Primal-dual approximation algorithms for node-weighted network design in planar graphs

Piotr Berman[a], Grigory Yaroslavtsev[a]

[a]{berman, grigory}@cse.psu.edu, Pennsylvania State University, University Park, PA 16802.

Abstract

We present primal-dual algorithms which give a 2.4 approximation for a class of node-weighted network design problems in planar graphs, introduced by Demaine, Hajiaghayi and Klein (ICALP'09). This class includes Node-Weighted Steiner Forest problem studied recently by Moldenhauer (ICALP'11) and other node-weighted problems in planar graphs that can be expressed using $(0, 1)$-proper functions introduced by Goemans and Williamson. We show that these problems can be equivalently formulated as feedback vertex set problems and analyze approximation factors guaranteed by different violation oracles within primal-dual framework developed for such problems by Goemans and Williamson.

For edge-weighted versions of these problems, as well as versions with uniform node weights, there has been significant progress in obtaining PTASs recently. There were proposed general techniques, such as bidimensionality of Demaine and Hajiaghayi, algorithmic theory of vertex separators of Feige, Hajiaghayi and Lee and contraction decomposition of Demaine, Hajiaghayi and Kawarabayashi. For Edge-Weighted Steiner Forest PTAS was obtained by Bateni, Hajiaghayi and Marx and by Eisenstat, Klein and Mathieu. In contrast, for more general node-weighted versions constant factor approximations via primal-dual algorithms remain the state of the art, while no APX-hardness is known.

1. Introduction

In feedback vertex set problems input is a graph $G = (V, E)$, a family of cycles $\mathcal{C}$ in G and a non-negative weight function $w: V \to \mathbb{R}^{\geq 0}$ on the set of vertices of G. The goal is to find a set of vertices $H \subset V$ which contains a node in every cycle in $\mathcal{C}$ such that the total weight of vertices in H is minimized. This is a special case of a general hitting set problem, when sets correspond to cycles in the graph.

Feedback Vertex Set (FVS) problem in a graph is the problem of finding a hitting set for all cycles. We consider the problem of hitting sets for cycles that satisfy some special properties. There are four natural examples.

- Odd cycles. If $H \subset V$ is a hitting set for all odd-length cycles then the subgaph of G, induced by the vertex set $V \setminus H$ is bipartite and removal of H is called *graph bipartization*. The corresponding hitting set problem is denoted as Bipartization (BIP).
- The set of all cycles which contain at least one node from a given set of nodes. The corresponding hitting set problem is known as Subset Feedback Vertex Set (S-FVS).
- The set of all directed cycles of a given directed graph. The corresponding problem is called Directed Feedback Vertex Set (D-FVS).
- In Node-Weighted Steiner Forest problem we are given a weighted undirected graph and a set of terminal pairs (s_i, t_i). The goal is to select a subset of vertices of the graph of minimum weight such that in the subgraph induced by these vertices all pairs of terminals are connected. In Section 2.1 we show that Node-Weighted Steiner Forest belongs to a class of problems which can be expressed as hitting set problems for some collection of cycles.

While in general graphs Feedback Vertex Set can be approximated within factor of 2 for all graphs, as shown by Becker and Geiger [4] and Bafna, Berman and Fujito [1], hitting a restricted family of cycles can be much harder. For example, the best known approximation ratio for graph bipartization in general graphs is

Problem	Approximation	Hardness of approximation
Feedback Vertex Set	2 [4, 1]	MAX-SNP complete [24, 28]:
Bipartization	$O(\log n)$ [14]	1.3606, if $P \neq NP$ [8]
Directed Feedback Vertex Set	$O(\log n \log\log n)$ [11]	$2 - \epsilon$, assuming UGC [21]
Subset Feedback Vertex Set	8 [10]	
Node-Weighted Steiner Forest	$O(\log k)$ [23]	$(1 - o(1)) \log k$ [12], if $NP \not\subseteq ZTIME(2^{n^\eta})$ for some $\eta > 0$

Table 1: General graphs

Problem	Previous work (our analysis)[1]	Our work	Hardness
FVS	10 [2], 3(18/7) [16]	2.4	NP-hard [30]
BIP, D-FVS, S-FVS	3(18/7) [16]		
Node-Weighted Steiner Forest	6 [5], 3(18/7) [26]		

Table 2: Planar graphs

$O(\log n)$ by Garg, Vazirani and Yannakakis [14]. For D-FVS the best known approximation is $O(\log n \log\log n)$, as shown by Even, Naor, Schieber and Sudan [11]. These and other results for general graphs are summarized in Table 1.

Yannakakis [30] has given an NP-hardness proof for many vertex deletion problems restricted to planar graphs which applies to all problems that we consider. Also, it is known that D-FVS is NP-hard even if both indegree and outdegree of every vertex are at most 3 (Garey and Johnson, [13, p. 191]). For planar graphs, the unweighted Feedback Vertex Set problem admits a PTAS, as shown by Demaine and Hajiaghayi [6] using a bidimensionality technique. Goemans and Williamson [16] created a framework for primal-dual algorithms that for planar instances of all above problems provide approximation algorithms with constant approximation factors. More specifically, they showed $9/4$-approximations for FVS, S-FVS, D-FVS and BIP. For Node-Weighted Steiner Forest it was shown by Demaine, Hajiaghayi and Klein [5] that the generic framework of Goemans and Williamson gives a 6-approximation which was improved to $9/4$-approximation by Moldenhauer [26]. However, the original paper by Goemans and Williamson [16] contains a mistake in the analysis. Similar mistake was repeated in [26]. We exhibit the mistake on an example and prove that no worse example exists. More precisely, primal-dual approximation algorithms of Goemans and Williamson for all problems described above give approximation factor $18/7$ rather than $9/4$. We also give an improved version of the violation oracle which can be used within the primal-dual framework of Goemans and Williamson and guarantees approximation factor 2.4. Results for planar graphs are summarized in Table 2.

The edge counterparts of the given problems, i.e. finding a minimum-weight subset of edges which intersects with every cycle in a given collection, are also well-studied. They reduce to vertex-weighted versions by adding a new vertex on each edge and assigning its weight to be equal to the weight of the edge. These problems are also significantly simpler, especially in planar graphs. Feedback Edge Set problem is a complement of the maximum spanning tree problem. Minimum-weight graph bipartization by edge removals is complementary to the maximum-weight cut problem which in planar graphs can be solved in polynomial time (Hadlock [19], Dorfman and Orlova [9]). Directed Feedback Edge Set problem in planar graphs reduces to finding a minimum-weight dijoin in the dual graph which can be solved in polynomial time (see, e.g. Grötschel, Lovász and Schrijver [17, p.253-254]). Edge-weighted Steiner Forest problem in planar graphs is NP-hard [13], but admits a PTAS, as shown recently by Bateni, Hajiaghayi and Marx [3].

[1] See discussion in the text.

Applications and ramifications. Node-weighted Steiner problems have been studied theoretically in many different settings, see e.g. [23, 27, 29, 25]. Applications of such problems range from maintenance of electric power networks [18] to computational sustainability [7]. Experimental evaluation of primal-dual algorithms for feedback vertex set problems in planar graphs in applications to VLSI design was shown by Kahng, Vaya and Zelikovsky [20].

As observed by Goemans and Williamson primal-dual algorithms for feedback vertex set problems in planar graphs have close connections to conjectures of Akiyama and Watanabe and Gallai and Younger about the size of minimum feedback vertex set in planar graphs. See [16] for more details.

Organization. We give basic definitions and preliminary observations in Section 2. In Section 2.1 we show that a wide class of node-weighted network design problems in planar graphs, introduced by Demaine, Hajiaghayi and Klein [5], can be equivalently defined as a class of hitting set problems for appropriately defined collections of cycles satisfying *uncrossing property*, as introduced by Goemans and Williamson [16]. In Section 3 we introduce local-ratio analog of primal-dual framework of Goemans and Williamson for such problems and give examples of violation oracles which can be used within this framework.

