fixed cluster pair (C_i, C_j) is as follows.

(1) Around the clusters C_i and C_j we build augmented clusters C_i' and C_j' as given by the proof of Theorem 4.4. Let $\sigma_i = \sigma(C_i', C_j)$ and $\sigma_j = \sigma(C_j', C_i)$.

(2) We use Corollary 4.9 to build a simple σ_i-σ_j crossbar. First we compute a maximum σ_i-σ_j flow value f_{ij} in the graph G'' as defined above the Corollary with $A = C_i'$ and $B = C_j'$. Since the cardinalities of σ_i and σ_j are at most $25\beta\gamma d$, the value of this flow is at most

$$25(\Delta - 1)\beta\gamma d \leq 25\beta^2\gamma d < d(C_i', C_j')/L - 3.$$

Thus we can invoke Corollary 4.9 to obtain a simple σ_i-σ_j crossbar of value $f_{ij}/2\Delta$. Let $U \subset \sigma_i$ and $W \subset \sigma_j$ denote the endpoints of the flow paths in the crossbar; these will be called the *crossbar ports.*

(3) Let $U' \subset \sigma_i$ and $W' \subset \sigma_j$. A *$\mathcal{T}_i$-$U'/\mathcal{T}_j$-W'-coordinated flow* consists of two flows: a flow in C_i' from a subset of terminals $\mathcal{T}_i$ to the set U', and a flow in C_j' from the corresponding set of terminals in $\mathcal{T}_j$ to the set W'. A coordinated flow is *simple* if both of the flows involved are simple; its *value* is the value of the two flows. In this step we compute a simple $\mathcal{T}_i$-$U/\mathcal{T}_j$-W-coordinated flow of maximum value.

Computing a maximum simple coordinated flow can be reduced to the matroid intersection theorem. The subsets of terminals in C_i that can be routed by disjoint paths to the cross-bar ports U form a matroid $M_{C_i'}(\mathcal{T}_i, U)$, as discussed in Section 3.2. Thus the problem of constructing a maximum simple coordinated flow is a special case of Edmonds' matroid intersection theorem; another consequence of this matroid structure is that the greedy algorithm gives a 2-approximation to the maximum simple coordinated flow. Alternately, Aumann and Rabani [2] observed that computing a maximum simple coordinated flow can be reduced to a maximum flow computation. Add a directed edge of capacity one from every terminal in C_i to the corresponding terminal in C_j. Now the maximum simple coordinated flow value is equal to the maximum simple flow from U, the crossbar ports of C_i', to W, the crossbar ports of C_j', in the graph consisting of C_i', C_j' and the directed edges connecting the terminals.

(4) We route at least a third of the connections that have reached the crossbar ports at U and W by the coordinated flow in step (3) on the edges of the crossbar using Lemma 4.5.

We claim that the resulting solution is at least an $\varepsilon_1 = \frac{\varepsilon}{6\Delta(1+\varepsilon)}$ fraction of the optimal.

Theorem 4.10 *The above procedure is an ε_1^{-1}-approximation algorithm for the subproblem involving terminal pairs $\mathcal{T}_{ij}$. The paths constructed by the procedure stay within distance at most $2d$ of C_i and C_j.*

Proof. The second statement is obvious. To prove the first one, consider a realizable subset $\mathcal{T}^*$ of $\mathcal{T}_{ij}$ of maximum cardinality, and let $\mathcal{P}^*$ denote the associated collection of edge disjoint paths. Obviously, the maximum value f_{ij} of a σ_i-σ_j flow in $G' = G - (C_i')^o - (C_j')^o$ is an upper bound on $|\mathcal{T}^*|$. By Corollary 4.9 we see that the simple crossbar constructed in step (2) is of size at least $f_{ij}/2\Delta$ and hence at least $|\mathcal{T}^*|/2\Delta$.

