

A PRIMAL-DUAL EXTENSION OF THE GOEMANS–WILLIAMSON ALGORITHM FOR THE WEIGHTED FRACTIONAL CUT-COVERING PROBLEM

NATHAN BENEDETTO PROENÇA^{1¶}, MARCEL K. DE CARLI SILVA^{2*}, CRISTIANE M. SATO^{3*},
AND LEVENT TUNÇEL^{1¶}

ABSTRACT. We study a weighted generalization of the fractional cut-covering problem, which we relate to the maximum cut problem via antiblocker and gauge duality. This relationship allows us to introduce a semidefinite programming (SDP) relaxation whose solutions may be rounded into fractional cut covers by sampling via the random hyperplane technique. We then provide a $1/\alpha_{\text{GW}}$ -approximation algorithm for the weighted fractional cut-covering problem, where $\alpha_{\text{GW}} \approx 0.878$ is the approximation factor of the celebrated Goemans–Williamson algorithm for the maximum cut problem. Nearly optimal solutions of the SDPs in our duality framework allow one to consider instances of the maximum cut and the fractional cut-covering problems as primal-dual pairs, where cuts and fractional cut covers simultaneously certify each other’s approximation quality. We exploit this relationship to introduce new combinatorial certificates for both problems, as well as a randomized polynomial-time algorithm for producing such certificates. In particular, we show how the Goemans–Williamson algorithm implicitly approximates a weighted instance of the fractional cut-covering problem, and how our new algorithm explicitly approximates a weighted instance of the maximum cut problem. We conclude by discussing the role played by geometric representations of graphs in our results, and by proving our algorithms and analyses to be optimal in several aspects.

1. INTRODUCTION

Let $G = (V, E)$ be a simple graph. For every $S \subseteq V$, the *cut* with *shore* S is the set $\delta(S) \subseteq E$ of edges which have precisely one vertex in S . For every nonnegative vector $z \in \mathbb{R}_+^E$ indexed by the edges, the *fractional cut-covering number* of (G, z) is

$$(1) \quad \text{fcc}(G, z) := \min \left\{ \mathbb{1}^\top y : y \in \mathcal{P}(V), \sum_{S \subseteq V} y_S \mathbb{1}_{\delta(S)} \geq z \right\},$$

where the power set of V is denoted by $\mathcal{P}(V)$, the incidence vector of $T \subseteq U$ is $\mathbb{1}_T \in \{0, 1\}^U$, and the vector of all-ones is $\mathbb{1}$. When z is integer-valued, the integer solutions of (1) correspond to multisets of cuts which cover each edge $e \in E$ at least z_e times, thus explaining the name “fractional cut-covering”. The unweighted version of this graph parameter — i.e., $\text{fcc}(G) := \text{fcc}(G, \mathbb{1})$ — is used by Šámal [39] to prove non-existence of cut-continuous functions between certain graphs. Such functions are maps between the edge sets of graphs that arise in the study of certain graph flow conjectures [12, 39]. A *fractional cut cover* of (G, z) is a feasible solution of (1).

For every $w \in \mathbb{R}_+^E$, the maximum weight of a cut with respect to w is

$$(2) \quad \text{mc}(G, w) := \max \{ w^\top \mathbb{1}_{\delta(S)} : S \subseteq V \}.$$

As larger cuts intuitively give rise to smaller covers, this suggests a combinatorial relationship between (2) and (1). The problem of computing $\text{mc}(G, w)$ is known as the *maximum cut problem*, and it is one of Karp’s original NP-hard problems [26]. Goemans and Williamson’s approximation algorithm [19] for this problem is one of the most celebrated applications of semidefinite programming. We denote by \mathbb{S}^V the Euclidean space of real symmetric $V \times V$ matrices, and by $\mathbb{S}_+^V \subseteq \mathbb{S}^V$ the cone of *positive semidefinite matrices*, i.e., the set of

¹DEPARTMENT OF COMBINATORICS AND OPTIMIZATION, UNIVERSITY OF WATERLOO

²INSTITUTE OF MATHEMATICS AND STATISTICS, UNIVERSITY OF SÃO PAULO

³CENTER FOR MATHEMATICS, COMPUTING AND COGNITION, FEDERAL UNIVERSITY OF THE ABC REGION

Date: November 24, 2023.

[¶]Research of this author was supported in part by a Discovery Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada .

^{*}This work was partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

symmetric matrices with nonnegative eigenvalues. The *Laplacian of G* is the linear function $\mathcal{L}_G: \mathbb{R}^E \rightarrow \mathbb{R}^{V \times V}$ defined by

$$(3) \quad \mathcal{L}_G(w) := \sum_{ij \in E} w_{ij}(e_i - e_j)(e_i - e_j)^\top \in \mathbb{R}^{V \times V} \quad \text{for each } w \in \mathbb{R}^E,$$

where $\{e_i : i \in V\} \subseteq \{0, 1\}^V$ are the canonical basis vectors. The *trace inner product* of $A, B \in \mathbb{R}^{V \times V}$ is $\langle A, B \rangle := \text{Tr}(A^\top B)$. The linear function $\text{diag}: \mathbb{R}^{V \times V} \rightarrow \mathbb{R}^V$ extracts the diagonal of a square matrix, and its adjoint $\text{Diag}: \mathbb{R}^V \rightarrow \mathbb{R}^{V \times V}$ builds a diagonal matrix from its argument such that $\text{Diag}(x)_{ii} = x_i$ for every $i \in V$. Write $X \succeq Y$ or $Y \preceq X$ for symmetric matrices X and Y if $X - Y$ is positive semidefinite. Goemans and Williamson's approximation algorithm implies that the optimal value of the semidefinite program (SDP)