In Section 4 we give a corrected version of the analysis of the approximation factor achieved by the generic primal-dual algorithm with a violation oracle, presented by Goemans and Williamson in [16]. In Appendix B we present analysis of primal-dual algorithms with a new violation oracle which gives approximation factor 2.4. In Section 4.4 we show examples, on which these approximation factors are achieved.

2. Preliminaries

A *simple* cycle of length k is a sequence of vertices $v_1, \dots, v_{k+1}$, where $v_{k+1} \equiv v_1$, all vertices $v_1, \dots v_k$ are distinct, $(v_i, v_{i+1}) \in E$ for all $1 \le i \le k$ and all these edges are distinct. When working with simple graphs, the edge set above is uniquely defined. Note that in undirected simple graphs a simple cycle has length at least three. For a cycle C, the edge set of C is denoted as $E(C)$, although to simplify presentation we will abuse the notation slightly and sometimes refer to it as just C.

Every planar graph has a combinatorial embedding that for each vertex specifies a cyclic ordering of edges that are adjacent to it. A subset $U \subset V$ defines $G[U]$, the *induced subgraph* of G, with node set U and edges $\{(u, v) \in E : u, v \in U\}$. An embedding of a planar graph naturally defines embeddings of all its induced subgraphs. We denote the set of faces of a planar graph as F (for a standard definition of the set of faces via a combinatorial embedding, see e.g. [22]). The *planar dual* of a graph G is graph $G^* = (F, E')$ where F is the set of faces of G, and E' is the set of pairs of faces that share an edge. We select one face F_0 as the *outer face*.

For a simple cycle $C = (v_1, \dots, v_{k+1})$ we denote the set of faces that are surrounded by C as $\mathsf{Faces}(C)$. More formally, let E'' be the set of pairs of faces that share an edge that is not on C then in (F, E'') has exactly two connected components. We denote as $\mathsf{Faces}(C)$ the connected component of (F, E'') that does not contain the outer face F_0. A family of cycles $\mathcal{Z}$ is *laminar* iff for every $C, D \in \mathcal{Z}$ either $\mathsf{Faces}(C) \subset \mathsf{Faces}(D)$, or $\mathsf{Faces}(D) \subset \mathsf{Faces}(C)$, or $\mathsf{Faces}(D) \cap \mathsf{Faces}(C) = \varnothing$.

We use notation $\bullet$ to denote *contact* between two objects. More formally $u \bullet V$, if $u \subseteq V$. For example, we can have nodes and edges in contact with faces and cycles.

2.1. Uncrossable families of cycles and proper functions

Our algorithms apply to every family of cycles that satisfies the following:

Definition 2.1 (Uncrossing property [16]). *For any two cycles $C_1, C_2 \in \mathcal{C}$ such that there exists a path P_2 in C_2 which is edge-disjoint from C_1 and which intersects C_1 only at the endpoints of P_2, the following must hold. Let P_1 be a path in C_1 between the endpoints of P_2. Then either $P_1 \cup P_2 \in \mathcal{C}$ and $(C_1 \setminus P_1) \cup (C_2 \setminus P_2)$ contains a cycle in $\mathcal{C}$, or $(C_1 \setminus P_1) \cup P_2 \in \mathcal{C}$ and $(C_2 \setminus P_2) \cup P_1$ contains a cycle in $\mathcal{C}$.*

Many natural families of cycles satisfy the *uncrossing property*. Goemans and Williamson [16] showed this for FVS, D-FVS, BIP, and S-FVS. We show that a certain class of node-weighted connectivity problems in planar graphs can be expressed as problems of finding hitting sets for families of cycles satisfying the uncrossing property. To formalize this statement we introduce some definitions.

Definition 2.2 ($(0,1)$-proper function). *A Boolean function* $f\colon 2^V \to \{0,1\}$ *is* proper *if* $f(\emptyset) = 0$ *and it satisfies the following two properties:*

1. *(Symmetry)* $f(S) = f(V \setminus S)$.

2. *(Disjointness) If* $S_1 \cap S_2 = \emptyset$ *and* $f(S_1) = f(S_2) = 0$ *then* $f(S_1 \cup S_2) = 0$.

These two properties imply the property known as *complementarity*: if $A \subseteq S$ and $f(S) = f(A) = 0$ then $f(S \setminus A) = 0$.

For a set $S \subseteq V$, let $\Gamma(S)$ be its boundary, i.e. the set of nodes not in S which have a neighbor in S, or formally $\Gamma(S) = \{v \in V | v \notin S, \exists u \in S\colon (u,v) \in E\}$. As observed by Demaine, Hajiaghayi and Klein [5], a wide class of node-weighted network design problems can be formulated as the following generic integer program, where $f\colon 2^V \to \{0,1\}$ is a proper function:

$$\begin{aligned}
\text{Minimize:} \quad & \sum_{v \in V} w(v)x(v) & \\
\text{Subject to:} \quad & \sum_{v \in \Gamma(S)} x(v) \geq f(S) & \text{for all } S \subseteq V \\
& x(v) \in \{0,1\} & \text{for all } v \in V,
\end{aligned}$$

For example, for Node-Weighted Steiner Forest the corresponding $(0,1)$-proper function is defined as follows: $f(S) = 1$ iff there exists a pair of terminals (s_i, t_i), such that $|S \cap \{s_i, t_i\}| = 1$. The edge-weighted version of this program was introduced by Goemans and Williamson in [15]. Note that without loss of generality we can assume that the input graph is triangulated. Otherwise we add extra nodes of infinite cost inside each face and connect these new nodes by edges to all nodes on their faces without changing the cost of the optimum solution.

In Theorem 2.1 we will show that problems which can be expressed by a generic integer program above with some $(0,1)$-proper function f can also be expressed as problems of hitting uncrossable collections of cycles. We give some definitions and simplifying assumptions first.

Definition 2.3 (Active sets and boundaries). *For a proper function* $f\colon 2^V \to \{0,1\}$ *we say that sets* S*, such that* $f(S) = 1$ *are* active. *For an active set* S *we refer to its boundary* $\Gamma(S)$ *as* active boundary. *If an active boundary forms a simple cycle we call it* active simple boundary. *We denote the collection of all active simple boundaries as* $\mathcal{C}_A^f$.

Using this terminology, the generic integer program expresses the problem of finding a minimum weight hitting set for the collection of all active boundaries. Note that we can assume that all active singleton sets are included into the solution because each such set $\{s\}$ forms a boundary of its complement $V \setminus \{s\}$, which is active by symmetry, and thus $\{s\}$ has to be hit. We simplify the generic integer program using this observation so we can assume that all active boundaries don't contain active singleton sets. In particular, this implies by disjointntess of f that all active boundaries $\Gamma(S)$ are not active, namely $f(\Gamma(S)) = 0$.

We first show that finding a minimim weight hitting set for all active boundaries is equivalent to finding a minimum weight hitting set for $\mathcal{C}_A^f$ by showing that every active boundary contains an active simple boundary as a subset in Lemma Appendix A.1. Then we show that the family of active simple boundaries $\mathcal{C}_A^f$ satisfies the uncrossing property.

Theorem 2.1. *Let* $G(V,E)$ *be a triangulated planar graph. For every proper function* $f\colon 2^V \to \{0,1\}$ *the collection of active simple boundaries* $\mathcal{C}_A^f$ *forms an uncrossable family of cycles.*

Proof. Consider two active simple boundaries $\Gamma(S_1)$ and $\Gamma(S_2)$. If $\Gamma(S_2)$ crosses $\Gamma(S_1)$ then there exists a collection of edge-disjoint paths in $\Gamma(S_2)$ which we denote as P, such that each path $P_i \in P$ has only two nodes in common with $\Gamma(S_1)$. Each path $P_i \in P$ partitions $S_1 \setminus P_i$ into two parts which we denote as A_i^1 and A_i^2 respectively. Let's fix a path $P_i \in P$, such that at A_i^1 doesn't contain any other paths from P.