Consider the structure of the optimal solution $\mathcal{P}^*$. A $\mathcal{T}_i$-$\sigma_i/\mathcal{T}_j$-σ_j-coordinated flow of value $|T^*|$ is defined by the parts of the paths $P_\ell \in P^*$ from s_ℓ to the first intersection with σ_i, and from the last intersection with σ_j to t_ℓ. By the same bipartite coloring argument used in Corollary 4.9, there is a simple $\mathcal{T}_i$-$\sigma_i/\mathcal{T}_j$-σ_j-coordinated flow $\mathcal{G}$ of value $|T^*|/2\Delta$, which is at most $|U|$.

We now use the ε-linkage of the boundaries σ_i and σ_j to prove that the value of the maximum simple coordinated flow constructed in step (3) is at least $\frac{\varepsilon}{2\Delta(1+\varepsilon)}|T^*|$. Let $\mathcal{T}_i'$ and $\mathcal{T}_j'$ denote the set of terminals routed by the coordinated flow $\mathcal{G}$, and let U' and W' denote the endpoints of these flow paths on σ_i and σ_j respectively. Recall that $|U'| \le |U|$. For any $\mathcal{T}_i'' \subset \mathcal{T}_i'$, we have $f_s(\mathcal{T}_i'', U') = |\mathcal{T}_i''|$; since σ_i is ε-linked, Lemma 4.2 implies

$$f_s(\mathcal{T}_i'', U) \ge \frac{\varepsilon}{1+\varepsilon} \cdot f_s(\mathcal{T}_i'', U') = \frac{\varepsilon}{1+\varepsilon} \cdot |\mathcal{T}_i''|.$$

So in the matroid $M_{G''}(\mathcal{T}_i', U)$, every subset has rank at least $\frac{\varepsilon}{1+\varepsilon}$ times its cardinality. Since the same argument applies to $M_{G''}(\mathcal{T}_j', W)$, the matroid intersection theorem implies there is a simple $\mathcal{T}_i$-$U/\mathcal{T}_j$-W coordinated flow of value at least $\frac{\varepsilon}{1+\varepsilon} \cdot |\mathcal{T}_i'| = \frac{\varepsilon}{2\Delta(1+\varepsilon)} \cdot |\mathcal{T}^*|$. Thus the size of the coordinated flow that reaches the crossbar ports in step (3) is at least $3\varepsilon_1|\mathcal{T}^*|$. By Lemma 4.5, step (4) routes at least a third of this flow. ■

Next we give an $O(1)$-approximation algorithm for the subproblem associated with distance d. By Lemma 4.10 the solution to the (C_i, C_j) subproblem interferes only with other subproblems at most $4d$ distance away. We build an *interference graph* $\mathcal{K}$ on the set of pairs (C_i, C_j), joining two pairs if there is some edge within distance $2d$ of each. The uniform-diameter condition (ii) and Lemma 4.1 imply that each cluster pair has a constant number of neighbors in the graph $\mathcal{K}$. So by Brooks' Theorem, we can color the cluster pairs with a constant number of colors, so that no two pairs in the same color class interfere. Thus the above algorithm can be applied to all cluster pairs in one color class simultaneously. Taking the maximum number of terminals routed in any color class, we get an $O(1)$-approximation for the subproblem associated with distance d.

Theorem 4.11 *There is a polynomial-time $O(1)$-approximation for the problem of finding a maximum realizable subset of $\mathcal{T}$ in nearly-Eulerian uniformly high-diameter plane graphs in which all terminal pairs satisfy $d/2 \leq dist(s_\ell, t_\ell) \leq d$ for some parameter d.*

Finally, the original routing problem consists of at most $O(\log n)$ subproblems associated with a fixed distance d, so by taking the maximum value found in any subproblem, we obtain the main result of this section.