$$(4a) \quad \eta(G, w) := \max\{\langle \tfrac{1}{4}\mathcal{L}_G(w), Y \rangle : Y \in \mathbb{S}_+^V, \text{diag}(Y) = \mathbb{1}\}$$

$$(4b) \quad = \min\{\rho : \rho \in \mathbb{R}_+, x \in \mathbb{R}^V, \rho \geq \mathbb{1}^\top x, \text{Diag}(x) \succeq \tfrac{1}{4}\mathcal{L}_G(w)\}$$

satisfies

$$(5) \quad \alpha_{\text{GW}}\eta(G, w) \leq \text{mc}(G, w) \leq \eta(G, w) \quad \text{for each } w \in \mathbb{R}_+^E,$$

where

$$(6) \quad \alpha_{\text{GW}} := \min_{0 < \theta \leq \pi} \frac{2}{\pi} \frac{\theta}{1 - \cos \theta} \approx 0.878$$

is the approximation factor. Equation (4b) follows from SDP Strong Duality, since both primal and dual SDPs have Slater points.

The *norm* of a vector $u \in \mathbb{R}^d$ is $\|u\| := \sqrt{u^\top u}$. For a fixed real number $t \geq 1$, a *vector t -coloring of G* is a function $f: V \rightarrow \mathbb{R}^d$ assigning a unit-norm vector $f(i) \in \mathbb{R}^d$ to each $i \in V$ such that $(t-1)f(i)^\top f(j) \leq -1$ for every $ij \in E$. Vector colorings were first introduced in [24]. The smallest value t for which a graph has a vector t -coloring is called the *vector chromatic number* of G , denoted by $\chi_{\text{vec}}(G)$. Šámal [38, Theorem 5.2] defines a map from fractional cut covers to vector colorings, thus proving that G has a vector t -coloring such that

$$(7) \quad 2\left(1 - \frac{1}{t}\right) \leq \text{fcc}(G).$$

Neto and Ben-Ameur [36, Proposition 17] tightened the relationship between fractional cut covers and vector colorings by showing that

$$(8) \quad \text{fcc}(G) \leq \frac{\pi}{\arccos(1/(1-t))}$$

for every vector t -coloring of G such that $t > 1$. Assume $E \neq \emptyset$, so $t \geq 2$, and set $\zeta := 1/(1-t)$, so $\zeta \in [-1, 0)$. Then (6) implies that

$$(9) \quad \frac{\pi}{\arccos(1/(1-t))} = \frac{\pi}{\arccos(\zeta)} = \frac{\pi}{2} \frac{1-\zeta}{\arccos(\zeta)} \frac{2}{1-\zeta} \leq \frac{1}{\alpha_{\text{GW}}} \frac{2}{1-\zeta} = \frac{1}{\alpha_{\text{GW}}} 2\left(1 - \frac{1}{t}\right).$$

Putting it all together, [36] combines (7), (8), (9), and monotonicity of $x \mapsto (1 - 1/x)$ to conclude that

$$(10) \quad 2\left(1 - \frac{1}{\chi_{\text{vec}}(G)}\right) \leq \text{fcc}(G) \leq \frac{1}{\alpha_{\text{GW}}} 2\left(1 - \frac{1}{\chi_{\text{vec}}(G)}\right).$$

As [36, Corollary 4] points out, the inequalities in (10) provide a polynomial-time computable approximation for the unweighted *number* $\text{fcc}(G)$, since $\chi_{\text{vec}}(G)$ is the optimal value of an SDP which can be approximated to any given precision in polynomial time.

We invite the reader to compare (5) and (10). Both describe constant-factor approximations that are computable from the optimal values of SDPs, and furthermore, the approximation factors are reciprocal. This work exploits and extends the ideas underlying (10). For every $z \in \mathbb{R}_+^E$, define

$$(11) \quad \eta^\circ(G, z) := \min\{\mu : \mu \in \mathbb{R}_+, Y \in \mathbb{S}_+^V, \text{diag}(Y) = \mu \mathbb{1}, \tfrac{1}{4}\mathcal{L}_G^*(Y) \geq z\}$$

where $\mathcal{L}_G^*: \mathbb{R}^{V \times V} \rightarrow \mathbb{R}^E$ is the adjoint of the Laplacian, i.e.,

$$(12) \quad (\mathcal{L}_G^*(Y))_{ij} = Y_{ii} + Y_{jj} - Y_{ij} - Y_{ji} \quad \text{for each } Y \in \mathbb{R}^{V \times V} \text{ and } ij \in E.$$