There are two cases: $A_i^1 \cap S_2 = \emptyset$ and $A_i^1 \subseteq S_2$. They are symmetric because if $A_i^1 \subseteq S_2$ we can replace the set S_2 by a set $S_2' = V \setminus S_2 \setminus \Gamma(S_2)$, ensuring that $A_i^1 \cap S_2 = \emptyset$. Note that the boundary doesn't change after such replacement, because $\Gamma(S_2) = \Gamma(S_2')$. By symmetry of f we have that $f(S_2) = f(V \setminus S_2) = 1$. Because $f(\Gamma(S_2)) = 0$ by disjointness we have $f(V \setminus S_2 \setminus \Gamma(S_2)) = f(S_2') = 1$, so S_2' is also an active set.

This is why it is sufficient to consider only the case when $A_i^1 \cap S_2 = \emptyset$. We will show the following auxiliary lemma:

Lemma 2.2. *Let $A_1, A, B \subseteq V$ be such that $A_1 \subseteq A$, $A_1 \cap B = \emptyset$ and $f(A) = f(B) = 1$. Then at least one of the following two statements holds:*

1. $f(A_1 \cup B) = f(A \setminus A_1) = 1$.
2. $f(A_1) = \max\,[f(B \setminus (A \setminus A_1)), f((A \setminus A_1) \setminus B)] = 1$.

The proof of the lemma follows from the properties of $(0,1)$-proper functions and is given in Appendix A.3

To show the uncrossing property for cycles $C_1 = \Gamma(S_1)$ and $C_2 = \Gamma(S_2)$ we select the paths in the definition of the uncrossing property as $P_1 = \Gamma(A_i^2) \setminus P_i$ and $P_2 = P_i$. Now we can apply Lemma 2.2 to sets A_i^1, S_1 and S_2, because $A_i^1 \subseteq S_1$, $A_i^1 \cap S_2 = \emptyset$ and $f(S_1) = f(S_2) = 1$. Thus, by Lemma 2.2 either $f(A_i^1 \cup S_2) = f(S_1 \setminus A_i^1) = 1$ or $f(A_i^1) = \max(f(S_2 \setminus (S_1 \setminus A_i^1)), f((S_1 \setminus A_i^1) \setminus S_2)) = 1$. In the first case we have $f(A_i^2) = f(A_i^1 \cup S_2) = 1$ and thus both cycles $P_1 \cup P_2 = \Gamma(A_i^2)$ and $(C_1 \setminus P_1) \cup (C_2 \setminus P_2) = \Gamma(A_i^1 \cup S_2)$ are active simple boundaries. In the second case $f(A_1) = 1$ and thus the cycle $(C_1 \setminus P_1) \cup P_2 = \Gamma(A_1)$ is an active simple boundary. The cycle $(C_2 \setminus P_2) \cup P_1$ is not necessarily simple, but it forms a boundary of an active set $(S_2 \setminus (S_1 \setminus A_i^1)) \cup ((S_1 \setminus A_i^1) \setminus S_2)$. Thus, by Lemma Appendix A.1 it contains an active simple boundary, which is a cycle in $\mathcal{C}_A^f$.

□

3. Algorithm

3.1. Generic local-ratio algorithm

We will use a local-ratio analog of a generic primal-dual algorithm formulated by Goemans and Williamson [16] which we state as Algorithm 1

Algorithm 1: Generic local-ratio algorithm $(G(V,E), w, \mathcal{C})$.

1 $\bar{w} \leftarrow w$.
2 $S \leftarrow \{u \in V : \ \bar{w}(u) = 0\}$.
3 **while** S *is not a hitting set for* $\mathcal{C}$ **do**
4 $\quad\mathcal{M} \leftarrow$ a collection of cycles returned by a violation oracle $\textsc{Violation}(G, \mathcal{C}, S)$.
5 $\quad c_{\mathcal{M}}(u) \leftarrow |\{M \in \mathcal{M} : \ u \bullet M\}|$, for all $u \in V$.
6 $\quad\alpha \leftarrow \min_{u \in V \setminus S} \frac{\bar{w}(u)}{c_{\mathcal{M}}(u)}$.
7 $\quad\bar{w}(u) \leftarrow \bar{w}(u) - \alpha c_{\mathcal{M}}(u)$, for all $u \in V$.
8 $\quad S \leftarrow \{u \in V : \ \bar{w}(u) = 0\}$.
end
9 **return** a minimal hitting set $H \subset S$ of $\mathcal{C}$.

Note that we don't need to specify the collection of cycles $\mathcal{C}$ explicitly. instead the generic algorithm requires that we specify an oracle $\textsc{Violation}(G, \mathcal{C}, S)$ used in Step 4. Given a graph G, collection of cycles $\mathcal{C}$ and a solution S if there are some cycles in $\mathcal{C}$ which are not hit by S this oracle should return a non-empty collection of such cycles, otherwise it should return the empty set. Such an oracle also allows to perform Step 3 and Step 9 without explicitly specifying $\mathcal{C}$.

The performance guarantee of the generic algorithm depends on the oracle used as described below. If $z : Z \to \mathbb{R}$ we use $z(Z)$ to denote $\sum_{a \in Z} z(a)$.

Theorem 3.1 (Local-ratio analog of Theorem 3.1 in [16]). *If the set $\mathcal{M}$ returned by a violation oracle used in Step 4 of the generic local-ratio Algorithm 1 satisfies that for any minimal solution $\breve{H}$:*

$$c_{\mathcal{M}}(\breve{H}) \leq \gamma |M|,$$

then Algorithm 1 returns a hitting set H of cost $w(H) \leq \gamma w(H^*)$, where H^* is the optimum solution.

We give the proof of this theorem for completeness in Appendix A.2.

The simplest violation oracles return a single cycle. Bar-Yehuda, Geiger, Naor and Roth [2] show that for FVS this approach can give a 10-approximation for planar graphs and Goemans and Willamson [16] improve it to a 5-approximation. They also analyzed an oracle, which returns a collection of all faces in $\mathcal{C}$, which are not hit by the current solution, and showed such oracle gives a 3-approximation for all families of cycles satisfying uncrossing property. Thus, by Theorem 2.1 such oracle gives a 3-approximation for all problems that we consider. We now give more complicated examples of violation oracles which give better approximation factors.

3.2. Minimal pocket violation oracles

The following oracle, introduced by Goemans and Williamson [16], returns a collection of faces in $\mathcal{C}$ inside a minimal *pocket* not hit by the current solution H.

Definition 3.1. *A* pocket *for a planar graph $G(V, E)$ and a cycle collection $\mathcal{C}$ is a set $U \subseteq V$ such that:*

1. *The set U contains at most two nodes with neighbors outside U.*
2. *The induced subgraph $G[U]$ contains at least one cycle in $\mathcal{C}$.*

Algorithm 2: Minimal-Pocket-Violation $(G, \mathcal{C}, S)$.

$\mathcal{C}_0 \leftarrow \{c \in \mathcal{C} : c \text{ not hit by } S\}$
Construct a graph G_S by removing from G:
  All edges which do not belong to any cycle in $\mathcal{C}_0$.
  All vertices which are not adjacent to any edges.
Let U_0 be a pocket for G_S and $\mathcal{C}_0$ which doesn't contain any other pockets.
return A collection of all cycles in $\mathcal{C}_0$ which are faces of $G_S[U_0]$.

As in the generic algorithm, we will not specify $\mathcal{C}$ and $\mathcal{C}_0$ explicitly, but will rather use an oracle to check relevant properties with respect to them. We show analysis of the approxiamtion factor obtained with this oracle in Section 4.

We will obtain a better approximation ratio by analyzing the following oracle in Section Appendix B.

Definition 3.2. *A* triple pocket *for a planar graph $G(V, E)$ and a cycle collection $\mathcal{C}$ is a set $U \subseteq V$ such that:*

1. *The set U contains at most three nodes with neighbors outside U.*
2. *The induced subgraph $G_S[U]$ has at least three faces in $\mathcal{C}$.*

The violation oracle Minimal-3-Pocket-Violation finds a minimal U_0 that is either a pocket or a triple pocket, and otherwise works like Minimal-Pocket-Violation.