Corollary 4.12 *There is a polynomial-time $O(\log n)$-approximation for the problem of finding a maximum realizable subset of $\mathcal{T}$ in a nearly-Eulerian uniformly high-diameter plane graph.*

5 Optical Routing

The techniques developed in the previous section allow us to give an $O(\log n)$-approximation for optical routing, as well as an $O(1)$ approximation in the special case when all terminal pairs are roughly the same distance apart. Recall that given G and a set of terminal pairs $\mathcal{T}$, we seek to minimize the number of subsets into which $\mathcal{T}$ must be partitioned such that each subset is realizable in G. This minimum is denoted $\chi(\mathcal{T})$.

One way to approach this partitioning problem is as a set-cover problem. That is, one greedily uses an approximation for the maximization problem, assigning a new color for each realizable subset that is found this way. This approach was used in the case of the mesh by Aumann and Rabani [2] and leads to an $O(\log^2 n)$-approximation algorithms for the problem. Here we show how the matroidal tools mentioned in Section 3.2 lead to an $O(\log n)$-approximation algorithm.

As in the previous section, we break $\mathcal{T}$ into $O(\log n)$ subsets, such that $d(s_i, t_i)$ and $d(s_j, t_j)$ are within a constant factor of each other if they belong to the same subset. The $O(\log n)$-approximation will be a corollary of an $O(1)$-approximation for each subset. We break a single subproblem into problems $\mathcal{T}_{ij}$ for each pair of clusters (C_i, C_j). The Brooks' Theorem argument of the previous section (above Theorem 4.11) shows that it is enough to get a constant factor approximation for each $\mathcal{T}_{ij}$ subproblem; using different colors on different color classes of the interference graph $\mathcal{K}$, we thereby get a constant-factor approximation for the subproblem associated with a given distance and hence an $O(\log n)$-approximation for $\chi(\mathcal{T})$. We therefore turn to the problem of obtaining a constant approximation for $\chi(\mathcal{T}_{ij})$.

(1-2) As was done in the routing algorithm, we first build the augmented clusters C_i' and C_j', and a simple σ_i-σ_j crossbar of value at least $f_{ij}/2\Delta$, with endpoints $U \subseteq \sigma_i$ and $W \subseteq \sigma_j$.

(3) We next compute a minimum cardinality cover of $\mathcal{T}_{ij}$ by simple $\mathcal{T}_i$-$U/\mathcal{T}_j$-W-coordinated flows — since this involves covering a set by common independent sets in the two strongly base-orderable matroids $M_{G''}(\mathcal{T}_i, U)$ and $M_{G''}(\mathcal{T}_j, W)$, we can use the algorithm of Davies and McDiarmid to find such a minimum cover in polynomial time. Terminal pairs routed by different coordinated flows will receive different colors.

(4) We use Lemma 4.5 to break each coordinated flow into three color classes, each of which is routable using edge-disjoint paths on the crossbar. This is the coloring of the terminal pairs, with associated paths constructed in the obvious way by pasting the coordinated flow paths to the paths in the crossbar.

Theorem 5.1 *The number of colors used by the above algorithm is at most $6\Delta(1+\varepsilon^{-1})\chi(\mathcal{T}_{ij})$. The paths constructed by the procedure stay within distance at most $2d$ of C_i and C_j.*

Proof. Parts of the paths in each color class in an optical routing constitute a $\mathcal{T}_i$-$\sigma_i/\mathcal{T}_j$-σ_j-coordinated flow. Clearly the size of such a coordinated flow is at most f_{ij}. Therefore, $\mathcal{T}_{ij}$ can be covered by at most $\chi(\mathcal{T}_{ij})$ (not necessarily simple) $\mathcal{T}_i$-$\sigma_i/\mathcal{T}_j$-σ_j-coordinated flows each of which is of size at most f_{ij}. By a simple bipartite coloring argument this implies that $\mathcal{T}_{ij}$ can be covered by at most $2\Delta\chi(\mathcal{T}_{ij})$ simple $\mathcal{T}_i$-$\sigma_i/\mathcal{T}_j$-σ_j-coordinated flows each of which is of size at most $f_{ij}/2\Delta$.