If y is a fractional cut cover for (G, z) , then $(\mu, Y) := (\mathbb{1}^\top y, \sum_{S \subseteq V} y_S (\mathbb{1} - 2\mathbb{1}_S)(\mathbb{1} - 2\mathbb{1}_S)^\top)$ is feasible for the SDP (11). Note that the RHS of (11) is an SDP. If $f: V \rightarrow \mathbb{R}^d$ is a vector t -coloring for some $t > 1$, we may define $\mu := 2(1 - 1/t)$ and $Y \in \mathbb{S}_+^V$ by $Y_{ij} := \mu f(i)^\top f(j)$ for every $i, j \in V$. Then $\text{diag}(Y) = \mu \mathbb{1}$ and $\frac{1}{4} \mathcal{L}_G^*(Y) \geq \mathbb{1}$, as $\frac{1}{4} (\mathcal{L}_G^*(Y))_{ij} = \frac{1}{2} \mu (1 - f(i)^\top f(j)) \geq \frac{1}{2} \mu (1 + \frac{1}{t-1}) = 1$ for every $ij \in E$. In this manner, the feasible solutions for (11) capture the geometry of vector colorings which enables (10). Using η° , we strengthen (10) to all nonnegative weights:

$$(13) \quad \eta^\circ(G, z) \leq \text{fcc}(G, z) \leq \frac{1}{\alpha_{\text{GW}}} \eta^\circ(G, z) \quad \text{for each } z \in \mathbb{R}_+^E.$$

This weighted generalization of (10) stands as the proper fractional cut-covering analogue to (5) for the maximum cut problem.

The similarity between (5) and (13) is the starting point of this work, whose main contributions include

- (i) pinpointing the relationship between the maximum cut problem (2) and the fractional cut-covering number (1) to gauge and antiblocker duality [16, 17];
- (ii) introducing (11) as the dual parameter to (4a), immediately obtaining that (13) is equivalent to (5) via a precisely defined bound conversion procedure [6, Sections 6 and 7];
- (iii) describing a randomized approximation algorithm, dual to the Goemans-Williamson algorithm, which rounds any nearly optimal solution of the SDP (11) to a $(1/\alpha_{\text{GW}})$ -approximately optimal fractional cut cover of (G, z) with very sparse support;
- (iv) pairing instances of the maximum cut and fractional cut-covering problems so that one can obtain approximately optimal solutions for *both* instances by solving a *single* SDP, and so that their approximate optimality can be certified by a simultaneous, (mostly) combinatorial certificate;
- (v) showing our algorithms to be best possible in several aspects;
- (vi) clarifying the role played by geometric representation of graphs in the aforementioned results.

Our algorithms run in polynomial time on the real-number machine model (see [7]) and hence in strongly polynomial time on a Turing machine, assuming rational input for the latter. A sufficient condition for our randomized algorithms to satisfy these complexity bounds is the existence of random number generation and Cholesky factorization oracles which run in polynomial time.

1.1. Organization of the Text. In order to facilitate reading, we unveil these results in increasing order of abstraction. We start at Section 2 by exhibiting a novel randomized approximation algorithm for the weighted fractional cut covering problem. The connection between η and η° is the main theme of Section 2. In Section 3.1 we express the relationship between both optimization problems via antiblocker [16, 17] and gauge duality [6]. We then show how computing either one of the parameters η or η° implicitly computes the other parameter, and how this can be leveraged to provide simultaneous combinatorial certificates for the approximate optimality of certain cuts and fractional cut covers. The existential results for certificates we prove in Section 3.3 are refined into efficient algorithms in Section 3.4. We recover the role played by vector colorings in this introduction by relating our approach to geometric representations of graphs in Section 4. Section 5 discusses possible improvements to our approximation algorithms by collecting noteworthy instances of our optimization problems: either instances where simpler approaches lead to degenerate behavior, or instances which show our bounds to be tight.

1.2. Notation. For each $n \in \mathbb{N}$, denote as usual $[n] := \{1, \dots, n\}$. The set of nonnegative real numbers is denoted by \mathbb{R}_+ , and the set of positive real numbers is \mathbb{R}_{++} . Let U be a finite set. We denote by \mathbb{R}^U the real vector space indexed by entries in U . For each $i \in U$, we denote by $e_i \in \{0, 1\}^U$ the i -th canonical basis vector. The 1-norm of a vector $z \in \mathbb{R}^U$ is $\|z\|_1 := \sum_{i \in U} |z_i|$, the ∞ -norm of z is $\|z\|_\infty := \max\{|z_i| : i \in U\}$, and the support of z is $\text{supp}(z) := \{i \in U : z_i \neq 0\}$.

2. A RANDOMIZED ROUNDING ALGORITHM FOR WEIGHTED FRACTIONAL CUT COVERING

Let V be a finite set. The set

$$\mathcal{E}^V := \{Y \in \mathbb{S}_+^V : \text{diag}(Y) = \mathbb{1}\}$$

is commonly referred to as the *elliptope*. We adopt (extended) Minkowski set operations and write $\mu \mathcal{E}^V := \{Y \in \mathbb{S}_+^V : \text{diag}(Y) = \mu \mathbb{1}\}$ for every $\mu \in \mathbb{R}$ and $\mathbb{R}_{++} \mathcal{E}^V := \{\mu Y : \mu \in \mathbb{R}_{++}, Y \in \mathcal{E}^V\}$. For every $Y \in \mathbb{R}_{++} \mathcal{E}^V$