4. Proof of 18/7 approximation ratio with pocket oracle

According to Theorem 3.1, to show that Algorithm 1 has approximation factor 18/7 it suffices to prove the following:

Theorem 4.1. *In every iteration of the generic local-ratio algorithm (Algorithm 1) with oracle* MINIMAL-POCKET-VIOLATION *for every minimal hitting set* $\breve{H}$ *of* $\mathcal{C}$ *we have* $c_{\mathcal{M}}(\breve{H}) \le \gamma|\mathcal{M}|$ *for* $\gamma = 18/7$.

The rest of this section is the proof of this theorem. The strategy of the proof is a variant of amortized analysis. We consider an arbitrary minimal hitting set $\breve{H}$ of $\mathcal{C}$-cycles in the residual graph G_S (as defined in Algorithm 2). For every cycle in $\mathcal{M}$ we get 1 unit of credit, and for a node $h \in \breve{H}$ we get $c_{\mathcal{M}}(h)/\gamma$ units of debit (*i.e.* negative), and we need to show that overall balance is non-negative. We start by decomposing the balance into smaller parts which are simpler to analyze than the balance of the entire pocket. The goal is to limit the impact of the nodes in $\breve{H}$ for which witness cycle A_h is not in $\mathcal{M}$. We decompose the pocket into parts that have at most two such nodes (Section 4.1). Further analysis refers to one such part.

We further simply the analysis of the balance by applying a pruning rule, each application of the pruning rule makes the instance smaller while the balance decreases. Thus it is enough to prove the claim when the pruning rule cannot be applied. In particular, this proves the claim if pruning produces an instance with no credits and debits. (Section 4.2).

Finally (Section 4.3) we define objects called envelopes and we assign all credits and debits to the envelopes. Then we show that each envelope has a non-negative balance. The nature of the pocket oracle eliminates conceivable envelopes with negative balance. In the next section we show that we eliminate more types of envelopes with the oracle MINIMAL-3-POCKET-VIOLATION which gives an approximation factor $12/5$.

Before we proceed we need several definitions.

Definition 4.1 (See also [16])**.**

(a) *Given a hitting set* $\breve{H}$ *for* $\mathcal{C}$ *we say that* $A \in \mathcal{C}$ *is a* witness cycle *for* $h \in \breve{H}$ *if* $A \cap \breve{H} = \{h\}$.
(b) *If* $\breve{H}$ *is a minimal hitting set, we can select* $\mathcal{A}_{\breve{H}} = \{A_h : \ h \in \breve{H}\}$*, a* family of witness cycles *for* $\breve{H}$.
(c) *Given a pocket* $G_S[U_0]$ *with* $\mathcal{M}$ *being set of faces of* $G_S[U_0]$ *that are in* $\mathcal{C}$ *we define* debit graph*, a bipartite graph* $\mathcal{G} = (\mathcal{M} \cup \breve{H}, \mathcal{E})$ *with edges* $\mathcal{E} = \{(M, h) \in \mathcal{M} \times \breve{H} : \ M \bullet h\}$.
(d) *For* $\mathcal{N} \subset \mathcal{M}$ *we define* $\mathcal{E}_{\mathcal{N}} = \{(M, h) \in \mathcal{E} : \ M \in \mathcal{N}\}$ *and* $\mathtt{balance}(\mathcal{N}) = |\mathcal{N}| - |\mathcal{E}_{\mathcal{N}}|/\gamma$.

Goemans and Williamson showed the following:

Lemma 4.2 (Lemma 4.2 in [16])**.** *For every collection of cycles* $\mathcal{C}$ *and every minimal hitting set* $\breve{H}$ *there exists a laminar family of witness cycles* $\mathcal{A}_{\breve{H}}$.

Observe also that the planar embedding of G_S defines a planar embedding of $\mathcal{G}$. We are going to use the fact that $\mathcal{G}$ is planar. By the definition, $c_{\mathcal{M}}(H) = |\mathcal{E}|$, so to prove Lemma 4.1 it suffices to show that

$$\mathtt{balance}(\mathcal{M}) \ge 0 \tag{1}$$

We prove inequality (1) using mathematical induction on $|\mathcal{M} \cup \mathcal{A}_{\breve{H}}|$.

4.1. Complex witness cycles and decomposition of the debit graph

In this subsection we will show a sufficient condition for inequality (1) and thus for our theorem.

Definition 4.2. *If* $A_g \in \mathcal{A}_h\mathtt{min}$ *and* $A_g \notin \mathcal{M}$ *we say that* $A_g \in \mathcal{A}_{\breve{H}}$ *is a* complex witness cycle *and that* g *is an outer (hit) node.*

Complex witness cycle A_h makes the analysis more complicated because there exist debits associated with pairs (M, h) but there is no credit for A_h. We reduce the problem of proving the non-negative balance of the debit graph to the problem of proving sufficiently high balances in simpler parts of that graph, where a part may have at most two complex witness cycles. There are two types of complex witness cycles:

Definition 4.3. *Let* $\mathcal{C}_g = \mathsf{Faces}(A_g) \cap \mathcal{C}$ *and* $\mathcal{M}_g = \mathcal{C}_g \cap \mathcal{M}$.

If all nodes of a complex witness cycle A_g *are in the pocket* U_0 *we say that* A_g *is a* hierarchical witness cycle. *Otherwise both contact nodes of* U_0 *belong to* A_g *and we say that* A_g *is a* crossing witness cycle.

First we discuss how to handle the hierarchical witness cycles. If there are such cycles, one of them, say A_g, has a minimal set $\mathsf{Faces}(A_g)$, and for such a cycle A_g we will show that

$$\mathsf{balance}(\mathcal{M}_g) \geq 1 - 1/\gamma \tag{2}$$

As a consequence of this inequality we can simplify the debit graph $\mathcal{G}$ by removing $\mathcal{M}_g$ and inserting in its place A_g. After that replacement $\mathcal{M}_g$ becomes $\{A_g\}$ and $\mathcal{E}_{\mathcal{M}_g}$ becomes $\{(A_g, g)\}$, so $\mathsf{balance}(\mathcal{M}_g)$ changes to $1 - 1/\gamma$. The replacement reduces $|\mathcal{A}_{\check{H}} \cup \mathcal{M}|$ while the inequality (2) assures that the $\mathsf{balance}(\mathcal{E})$ does not increase. This allows to invoke the inductive hypothesis.

We can repeat the simplification as long as there exists a hierarchical witness cycle. Note that we in the set of pairs $\mathcal{E}_{\mathcal{M}_g}$ (of the form "cycle M, hit node h") only g is a hit node with a complex witness cycle.

Now it remains to consider the case when for every hit node g that occurs in pairs of $\mathcal{E}$, if its witness cycle A_g is complex then A_g is a crossing witness cycle. Each such A_g contains a path $P_g \subset U_0$ between the two contact nodes of U_0, these paths cannot cross each other, thus they split pocket U_0 into *subpockets*; a subpocket has the boundary contained in two paths that either are of the form P_g or form a part of the outer face of $G[U_0]$.

Let $\mathcal{X}$ be the set of faces of $\mathcal{M}$ that are contained in a subpocket. It is sufficient to show that if $\mathcal{X} \neq \varnothing$ then

$$\mathsf{balance}(\mathcal{X}) \geq 1 - 2/\gamma. \tag{3}$$

We established that it suffices two consider two types of simplified debit graphs and to prove that they respectively satisfy inequalities (2) and 3. We can describe a more general type of a debit graph that includes both types as special cases. In our first type, A_g is a hierarchical witness cycle that does not contain other hierarchical witness cycles in its interior and we can define W, the set of nodes of G_S that are on A_g or its interior. In our second type, P_g and P_h are two paths connecting the contacts of U_0 and W is the set of nodes of G_S that are on these two paths or in the interior of the cycle formed by these two paths; W may contains outer nodes only on these paths. Let $\mathcal{M}_W$ be the set of faces of $G[W]$ that are in $\mathcal{M}$, the inequalities (2) and (3) have LHS equal to $\mathsf{balance}(\mathcal{M}_W)$.

Moreover, RHS of (2) and (3) also can be written with the same expression: if a is the number of the outer nodes in W than in both cases RHS $= 1 - a/\gamma$. The next lemma summarizes these observations.