We use the ε-linked property of the boundaries σ_i and σ_j and the Davies-McDiarmid theorem to show that the number of coordinated flows used in step (3) is at most $2\Delta(\varepsilon^{-1}+1)\chi(\mathcal{T}_{ij})$. Consider a simple coordinated $\mathcal{T}_i$-$\sigma_i/\mathcal{T}_j$-σ_j flow of size at most $f_{ij}/2\Delta$, and let $\mathcal{T}_i' \subseteq \mathcal{T}_i$ and $\mathcal{T}_j' \subseteq \mathcal{T}_i$ denote the set of terminals covered by this flow. By Lemma 4.2 the rank of any subset $\mathcal{T}_i'' \subset T_i'$ in the matroid $M_{G''}(\mathcal{T}_i, U)$ is at least $\frac{\varepsilon}{1+\varepsilon}|T_i''|$. The analogous statement holds for subsets of $\mathcal{T}_j'$. Therefore by the Davies-McDiarmid theorem the $\mathcal{T}_i'$-$\mathcal{T}_j'$ terminal pairs can be covered by at most $(1+\varepsilon^{-1})$ simple $\mathcal{T}_i'$-$U/\mathcal{T}_j'$-W coordinated flows, and hence $\mathcal{T}_{ij}$ can be covered by at most $2\Delta(1+\varepsilon^{-1})\chi(\mathcal{T}_{ij})$ simple $\mathcal{T}_i$-$U/\mathcal{T}_j$-W coordinated flows. Therefore, step (3) covers $\mathcal{T}_{ij}$ by at most this many coordinated flows. Finally, step (4) splits every coordinated flow into at most 3 colors. ■

Theorem 5.2 *There is a polynomial-time algorithm that in any nearly-Eulerian uniformly high-diameter plane graph routes a set of terminal pairs $\mathcal{T}$, using at most $O(\log n)$ times $\chi(\mathcal{T})$ colors. The algorithm uses at most $O(1)$ times $\chi(\mathcal{T})$ colors if all terminal pairs satisfy $d/2 \leq dist(s_\ell, t_\ell) \leq d$ for some parameter d.*

6 Conclusion

We have given $O(\log n)$-approximations for the maximum disjoint paths problem and the optical routing problem in a fairly general class of planar graphs. In our subsequent paper [14] we give an $O(1)$-approximation algorithm, and an on-line $O(\log n)$-approximation algorithm, for the maximum disjoint paths problem in a related class of graphs. The improved algorithm relies on two of the techniques developed in this paper — ε-linking and the crossbar construction. It is our hope that the techniques developed will be useful in attacking more general cases of this problem, and in improving the approximation ratio for optical routing to $O(1)$.

One natural goal is an approximation for edge-disjoint paths in an arbitrary even-degree planar graph. Though this seems quite difficult, we note that our main tools apply to any even-degree planar graph. If we remove the evenness condition, then we must be prepared to deal with 3-regular graphs, for which an approach based on crossbars is clearly of no use.

Beyond this lies the prospect of a polylogarithmic approximation for general graphs. Essentially the only work that has been done in the setting of general graphs has been on exact algorithms for a fixed number of terminal pairs or for special cases involving a graph embedded on a fixed surface, as in work of Robertson and Seymour and of Schrijver (see [11]). Most of the techniques presented in this paper cannot be directly applied in the non-planar case; but it is possible that analogous notions could prove useful.

Acknowledgments

We thank Yonatan Aumann and Yuval Rabani for showing us a preliminary version of their work on optical routing.

References

[1] A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, M. Sudan, "Efficient routing and scheduling algorithms for optical networks," *Proc. 5th ACM-SIAM SODA*, 1994, pp. 412–423.

[2] Y. Aumann, Y. Rabani, "Improved bounds for all-optical routing," *Proc. 6th ACM-SIAM SODA*, 1995, pp. 567–576.