and nonzero $h \in \mathbb{R}^V$, define

$$(14) \quad \text{GW}(Y, h) := \{i \in V : e_i^\top Y^{1/2} h \geq 0\},$$

where $Y^{1/2}$ is the unique positive semidefinite square root of Y . Equation (14) describes a possible implementation of the random hyperplane technique used by Goemans and Williamson [19] to sample a shore of a cut. Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and let the random variable $h: \Omega \rightarrow \mathbb{R}^V$ be a uniformly distributed unit vector. For every $Y \in \mathbb{R}_{++}\mathcal{E}^V$, we denote by $\text{GW}(Y): \Omega \rightarrow \mathcal{P}(V)$ the random variable given by

$$\text{GW}(Y): \omega \in \Omega \mapsto \text{GW}(Y, h(\omega)) \subseteq V,$$

which samples shores. It is proved in [19] that

$$(15) \quad \mathbb{P}(ij \in \delta(\text{GW}(Y))) = \frac{\arccos(Y_{ij})}{\pi} \quad \text{for every } ij \in E \text{ and } Y \in \mathcal{E}^V.$$

Algorithm 1 leverages the connection between the ellipsope and probability distributions on $\mathcal{P}(V)$ given by GW. We will exploit semidefinite programming to produce a matrix $Y \in \mathbb{R}_{++}\mathcal{E}^V$ from which we sample a fractional cut cover in (randomized) polynomial time by repeated sampling from $\text{GW}(Y)$. This section is devoted to proving correctness of Algorithm 1: namely, that it produces an approximately optimal fractional cut cover with high probability in polynomial time.

The algorithm works roughly as follows. First it creates a new weight vector \hat{z} from the input weight vector $z \in \mathbb{R}_+^E$ by rounding up the edge weights that are too small relative to $\|z\|_\infty$. Then it obtains a nearly optimal solution (μ, Y) for the SDP relaxation (11), which is used to sample cuts $\delta(\text{GW}(Y))$. The rounding up of the weights ensures that every edge has a significant probability of being in the random cut $\delta(\text{GW}(Y))$. The algorithm builds a fractional cut cover by using this sampling procedure T times independently, obtaining a vector of support size T , which is then scaled. The parameter T is defined in the range $\Theta(\ln(|V|))$ so that it is large enough to guarantee that we obtain a fractional cut cover with high probability. Our pseudocode abstracts away the important work of carefully choosing data structures: in particular, one needs to exploit the sparse nature of the fractional cut cover produced in its representation.

Algorithm 1 SDP-based randomized approximation algorithm for fcc

Parameters: a constant approximation factor $\beta \in (0, \alpha_{\text{GW}})$ parameterizes the algorithm APPROXFCC_β .

As in (27) and (28), define the following constants in terms of β :

$$\tau := 1 - \frac{\beta}{\alpha_{\text{GW}}} \in (0, 1), \quad \sigma := \varepsilon := \gamma := \frac{\tau}{3} \in (0, \frac{1}{3}), \quad \text{and} \quad C := 81\sqrt{2}\pi/\tau^{5/2}.$$

Input: a graph $G = (V, E)$ and edge weights $z \in \mathbb{R}_+^E$

Output: $\text{APPROXFCC}_\beta(G, z)$ is a fractional cut cover of (G, z) with high probability with objective value bounded above by $\frac{1}{\beta} \text{fcc}(G, z)$ and support size bounded above by $T := \lceil C \ln(|V|) \rceil$, as in (28)

```

1 procedure  $\text{APPROXFCC}_\beta(G, z)$ 
2    $(V, E) \leftarrow G$ 
3    $\hat{z}_e \leftarrow \text{MAX}(z_e, \frac{1}{2}\varepsilon\|z\|_\infty)$  for each  $e \in E$  ▷ Round up  $z$ 
4   Find a feasible  $(\mu, Y)$  for  $\eta^\circ(G, \hat{z})$  in (11) with objective value  $\mu \leq \eta^\circ(G, \hat{z}) + \sigma\|z\|_\infty$ 
5    $y \leftarrow 0 \in \mathbb{R}_+^{\mathcal{P}(V)}$ 
6   repeat  $T$  times
7      $S \leftarrow \text{GW}(Y)$  ▷ Sample a shore  $S \subseteq V$  via the random hyperplane technique
8      $y_S \leftarrow y_S + 1$ 
9   end
10  return  $\frac{\mu}{(1-\gamma)\alpha_{\text{GW}}} \frac{1}{T} y$ 
11 end procedure
```

Neto and Ben-Ameur [36, Section 4] already scale the probability distribution on $\mathcal{P}(V)$ given by $\text{GW}(Z)$ to define a fractional cut cover from some $Z \in \mathbb{R}_{++}\mathcal{E}^V$. By placing this construction in the weighted setting, we obtain Proposition 1, which directly justifies our definition of η° .