Lemma 4.3. *To prove (1) it suffices to prove the following. Consider* $W \subset U_0$ *such that* $\check{H} \cap W$ *contains* $a \leq 2$ *outer nodes, all on the outer face of* $G[W]$. *Let* $\mathcal{M}_W$ *be the set of faces in* $\mathcal{M}$ *that have all nodes in* W. *Then*

$$\mathsf{balance}(\mathcal{M}_W) \geq 1 - a/\gamma \tag{4}$$

In the remainder of this section we will refer to W and a as introduced in this lemma, and we will assume that we have at most a outer nodes, all on the outer face of W.

4.2. Pruning

The following operation simplifies $\mathcal{M}_W$ and thus facilitates the proof of inequality (4).

Pruning $\mathcal{M}$ cycles of small degree. If a cycle $M \in \mathcal{M}_W \setminus \mathcal{A}_{\check{H}}$ participates in at most two edges in $\mathcal{E}$, we remove M from $\mathcal{M}$. If $A_h \in \mathcal{M}_W$ for some $h \in \check{H}$ and h participates in at most two edges in $\mathcal{E}$, including (A_h, h), we remove h from $\check{H}$ and A_h from $\mathcal{M}$.

Lemma 4.4. *Each step of pruning decreases* $|\mathcal{M}_W|$ *by exactly 1 and* $\mathsf{balance}(\mathcal{M}_W)$ *by at least* $1 - 2/\gamma$.

If $\mathcal{M}_W \neq \varnothing$ *and applications of pruning lead to* $\mathcal{M}_W = \varnothing$ *then before pruning inequality (4) was true.*

Proof. The first claim follows from the fact that a step of pruning decreases $|\mathcal{M}_W|$ by 1 and $\mathcal{E}_{\mathcal{M}_W}$ by at most 2.

The second claim for $a = 2$ follows from the first claim because we apply at least one step of pruning.

The second claim for $a = 1$ follows from the fact that before the last step of pruning we have $|\mathcal{M}_g| = 1$, *i.e.* $\mathcal{M}_g = \{M\}$ for some face M. If M is not a witness cycle then all witness cycles except A_g were eliminated, hence $\mathcal{E}_{\mathcal{M}_W}$ consists of exactly one pair (M, g). If M is a witness cycle A_h, then $\mathcal{E}_{\mathcal{M}_W}$ consists of exactly one pair (A_h, h). In both cases $\mathtt{balance}(\mathcal{M}_g) = 1 - 1/\gamma$. □

4.3. Envelopes

It remains to prove inequality (4) when pruning cannot be applied and $\mathcal{E}_{\mathcal{M}_W}$ is non-empty. We decompose $\mathtt{balance}(\mathcal{M}_W)$ into parts that can be analyzed using Euler formula.

We partition the set of all faces of G[W] into three parts as follows: $\mathcal{A} = \mathcal{M}_W \cap \mathcal{A}_{\breve{H}}$, $\mathcal{B} = \mathcal{M}_W \setminus \mathcal{A}_{\breve{H}}$ and $\mathcal{Z}$ is the set of faces in W which are not in $\mathcal{M}$. Consider the dual graph $G^* = (\mathcal{A} \cup \mathcal{Z}, E^*)$; we have an edge between two vertices of G^* if the corresponding faces share an edge. For every connected component C_i of G^* let E_i be the set of its *boundary edges* which are edges that are adjacent to one face in $C_i \subset \mathcal{A} \cup \mathcal{Z}$ and one face in $\mathcal{B}$. Every E_i forms a cycle (not necessarily simple) and we call such cycles *envelopes*.

Definition 4.4 (Envelopes in W)**.** Envelopes *in W are cycles consisting of boundary edges of connected components in the dual of $\mathcal{A} \cup \mathcal{Z}$. We assign each $h \in \mathcal{A}_{\breve{H}_W}$ to an envelope which contains h; we select the one which is on the boundary of the component of $\mathcal{A} \cup \mathcal{Z}$ that contains A_h. If h is an outer node, select the envelope which is on the boundary of the component of $\mathcal{A} \cup \mathcal{Z}$ that contains the other face of* G[W]*. An* outer envelope *is an envelope to which we assigned an outer node.*

We will split $\mathtt{balance}(\mathcal{M}_W)$ into balances of envelopes; an edge (M, h) is assigned to the envelope where we assigned h together with its witness cycle A_h; below we explain how to assign faces from $\mathcal{B}$ to envelopes. We will show that each envelope has a non-negative balance and at least one envelope has a sufficient positive balance.

Envelopes do not have to be simple cycles, but balance for non-simple envelopes is positive, we omit the details in the preliminary version, instead, we assume that each envelope is a simple cycle.

Envelopes: normal form. For an envelope **S** we define *principal neighbors*, $\mathcal{B}$-faces that have edges on **S**. Because we do not have pockets inside W, **S** must have at least 3 principal neighbors, except for the outer envelope. One can see that a traversal of **S** is a face of $\mathcal{G}$ of the form

$$(h_0, B_0, h_1 \dots, h_{a-1}, B_{a-1}, h_0).$$

Envelopes: normalizing neighbors. First, suppose that some $B \in \mathcal{B}$ contains both h_i and h_{i+1} but $B \neq B_i$. This case is excluded, because if there are some $\mathcal{M}$ faces between B and B_i we would have a pocket, and if there are not, B_i would have at most two neighbors: h_i and h_{i+1}, and we have pruned $\mathcal{M}$-cycles with at most two neighbors in $\mathcal{E}_{\mathcal{M}_W}$. (Observe that h_i, h_{i+1} do not have to belong to $\breve{H}$.)

Now we modify $\mathcal{E}_{\mathcal{M}_W}$ to $\mathcal{E}'$ as follows:

- we eliminate pairs of the form (A_h, h);
- for each envelope **S** we contract nodes of $\breve{H}$ that were assigned to **S** to a single node **S**.

The resulting graph is bipartite, but it can be a multi-graph. Therefore we modify it further: if B_i is a principal neighbor of **S** and (B, h_i), $(B, h_{i+1}) \in \mathcal{E}_{\mathcal{M}_W}$, in $\mathcal{E}'$ these two edges produce a single double edge. The edge set $\mathcal{E}'$ may remain a multigraph, *e.g.* if $B = B_i$ the same as outer neighbor B_j then we have two double edges from B to **S**. However, the edges from B to **S** cannot be consecutive in the circular ordering of **S**, hence this multigraph does not have faces with two edges only.

Note that non-outer hit nodes assigned to an envelope **S** are contact nodes of **S**; otherwise such a node h would belong to exactly two edges of $\mathcal{E}$: (A_h, h) and (B, h) where B is the principal neighbor that contains **S**; in that case h and A_h would be eliminated by pruning.

Let $n_{\mathbf{S}}$ be the number of contact nodes of **S**, and let $n_{\mathbf{S}} - \ell_{\mathbf{S}}$ be the number of hit nodes assigned to **S**. Also, let $d_{\mathbf{S}}$ be the number of edges of the modified graph that are incident to **S**.

If $\mathbf{S}$ is not the outer envelope, $n_{\mathbf{S}} \geq 3$.

Balance of envelopes We define $\mathrm{balance}_{\mathbf{S}}$ as the part of $\mathrm{balance}(\mathcal{E}_{\mathcal{M}_W})$ that can be attributed to $\mathbf{S}$. We will consider $\mathcal{E}_{\mathcal{M}_W}$ edges that are incident to H-nodes of $\mathbf{S}$, and $\mathcal{A}$-cycles of these nodes, and a "share" of B-cycles that can be attributed to $\mathbf{S}$ using $d_{\mathbf{S}}$ and the Euler formula.

Suppose that the modified graph has m nodes, d edges and f faces. Note that $d = \sum d_{\mathbf{S}}$. Also, $m = b + s$ where b is the number of faces in $\mathcal{B}$ and s is the number of envelopes.

Because each face has at least 4 edges and each edge is in two faces, we have $f \leq d/2$, hence the Euler formula implies $m = d - f + 2 \geq d/2 + 2$.

Thus we can allocate $d_{\mathbf{S}}/2$ nodes to an envelope $\mathbf{S}$, because one of these nodes is $\mathbf{S}$, we allocate $d_{\mathbf{S}}/2 - 1$ B-nodes, while to the outer envelopes we can allocate extra two $\mathcal{B}$-nodes.