[3] B. Awerbuch, Y. Azar, S. Plotkin, "Throughput-Competitive Online Routing," *Proc. 34th IEEE FOCS*, 1993, pp. 32–40.

[4] B. Awerbuch, R. Gawlick, F.T. Leighton, Y. Rabani, "On-line admission control and circuit routing for high performance computing and communication," *Proc. 35th IEEE FOCS*, 1994, pp. 412–423.

[5] A. Broder, A. Frieze, S. Suen, E. Upfal, "Optimal construction of edge-disjoint paths in random graphs," *Proc. 5th ACM-SIAM SODA*, 1994, pp. 603–612.

[6] A. Broder, A. Frieze, E. Upfal, "Existence and construction of edge-disjoint paths on expander graphs," *Proc. 24th ACM STOC*, 1992, pp. 140–149.

[7] J. Davies, C. McDiarmid, "Disjoint Common Traversals in Exchange Structures," *J. London Math. Soc.*, 14(1976), pp. 55–62.

[8] J. Edmonds, "Minimum partition of a matroid into independent subsets,", *J. Res. Nat. Bur. Standards B*, 69(1965), pp. 67–72.

[9] J. Edmonds, "Matroid intersection," *Annals of Discrete Mathematics*, 4(1979), pp. 39–49.

[10] A. Frank, "Edge-disjoint paths in planar graphs," *J. Comb. Theory Ser. B*, 39(1985), pp. 164–178.

[11] A. Frank, "Packing paths, cuts, and circuits — a survey," in *Paths, Flows, and VLSI-Layout*, B. Korte, L. Lovász, H.J. Prömel, A. Schrijver, Eds., Berlin: Springer-Verlag, 1990.

[12] N. Garg, V. Vazirani, M. Yannakakis, "Primal-dual approximation algorithms for integral flow and multicut in trees, with applications to matching and set cover," *Proc. ICALP*, 1993, pp. 64–75.

[13] R.M. Karp, "Reducibility among combinatorial problems," *Complexity of Computer Computations*, R.E. Miller, J.W. Thatcher, Eds., New York: Plenum Press, 1972.

[14] J. Kleinberg, É. Tardos, "Disjoint Paths in Densely Embedded Graphs," *Proc. 36th IEEE FOCS*, 1995. Also appears as Cornell University School of OR&IE technical report TR-1127.

[15] M.E. Kramer, J. van Leeuwen, "The complexity of wire routing and finding the minimum area layouts for arbitrary VLSI circuits," *Advances in Computing Research 2: VLSI Theory*, F.P. Preparata, Ed., London: JAI Press, 1984.

[16] M. Middendorf, F. Pfeiffer, "On the complexity of the disjoint paths problem," *Combinatorica*, 13(1993).

[17] H. Okamura, P. Seymour, "Multicommodity flows in planar graphs," *J. Comb. Theory Ser. B*, 31(1981), pp. 75–81.

[18] D. Peleg, E. Upfal, "Disjoint paths on expander graphs," *Combinatorica*, 9(1989), pp. 289–313.

[19] P. Raghavan, "Probabilistic construction of deterministic algorithms: approximating packing integer programs," *J. Computer and System Sciences*, 37(1988), pp. 130–143.

[20] P. Raghavan, C.D. Thompson, "Randomized rounding: a technique for provably good algorithms and algorithmic proofs," *Combinatorica*, 7(1987), pp. 365–374.

[21] P. Raghavan, E. Upfal, "Efficient all-optical routing," *Proc. 26th ACM STOC*, 1994, pp. 134–143.

[22] D. Wagner, K. Weihe, "A linear time algorithm for multicommodity flow in planar graphs," *Proc. First European Symposium on Algorithms*, 1993, pp. 384–395.

[23] D.J.A. Welsh, *Matroid Theory*, London: Academic Press, 1976.