Proposition 1. Let $G = (V, E)$ be a graph and let $z \in \mathbb{R}_+^E$. Let (μ, Y) be feasible for (11). Set $y \in \mathbb{R}^{\mathcal{P}(V)}$ by

$$(16) \quad y_S := \frac{\mu}{\alpha_{\text{GW}}} \mathbb{P}(\text{GW}(Y) = S) \quad \text{for every } S \subseteq V.$$

Then y is a fractional cut cover for (G, z) with objective value $\mathbb{1}^\top y = \frac{1}{\alpha_{\text{GW}}} \mu$. In particular,

$$(17) \quad z_{ij} \leq \frac{\mu}{\alpha_{\text{GW}}} \mathbb{P}(ij \in \delta(\text{GW}(Y))) \quad \text{for every } ij \in E.$$

Proof. We may assume that $\mu > 0$. Set $\bar{Y} := \mu^{-1}Y \in \mathcal{E}^V$. By the definition in (14), we have that $\text{GW}(Y, h) = \text{GW}(\bar{Y}, h)$ for every $h \in \mathbb{R}^V$, which implies that $\text{GW}(Y) = \text{GW}(\bar{Y})$. Let $ij \in E$. Since $\bar{Y} \in \mathcal{E}^V$, we have that $\bar{Y}_{ij} \in [-1, 1]$. If $z_{ij} = 0$, then (17) holds trivially. Assume that $z_{ij} > 0$. Since $\text{diag}(\bar{Y}) = \mathbb{1}$, we get from (12) that

$$0 < z_{ij} \leq \frac{1}{4}(\mathcal{L}_G^*(Y))_{ij} = \mu \frac{1}{4}(\mathcal{L}_G^*(\bar{Y}))_{ij} = \mu \frac{1}{2}(1 - \bar{Y}_{ij}).$$

Thus $\bar{Y}_{ij} < 1$, so $\arccos(\bar{Y}_{ij}) > 0$, and $\mathbb{P}(ij \in \delta(\text{GW}(\bar{Y}))) > 0$ by (15). Hence

$$z_{ij} \leq \mu \frac{1}{2}(1 - \bar{Y}_{ij}) = \mu \frac{\pi}{2} \frac{1 - \bar{Y}_{ij}}{\arccos(\bar{Y}_{ij})} \mathbb{P}(ij \in \delta(\text{GW}(\bar{Y}))) \quad \text{by (15)}$$

$$\leq \mu \frac{1}{\alpha_{\text{GW}}} \mathbb{P}(ij \in \delta(\text{GW}(\bar{Y}))) \quad \text{by (6)}$$

$$= \sum_{\substack{S \subseteq V: \\ ij \in \delta(S)}} \frac{\mu}{\alpha_{\text{GW}}} \mathbb{P}(\text{GW}(\bar{Y}) = S) = \sum_{\substack{S \subseteq V: \\ ij \in \delta(S)}} y_S = \left(\sum_{S \subseteq V} y_S \mathbb{1}_{\delta(S)} \right)_{ij}.$$

As this holds for every $ij \in E$ for which $z_{ij} > 0$, we conclude that y is a fractional cut cover for (G, z) with objective value $\mathbb{1}^\top y = \frac{1}{\alpha_{\text{GW}}} \mu$, and (17) holds. \square

Corollary 2. Let $G = (V, E)$ be a graph, and let $z \in \mathbb{R}_+^E$. Then

$$(13) \quad \eta^\circ(G, z) \leq \text{fcc}(G, z) \leq \frac{1}{\alpha_{\text{GW}}} \eta^\circ(G, z).$$

Proof. As mentioned after (12), the first inequality holds since $(\mu, Y) := (\mathbb{1}^\top y, \sum_{S \subseteq V} y_S (\mathbb{1} - 2\mathbb{1}_S)(\mathbb{1} - 2\mathbb{1}_S)^\top)$ is feasible for the SDP (11), for every fractional cut cover y for (G, z) . The second inequality in (13) follows directly from Proposition 1. \square

Remark 3. We will define in (18) below the random variable capturing the repeated sampling behavior of Algorithm 1. One might question the purpose of employing a randomized approach when Proposition 1 defines a precise solution with a guaranteed approximation factor to the optimal value. We provide two compelling reasons: in Proposition 1,

- (i) it is a challenging task to compute the probabilities $\mathbb{P}(\text{GW}(Y) = S)$ in (16): while the marginal probability $\mathbb{P}(ij \in \delta(\text{GW}(Y)))$ is determined by (15), computing $\mathbb{P}(\text{GW}(Y) = S)$ requires joint probabilities for all the vertices in the shore S ;
- (ii) the fractional cut cover y obtained may have exponential support size: indeed there are instances (e.g., the complete graph with uniform weights) for which there exist optimal solutions for (11) that result in a vector y given as in (16) with exponential support size; see Appendix B.

The randomized procedure helps to address these two issues.

It is natural to consider the problem of, given a matrix Y and a real number μ as in Proposition 1, finding a fractional cut cover with polynomial support size and objective value bounded above by $(1/\alpha_{\text{GW}})\mu$; that such an exponential space saving exists follows from Carathéodory's theorem. The randomized approach below will in fact reduce the support size to $O(\ln(n))$, a reduction of *two* orders of magnitude compared to the fractional cut cover from Proposition 1.