To simplify the estimate of the number of pairs in $\mathcal{E}$ that are incident to $\mathbf{S}$ we assume first that $\ell_{\mathbf{S}} = 0$. For an non-outer envelope $\mathbf{S}$ we allocate $n_{\mathbf{S}}$ $\mathcal{A}$-faces, and this gives this estimate:

$$\begin{aligned}
\mathrm{balance}_{\mathbf{S}} &= \\
n_{\mathbf{S}} + d_{\mathbf{S}}/2 - 1 - (2n_{\mathbf{S}} + d_{\mathbf{S}})/\gamma &= \\
\frac{3}{2}n_{\mathbf{S}} - 1 + (d_{\mathbf{S}} - n_{\mathbf{S}})/2 - (3n_{\mathbf{S}} + (d_{\mathbf{S}} - n_{\mathbf{S}}))/\gamma &\geq && \gamma > 2 \\
\frac{3}{2}n_{\mathbf{S}} - 1 - 3n_{\mathbf{S}}/\gamma &\geq && n_{\mathbf{S}} \geq 3 \\
\frac{7}{2} - 9/\gamma &= 0 && \gamma = \frac{18}{7}
\end{aligned}$$

For every contact node h that is not a hit node (counted by $\ell_{\mathbf{S}}$) we decrease the number of credits in $\mathrm{balance}(\mathbf{S})$ by 1, as we do not have the credit of A_h, and the number of debits by $3/\gamma$ as we do not have (A_h, h), (B, h) and (B', h) where B, B' are the principal neighbors adjacent to h on $\mathbf{S}$. Thus $\mathrm{balance}(\mathbf{S}) \geq \ell_{\mathbf{S}}(3/\gamma - 1)$.

To estimate the balance of outer envelopes, observe that each outer envelope is in a separate connected component of the modified graph: all outer nodes are on the outer face of $G[W]$ so they are in the same connected component of G^*, hence an outer envelope is a connected component of edges of the boundary of that component of faces. In turn, $\mathcal{B}$-nodes of the modified graph are separated between the interiors of the outer envelopes. As a result, we can add 2 to the lower bound of each outer envelope.

Therefore an outer envelope $\mathbf{S}$ with a' outer nodes has balance at least

$$\begin{aligned}
&\frac{3}{2}n_{\mathbf{S}} - 1 - 3n_{\mathbf{S}}/\gamma + 2 - a'(1 - 1/\gamma) = \\
&3n_{\mathbf{S}}\left(\frac{1}{2} - 1/\gamma\right) + 1 - a'(1 - 1/\gamma) \geq && n_{\mathbf{S}} \geq 2 \\
&4 - 6/\gamma - a'(1 - 1/\gamma) > (2 - a')(1 - 1/\gamma) && \gamma \geq 2
\end{aligned}$$

For $a' < 2$ this gives the desired claim. For $a' = 2$ and $n_{\mathbf{S}}$ one can improve the estimate to $2(1 - 2/\gamma)$. Thus we have proven inequality (4) which suffices to prove Theorem 4.1, *i.e.* we showed that oracle Minimal-Pocket-Violation guarantees approximation factor $18/7$. We generalize this proof to show that oracle Minimal-3-Pocket-Violation gives approximation factor $12/5$ in Theorem Appendix B.1, which is deferred to Appendix B.

4.4. Tight examples

We show instances of graphs, on which the primal-dual algorithm with oracles Minimal-Pocket-Violation and Minimal-3-Pocket-Violation gives $18/7$ and $12/5$ approximations respectively.

Our examples are for the Subset Feedback Vertex Set problem. Recall that in this problem we need to hit all cycles which contain a specified set of "special" nodes. Our examples are graphs with no *pockets*

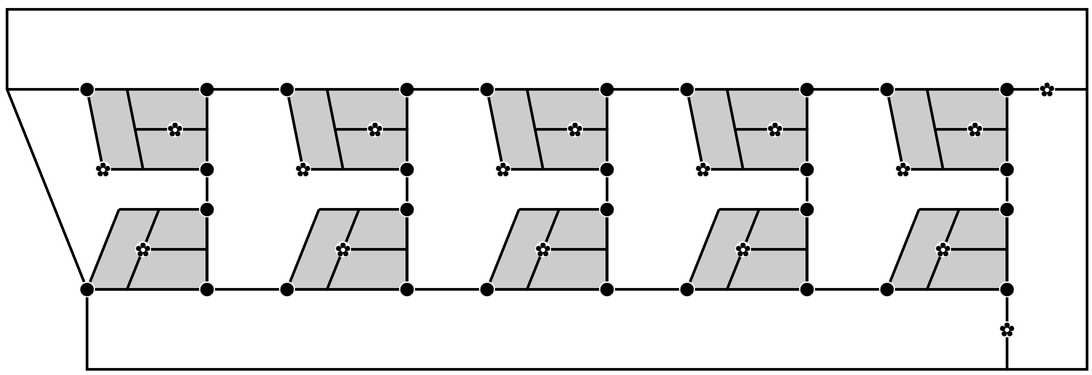

Figure 1: Family of instances of Subset Feedback Vertex Set with approximation factor 18/7 for the primal-dual algorithm with oracle Minimal-Pocket-Violation

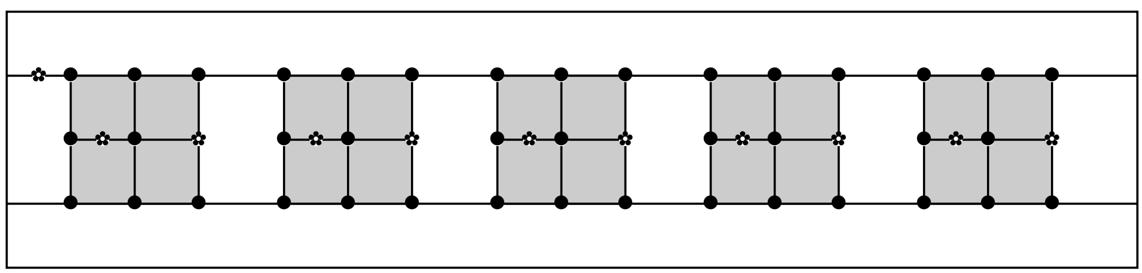

Figure 2: Family of instances of Subset Feedback Vertex Set with approximation factor 12/5 for primal-dual algorithm with oracle Minimal-3-Pocket-Violation

(or triple pockets), in which every face belongs to the family of cycles that we need to hit – this is ensured by selection of "special" nodes, which are marked with a star. The weights of vertices are assigned as follows. Given a node u with degree $d(u)$, its weight is $w(u) = d(u)$ if u is a solid dot and $w(u) = d(u) + \varepsilon$ otherwise (for some negligibly small value of ϵ).

First we show an example for the oracle Minimal-Pocket-Violation in Figure 1. Because there are no pockets, the first execution of the violation oracle returns the collection of all faces in the graph. Thus, in each building block of Picture 1 (which shows 5 such blocks from left to right), the primal-dual algorithm selects the black dots with total weight 18 while stars also form a valid solution with weight $7 + 3\epsilon$. Hence the ratio will be arbitrarily close to $18/7$, if we repeat the building block many times.

Similar family of examples for the oracle Minimal-3-Pocket-Violation is shown in Figure 2. In these examples there are no pockets or triple pockets, so the oracle Minimal-3-Pocket-Violation returns the collection of all faces in the graph. As above, the primal-dual algroithm selects the black dots with total weight 12 within each block, while the cost of the solution given by the stars is $5 + 2\epsilon$, so we can make the ratio arbitarily close to $12/5$.

[1] V. Bafna, P. Berman, and T. Fujito. A 2-approximation algorithm for the undirected feedback vertex set problem. *SIAM J. Discrete Math.*, 12(3):289–297, 1999.