Let $T \in \mathbb{N} \setminus \{0\}$ and let $\gamma \in (0, 1)$. Let $G = (V, E)$ be a graph, and let $Y \in \mathbb{S}^V$ and $\mu \in \mathbb{R}_+$ be such that $Y \in \mu \mathcal{E}^V$. Let $S_1, \dots, S_T \subseteq V$ be independent, identically distributed random shores sampled by $\text{GW}(Y)$. Define

$$(18) \quad \mathcal{A}_{T, \gamma}(G, Y) := (\mathcal{F}, y), \text{ where } \mathcal{F} := \{S_1, \dots, S_T\} \subseteq \mathcal{P}(V) \text{ and } y := \frac{\mu}{(1 - \gamma)\alpha_{\text{GW}}} \frac{1}{T} \sum_{t=1}^T e_{S_t} \in \mathbb{R}_+^{\mathcal{P}(V)}.$$

Informally, $\mathcal{A}_{T,\gamma}(G, Y)$ produces a (scaled version of a) sparse surrogate for the probability distribution of $\text{GW}(Y)$. It is immediate that

$$(19) \quad \mathbb{1}^\top y = \frac{1}{\alpha_{\text{GW}}} \frac{1}{1-\gamma} \mu \quad \text{and} \quad |\mathcal{F}| \leq T.$$

The parameter γ regulates the deviation from the objective value obtained in Proposition 1, while T needs to be chosen large enough so that concentration results imply that the desired level of accuracy is achieved with high probability.

Let $G = (V, E)$ be a graph. Šámal [38, Theorem 5.2] uses Chernoff's bound to show that, by sampling sufficiently many cuts, one can obtain a fractional cut cover for $(G, \mathbb{1})$. Proposition 4 improves on this work, by providing an explicit bound on the actual number of cuts that suffices in the general weighted setting. We will use that, for every $x \in [-1, 1]$ and $y \in [0, 2]$,

$$(20) \quad \text{if } x \leq 1 - y \text{ then } \arccos(x) \geq \sqrt{2y}.$$

This follows monotonicity of \arccos , and the inequality $\arccos(1 - y) \geq \sqrt{2y}$ for each $y \in [0, 2]$.

Proposition 4. Let $\xi, \kappa, \gamma \in \mathbb{R}$ be such that $0 < \xi \leq 1 \leq \kappa$ and $0 < \gamma < 1$. Let $G = (V, E)$ be a graph on n vertices, and let $z \in \mathbb{R}_+^E$ be nonzero. Let (μ, Y) be feasible in (11). Let $T \geq \lceil 3\pi(\frac{\kappa}{\xi})^{1/2} \frac{1}{\gamma^2} \ln(n) \rceil$ be an integer and set $(\mathcal{F}, y) := \mathcal{A}_{T,\gamma}(G, Y)$. If

$$(21a) \quad \mu \leq \kappa \|z\|_\infty$$

and

$$(21b) \quad z \geq \xi \|z\|_\infty \mathbb{1},$$

then

$$(22) \quad \mathbb{P}\left(\sum_{S \in \mathcal{F}} y_S \mathbb{1}_{\delta(S)} \geq z\right) \geq 1 - \frac{1}{n}.$$

Proof. Let S_1, \dots, S_T be random shores defined as in (18). For every $ij \in E$, set $X_{ij} := |\{t \in [T] : ij \in \delta(S_t)\}|$ and $p_{ij} := \mathbb{P}(ij \in \delta(\text{GW}(Y)))$. By construction, for each $ij \in E$, the random variable X_{ij} has binomial distribution on T trials with success probability p_{ij} and

$$\left(\sum_{S \in \mathcal{F}} y_S \mathbb{1}_{\delta(S)}\right)_{ij} = \frac{\mu}{(1-\gamma)\alpha_{\text{GW}}} \frac{1}{T} X_{ij}.$$

By (17), we have that $p_{ij} \geq \frac{\alpha_{\text{GW}}}{\mu} z_{ij}$ and so $\mathbb{E}(X_{ij}) \geq T \frac{\alpha_{\text{GW}}}{\mu} z_{ij}$. Thus, in order to prove (22), it suffices to show that $\mathbb{P}(\exists e \in E, X_e \leq (1-\gamma)\mathbb{E}(X_e)) \leq 1/n$.

Set $\bar{Y} := \mu^{-1}Y$ and let $ij \in E$. Using $z \neq 0$, (21b), (21a), feasibility of (μ, Y) in (11), and (12), we obtain

$$\frac{2\xi}{\kappa} = \frac{2\xi \|z\|_\infty}{\kappa \|z\|_\infty} \leq \frac{2z_{ij}}{\kappa \|z\|_\infty} \leq \frac{2z_{ij}}{\mu} \leq \frac{1}{2} \mathcal{L}_G^*(\bar{Y})_{ij} = 1 - \bar{Y}_{ij}.$$

Hence $\bar{Y}_{ij} \leq 1 - \frac{2\xi}{\kappa}$. Using (15) and (20) we see that

$$p_{ij} = \frac{\arccos(\bar{Y}_{ij})}{\pi} \geq \frac{2}{\pi} \sqrt{\frac{\xi}{\kappa}}.$$

By Chernoff's bound,

$$\mathbb{P}(X_{ij} \leq (1-\gamma)\mathbb{E}(X_{ij})) \leq \exp\left(-\frac{\gamma^2 \mathbb{E}(X_{ij})}{2}\right) = \exp\left(-\frac{\gamma^2 T p_{ij}}{2}\right) \leq \exp\left(-\frac{\gamma^2 T}{\pi} \sqrt{\frac{\xi}{\kappa}}\right).$$