[2] R. Bar-Yehuda, D. Geiger, J. S. Naor, and R. M. Roth. Approximation algorithms for the vertex feedback set problem with applications to constraint satisfaction and bayesian inference. In *Proceedings of the fifth annual ACM-SIAM symposium on Discrete algorithms*, SODA '94, pages 344–354, Philadelphia, PA, USA, 1994. Society for Industrial and Applied Mathematics.

[3] M. Bateni, M. Hajiaghayi, and D. Marx. Approximation schemes for steiner forest on planar graphs and graphs of bounded treewidth. *J. ACM*, 58:21:1–21:37, Oct. 2011.

[4] A. Becker and D. Geiger. Optimization of pearl's method of conditioning and greedy-like approximation algorithms for the vertex feedback set problem. *Artif. Intell.*, 83(1):167–188, 1996.

[5] E. Demaine, M. Hajiaghayi, and P. Klein. Node-weighted steiner tree and group steiner tree in planar graphs. In S. Albers, A. Marchetti-Spaccamela, Y. Matias, S. Nikoletseas, and W. Thomas, editors, *Automata, Languages and Programming*, volume 5555 of *Lecture Notes in Computer Science*, pages 328–340. Springer Berlin / Heidelberg, 2009.

[6] E. D. Demaine and M. Hajiaghayi. Bidimensionality: new connections between fpt algorithms and ptass. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, SODA '05, pages 590–601, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.

[7] B. Dilkina and C. Gomes. Solving connected subgraph problems in wildlife conservation. In A. Lodi, M. Milano, and P. Toth, editors, *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems*, volume 6140 of *Lecture Notes in Computer Science*, pages 102–116. Springer Berlin / Heidelberg, 2010.

[8] I. Dinur and S. Safra. On the hardness of approximating minimum vertex cover. *The Annals of Mathematics*, 162(1):pp. 439–485, 2005.

[9] Y. Dorfman and G. Orlova. Finding the maximal cut in a graph. *Engineering Cybernetics*, pages 502–506, 1972.

[10] G. Even, J. Naor, and L. Zosin. An 8-approximation algorithm for the subset feedback vertex set problem. *SIAM J. Comput.*, 30(4):1231–1252, 2000.

[11] G. Even, J. (Seffi) Naor, B. Schieber, and M. Sudan. Approximating minimum feedback sets and multicuts in directed graphs. *Algorithmica*, 20:151–174, 1998.

[12] U. Feige. A threshold of ln n for approximating set cover. *J. ACM*, 45(4):634–652, July 1998.

[13] M. R. Garey and D. S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., New York, NY, USA, 1990.

[14] N. Garg, V. V. Vazirani, and M. Yannakakis. Approximate max-flow min-(multi)cut theorems and their applications. *SIAM J. Comput.*, 25:235–251, February 1996.

[15] M. X. Goemans and D. P. Williamson. A general approximation technique for constrained forest problems. *SIAM J. Comput.*, 24(2):296–317, 1995.

[16] M. X. Goemans and D. P. Williamson. Primal-dual approximation algorithms for feedback problems in planar graphs. *Combinatorica*, 18:37–59, 1998.

[17] M. Grötschel, L. Lovász, and A. Schrijver. Geometric algorithms and combinatorial optimization. 1988.

[18] S. Guha, A. Moss, J. Naor, and B. Schieber. Efficient recovery from power outage (extended abstract). In *STOC*, pages 574–582, 1999.

[19] F. Hadlock. Finding a maximum cut of a planar graph in polynomial time. *SIAM J. Comput.*, 4(3):221–225, 1975.

[20] A. B. Kahng, S. Vaya, and A. Zelikovsky. New graph bipartizations for double-exposure, bright field alternating phase-shift mask layout. In *Proceedings of the 2001 Asia and South Pacific Design Automation Conference*, ASP-DAC '01, pages 133–138, New York, NY, USA, 2001. ACM.

[21] S. Khot and O. Regev. Vertex cover might be hard to approximate to within 2-epsilon. *J. Comput. Syst. Sci.*, 74(3):335–349, 2008.

[22] P. Klein. *Flatworlds: Optimization Algorithms for Planar Graphs.*

[23] P. N. Klein and R. Ravi. A nearly best-possible approximation algorithm for node-weighted steiner trees. *J. Algorithms*, 19(1):104–115, 1995.

[24] J. M. Lewis and M. Yannakakis. The node-deletion problem for hereditary properties is np-complete. *Journal of Computer and System Sciences*, 20(2):219 – 230, 1980.

[25] X. Li, X.-H. Xu, F. Zou, H. Du, P. Wan, Y. Wang, and W. Wu. A ptas for node-weighted steiner tree in unit disk graphs. In D.-Z. Du, X. Hu, and P. Pardalos, editors, *Combinatorial Optimization and Applications*, volume 5573 of *Lecture Notes in Computer Science*, pages 36–48. Springer Berlin / Heidelberg, 2009.

[26] C. Moldenhauer. Primal-dual approximation algorithms for node-weighted steiner forest on planar graphs. In *ICALP (1)*, pages 748–759, 2011.

[27] A. Moss and Y. Rabani. Approximation algorithms for constrained node weighted steiner tree problems. *SIAM J. Comput.*, 37(2):460–481, 2007.

[28] C. H. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *Journal of Computer and System Sciences*, 43(3):425 – 440, 1991.

[29] J. Remy and A. Steger. Approximation schemes for node-weighted geometric steiner tree problems. In C. Chekuri, K. Jansen, J. Rolim, and L. Trevisan, editors, *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques*, volume 3624 of *Lecture Notes in Computer Science*, pages 642–642. Springer Berlin / Heidelberg, 2005.

[30] M. Yannakakis. Node-and edge-deletion np-complete problems. In *Proceedings of the tenth annual ACM symposium on Theory of computing*, STOC '78, pages 253–264, New York, NY, USA, 1978. ACM.

Appendix A. Omitted proofs

Appendix A.1. Lemma Appendix A.1

Lemma Appendix A.1. *Assume that* (V, E) *is a connected triangulated planar graph,* f *is a proper function and* $\Gamma \subset V$ *has these properties: (a)* $f(A) = 0$ *for every* $A \subseteq \Gamma$ *and (b)* $f(B) = 1$ *for some* B *that is a connected component of* $V \setminus \Gamma$*. Then there exists* $C \subseteq \Gamma$ *with properies (a) and (b) such that* C *is a simple cycle.*

Proof. We will show the existence of C by taking any minimal subset of Γ, which has properties (a) and (b). The proof of the lemma follows from the Claim Appendix A.4, Claim Appendix A.5 and Claim Appendix A.6.

Claim Appendix A.2. *There exist two active connected components in* $V \setminus C$*.*

Proof. Let B_0 be a connected component of $V \setminus \Gamma$ such that $f(B_0) = 1$ which exists by property (b). By symmetry of f, we have $f(V \setminus B_0) = 1$, and by disjointness of f, we have $f(V \setminus B_0 \setminus \Gamma) = 1$. By applying disjointness again, there exists a connected component B_1 of $V \setminus B_0 \setminus \Gamma$ such that $f(B_1) = 1$. □

We will denote two active components, whose existence is guaranteed by Claim Appendix A.2 as B_0 and B_1.

Claim Appendix A.3. *Each* $c \in C$ *is adjacent to both* B_0 *and* B_1*.*

Proof. Otherwise, suppose that c is not adjacent to B_i and let $C' = C \setminus \{c\}$. Then B_i is a connected component of $V \setminus C'$, hence C' has property (b). Because C' is a subset of C it has property (a) and we get a contradiction with minimiality of C. □

Claim Appendix A.4. *Each* $c \in C$ *has at most 2 neighbors in* C*.*

Proof. Otherwise, let c_1, c_2, c_3 be three distinct nodes of C that are neighbors of c. Contract B_0 and B_1 to a single node, this allows to define a minor with nodes in two groups: c, B_0, B_1 in one group, c_1, c_2, c_3 in the second group, and this minor contains $K_{3,3}$. □

Claim Appendix A.5. *Each* $c \in C$ *has at least 2 neighbors in* C*, or* C *forms a cycle of length 2.*

Proof. In triangulated planar graphs neighbors of a node form a (not necessarily simple) cycle. Consider the cycle of neighbors of $c \in C$. On this cycle we can traverse a group of nodes from B_0, say $b_0^1, \ldots, b_0^k$. Because the cycle contains also nodes from other connected components of $V \setminus C$, this groups must be preceeded and followed by a node from C, say c_0 and c_1. If $c_0 \neq c_1$, Claim Appendix A.5 is true. Otherwise (c, c_0) and (c, c_1) are two separate edges from c to the same node, and the cycle $C' = (c, c_0 = c_1, c)$ has in its interior the group of nodes of B_0 that we have discussed and no other neighbors of c. Thus the nodes from B_1 that are neighbors of c must be in the exterior or C'; consequently there are no nodes of C in the interior of C', as they would violate Claim Appendix A.3. In the same way we can argue that the are no nodes of C in the exterior of C' and thus we conclude that $C = C'$. □

Claim Appendix A.6. *The subgraph* C *is connected.*

Proof. This is obvious when C is a 2-cycle, so it remains to consider the case when each node in C has exactly 2 neighbors in C. If C is not connected it forms a set of disjoint simple cycles and nodes of different cycles are not adjacent. Each such cycle is adjacent to exactly two connected components of $V \setminus C$ and we can consider a cycle C' with the minimal interior I. If $f(I) = 1$ then C' satisfies property (b), which is a contradiction with the minimality of C. If $f(I) = 0$ then nodes of C' don't satisfy Claim Appendix A.3, because they are adjacent to at most one active connected component. □

□

Appendix A.2. Analysis of the generic local-ratio algorithm

Here we give the proof of Theorem 3.1

Proof. When Algorithm 1 returns the hitting set H we have $\bar{w}(H) = 0$. Thus the cost of this hitting set $w(H)$ is the sum of decreases of $\bar{w}(H)$ that are caused by Step 7, and the same applies to an optimum solution H^*. To show approximation ratio γ it suffices to show that anytime we decrease $\bar{w}(H)$ by some γx we also decrease $\bar{w}(H^*)$ by at least x.

The decrease of $\bar{w}(H^*)$ is $\alpha c_{\mathcal{M}}(H^*)$. We can estimate $c_{\mathcal{M}}(H^*)$ it as follows: if $u \in H^*$ is responsible for hitting some m cycles in $\mathcal{M}$ then $c_{\mathcal{M}}(u) \geq m$, thus $c_{\mathcal{M}}(H^*) \geq |\mathcal{M}|$. Thus the decrease of $\bar{w}(H^*)$ is at least $\alpha|\mathcal{M}|$.

The decrease of $\bar{w}(H)$ is $\alpha c_{\mathcal{M}}(H)$. Thus to show the approximation ratio γ it suffices to show that for **every** minimal hitting set $\breve{H}$ we have $c_{\mathcal{M}}(\breve{H}) \leq \gamma|\mathcal{M}|$. □

Appendix A.3. Uncrossing proper sets (Lemma 2.2)

Here we give the proof of Lemma 2.2.

Proof. For a set $S \subseteq V$ we use notation $\bar{S}$ to denote $V \setminus S$, the complement of S in V. We will show that if the second statement is false then the first one is true. There are two cases.

Case 1. Suppose for the sake of contradiction that $f(A_1) = 0$. Then because $f(A) = 1$ we have $f(A \setminus A_1) = 1$ by disjointness.

By symmetry we have $f(B) = f(\bar{B}) = 1$. Because $A_1 \subseteq \bar{B}$ and $f(A_1) = 0$ by disjointness we have $f(\bar{B} \setminus A_1) = 1$. By symmetry this implies that $f(\overline{\bar{B} \setminus A_1}) = f(A_1 \cup B) = 1$, because $\overline{\bar{B} \setminus A_1} = B \cup A_1$.

Case 2. Suppose for the sake of contradiction that $f(B \setminus (A \setminus A_1)) = f((A \setminus A_1) \setminus B) = 0$.

Note that $A \setminus ((A \setminus A_1) \setminus B) = A_1 \cup (A \cap B)$. From $f((A \setminus A_1) \setminus B) = 0$ and $f(A) = 1$ we thus conclude by disjointness that $f(A \setminus ((A \setminus A_1) \setminus B)) = f(A_1 \cup (A \cap B)) = 1$. Because $A_1 \cap B = \emptyset$ the sets $B \setminus (A \setminus A_1)$ and $A_1 \cup (A \cap B)$ are disjoint. Using this together with the assumption that $f(B \setminus (A \setminus A_1)) = 0$ by disjointess we have that $f(B \setminus (A \setminus A_1) \cup A_1 \cup (A \cap B)) = f(A_1 \cup B) = 1$.

From the assumption that $f(B \setminus (A \setminus A_1)) = 0$ and $f(B) = 1$ by disjointness we have $f(B \setminus (B \setminus (A \setminus A_1))) = f(A \cap B) = 1$. Together with $f((A \setminus A_1) \setminus B) = 0$ by complementarity we have $f(A \setminus A_1) = 1$. □

Appendix B. Proof of 12/5 approximation ratio with triple pocket oracle

Theorem Appendix B.1. *In every iteration of the generic local-ratio algorithm (Algorithm 1) with oracle* MINIMAL-3-POCKET-VIOLATION *for every minimal hitting set* $\breve{H}$ *of* $\mathcal{C}$ *we have* $c_{\mathcal{M}}(\breve{H}) \leq \gamma|\mathcal{M}|$ *for* $\gamma = 12/5$.

Proof. We change the proof of Theorem 4.1 in only a few aspects.

We need to consider a triple pocket, with contact nodes c, d, e. The crossing witness cycles contain paths between contact nodes, now we have 3 pairs of contact nodes rather than one, hence 3 possible families of crossing paths. Thus we need to change the analysis of subpockets. A subpocket defined by two crossing paths from the same family has two contact nodes, as before. However, one subpocket can be defined by crossing paths from different families and this subpocket may have 3 contact nodes.

The subpocket with 3 contact nodes can have a negative balance if it contains less than two faces in $\mathcal{M}$. Thus it is important that if we have such subpocket we must also have nother subpockets that bring the total number of faces in $\mathcal{M}$ to at least 3, and those subpockets have positive balance.

If the subpocket with 3 contact nodes contains exactly 1 face in $\mathcal{M}$ its balance is at least $1 - 3/\gamma$, while the balance of the other pockets is at least $2(1 - 2/\gamma)$, for the total of $3 - 7/\gamma = 1/12$. If the subpocket with 3 contact nodes contains exactly 2 faces in $\mathcal{M}$, its balance is at least $2 - 5/\gamma$ while the other pockets have balance at least $1 - 2/\gamma$ and the estimate of the total is the same.

The effects of pruning are estimated exactly as before.

The balance of envelopes is estimated similarly, but we can make assumptions based on the fact that an envelope cannot be used to define a triple pocket. For a non-outer envelope the critical cases are when

$n_S \leq 4$. The case when $n_S \leq 2$ is excluded because we would define a pocket. The case when $n_S = 3$ and $\ell_S = 0$ is excluded because we would define a triple pocket. If $n_S = 3$ and $\ell_S \geq 1$ the balance is at least $(7/2-9/\gamma)+3/\gamma-1 = 5/2-6/\gamma = 5/2-5/2 = 0$. If $n_S \geq 4$ then the balance $\frac{3}{2}n_S - 1 - 3n_S/\gamma = 5-12/\gamma = 0$.

The estimate of the balance of an outer envelope is similar to an estimate of a non-outer envelope, so we just need to examine the differences to see that it cannot be lower. We can increase the credit given to an outer envelope by 2, because of Euler formula applied to the connected component of the debit graph that contains that envelope: one can see that each outer envelope is in a separate component.

The impact of an outer node on the balance can occur in two ways. We can replace the status of a hit node on the envelope from non-outer to outer. This decreases the credits by 1, and debits by $1/\gamma$, hence we subtract $7/12$ from the balance. We can also add an outer hit node that is not a contact node, this does not change credits but adds $1/\gamma$ to the debit, hence we subtract $5/12$ from the balance. Thus the worst case is that the balance increases by $2 - 3 \times 7/12 = 1/4$. □