Hence, by union bound and the lower bound on T ,

$$\begin{aligned} \mathbb{P}(\exists e \in E, X_e \leq (1-\gamma)\mathbb{E}(X_e)) &\leq n^2 \exp\left(-\frac{\gamma^2 T}{\pi} \sqrt{\frac{\xi}{\kappa}}\right) \\ &\leq \exp\left(2 \ln(n) - \frac{3\pi \kappa^{1/2} \ln(n)}{\xi^{1/2} \gamma^2} \frac{\gamma^2}{\pi} \sqrt{\frac{\xi}{\kappa}}\right) = \frac{1}{n}. \end{aligned} \quad \square$$

Remark 5. When applying Proposition 4, it is apparent that (21b) is the most stringent condition. Indeed, the other requirements can be satisfied by a nearly optimal solution to the SDP in (11). On the other hand, (21b) restricts the applicability of our procedure to certain well-behaved values of z . This is not an artifact of our analysis, but an unavoidable consequence of the repeated sampling approach. Appendix A shows that, for every $\varepsilon \in (0, 2)$, there exists an instance (G, z) which has an optimal solution $(\bar{\mu}, \bar{Y})$ for (26a) such that there is an edge ij with

$$\mathbb{P}(ij \in \delta(\text{GW}(\bar{Y}))) = \frac{\arccos(1 - 2\varepsilon + \varepsilon^2/2)}{\pi},$$

which can be made arbitrarily small. This, in turn, increases the number of samples needed to produce a cut cover, i.e., a set of cuts whose union is the whole edge set of the graph. Theorem 6 solves this issue by perturbing the edge weights z to a vector \hat{z} that satisfies (21b).

Let $G = (V, E)$ be a graph. We will use that

$$(23) \quad \eta^\circ(G, z_0) \leq \eta^\circ(G, z_1), \quad \text{for every } z_0, z_1 \in \mathbb{R}_+^E \text{ such that } z_0 \leq z_1,$$

$$(24) \quad \eta^\circ(G, z_0 + z_1) \leq \eta^\circ(G, z_0) + \eta^\circ(G, z_1) \quad \text{for every } z_0, z_1 \in \mathbb{R}_+^E, \text{ and}$$

$$(25) \quad \|z\|_\infty \leq \eta^\circ(G, z) \quad \text{for every } z \in \mathbb{R}_+^E.$$

These facts follow from SDP Strong Duality: if $G = (V, E)$ is a graph and $z \in \mathbb{R}_+^E$, then

$$(26a) \quad \eta^\circ(G, z) = \min \left\{ \mu : \mu \in \mathbb{R}_+, Y \in \mathbb{S}_+^V, \text{diag}(Y) = \mu \mathbb{1}, \frac{1}{4} \mathcal{L}_G^*(Y) \geq z \right\}$$

$$(26b) \quad = \max \left\{ z^\top w : w \in \mathbb{R}_+^E, x \in \mathbb{R}^V, \frac{1}{4} \mathcal{L}_G(w) \preceq \text{Diag}(x), \mathbb{1}^\top x \leq 1 \right\}.$$

The optimization problems (26a) and (26b) form a primal-dual pair of SDPs. Note that $(\hat{\mu}, \hat{Y}) := (2\|z\|_\infty, 2\|z\|_\infty I)$ and $(\hat{w}, \hat{x}) := (0, |V|^{-1} \mathbb{1})$ are relaxed Slater points of (26a), and (26b), respectively, so SDP Strong Duality ensures both problems have optimal solutions attaining a common optimal value; see, e.g., [34, Theorem 7.1.2]. The proofs of (23) and (24) are immediate from (26a) and (26b), respectively. To prove (25), let $z \in \mathbb{R}_+^E$ and note that, for every $ij \in E$,

$$2 \text{Diag}(e_i + e_j) - \mathcal{L}_G(e_{ij}) = 2(e_i e_i^\top + e_j e_j^\top) - (e_i - e_j)(e_i - e_j)^\top = (e_i + e_j)(e_i + e_j)^\top \in \mathbb{S}_+^V.$$

Hence, for every $ij \in E$, the pair $(\bar{w}, \bar{x}) := (e_{ij}, \frac{1}{2}(e_i + e_j))$ is feasible in (26b) for (G, z) . Thus (25) holds.

Theorem 6. Let $\beta \in (0, \alpha_{\text{GW}})$ and set

$$(27) \quad \tau := 1 - \frac{\beta}{\alpha_{\text{GW}}} \in (0, 1) \quad \text{and} \quad C := 81\sqrt{2}\pi/\tau^{5/2}.$$

There exists a randomized polynomial-time algorithm which takes as input a graph $G = (V, E)$ on n vertices and a vector $z \in \mathbb{R}_+^E$ and outputs (\mathcal{F}, y) , where $\mathcal{F} \subseteq \mathcal{P}(V)$ and $y \in \mathbb{R}_+^{\mathcal{F}}$ are such that

$$\mathbb{P}\left(\sum_{S \in \mathcal{F}} y_S \mathbb{1}_{\delta(S)} \geq z\right) \geq 1 - \frac{1}{n}, \quad \mathbb{1}^\top y \leq \frac{1}{\beta} \text{fcc}(G, z), \quad \text{and} \quad |\mathcal{F}| \leq \lceil C \ln(n) \rceil = O(\ln(n)).$$

That is, with high probability, y is a $\frac{1}{\beta}$ -approximately optimal solution for (1) with logarithmic support size.

Proof. We start by setting up the constants that will be used in the proof (and in Algorithm 1). Set

$$(28) \quad \sigma := \varepsilon := \gamma := \frac{\tau}{3} \in (0, \frac{1}{3}) \quad \text{and} \quad T := \lceil C \ln(n) \rceil,$$

as in the preamble to Algorithm 1. The constants σ , ε , and γ are chosen so that

$$(29) \quad \frac{1 - \gamma}{1 + \varepsilon + \sigma} \alpha_{\text{GW}} = \left(\frac{1 - \tau/3}{1 + 2\tau/3} \right) \alpha_{\text{GW}} = \left(1 - \frac{\tau}{1 + 2\tau/3} \right) \alpha_{\text{GW}} \geq (1 - \tau) \alpha_{\text{GW}} = \beta.$$

We define \hat{z} by rounding up entries that are smaller than $\frac{1}{2}\varepsilon\|z\|_\infty$. Set $\hat{z} \in \mathbb{R}_+^E$ by $\hat{z}_{ij} := \max\{z_{ij}, \frac{1}{2}\varepsilon\|z\|_\infty\}$ for every $ij \in E$. (Note that this done in Line 3 of Algorithm 1.)

Let (μ, Y) be a feasible solution for $\eta^\circ(G, \hat{z})$ in (11) with objective value

$$(30) \quad \mu \leq \eta^\circ(G, \hat{z}) + \sigma\|z\|_\infty = \eta^\circ(G, \hat{z}) + \sigma\|\hat{z}\|_\infty.$$

Note that this appears in Line 4 in Algorithm 1. Such a nearly optimal solution (μ, Y) can be computed in polynomial time due to the existence of strict Slater points for the SDPs in (26a); see Appendix C.

Problem	SDP Solution Properties	Rounding Procedure	Rounding Analysis	Algorithm
Fractional Cut-Covering	Proposition 1	(18)	Proposition 4	Theorem 6
Simultaneous Certificates	Proposition 15	(64)	Proposition 21	Proposition 22
Maximum Cut Certificates	Theorem 18			Theorem 23
Fractional Cut-Covering Certificates	Theorem 19			Theorem 24

FIGURE 1. Section 2 produces solutions accompanied by certificates of their approximate optimality. These developments parallel Section 2: we exploit the properties of optimal solutions to SDP relaxations in a rounding procedure. As an auxiliary step, we study *simultaneous* approximate solutions to both problems.

We now move to the final and randomized part of Algorithm 1. Set $(\mathcal{F}, y) := \mathcal{A}_{T, \gamma}(G, Y)$. Note that y is the solution produced by Algorithm 1. To finish the proof we will prove that y is a fractional cut cover for (G, z) with probability at least $1 - \frac{1}{n}$, and it has support size at most $\lceil C \ln(n) \rceil$ and objective value at most $(1/\beta) \text{fcc}(G, z)$.

We will apply Proposition 4 to show that y is a fractional cut cover for (G, \hat{z}) with probability at least $1 - \frac{1}{n}$. Since $\hat{z} \geq z$, this implies that y is a fractional cut cover for (G, z) with probability at least $1 - \frac{1}{n}$. Set $\xi := \varepsilon/2 = \tau/6$ and $\kappa := 3$, and recall the definition of γ in (28). By construction, (21b) holds. Note that

$$T = \lceil C \ln(n) \rceil = \left\lceil \frac{81\sqrt{2}}{\tau^{5/2}} \ln(n) \right\rceil = \left\lceil 3\pi \left(\frac{18}{\tau} \right)^{1/2} \left(\frac{3}{\tau} \right)^2 \ln(n) \right\rceil = \left\lceil 3\pi \left(\frac{\kappa}{\xi} \right)^{1/2} \left(\frac{1}{\gamma} \right)^2 \ln(n) \right\rceil,$$

so the lower bound on Proposition 4 is met. We will check that $\mu \leq \kappa \|\hat{z}\|_\infty$, that is, that (21a) holds. We claim that

$$(31) \quad \eta^\circ(G, \mathbb{1}) \leq 2.$$

This follows from feasibility of $(\bar{\mu}, \bar{Y}) := (2, 2I)$ in (11) for $(G, \mathbb{1})$, as $\frac{1}{4} \mathcal{L}_G^*(I) = \frac{1}{2} \mathbb{1}$ by (12). Hence, (30), (23), (31), and $\sigma < 1$ imply

$$\mu \leq \eta^\circ(G, \hat{z}) + \sigma \|\hat{z}\|_\infty \leq \eta^\circ(G, \|\hat{z}\|_\infty \mathbb{1}) + \sigma \|\hat{z}\|_\infty = (\eta^\circ(G, \mathbb{1}) + \sigma) \|\hat{z}\|_\infty \leq 3 \|\hat{z}\|_\infty,$$

so Proposition 4 applies.

The support size of y is $|\mathcal{F}| \leq T = \lceil C \ln(n) \rceil$ by (19). Finally, we bound $\mathbb{1}^\top y$:

$$\begin{aligned}
\mathbb{1}^\top y &= \frac{1}{\alpha_{\text{GW}}} \frac{1}{1-\gamma} \mu && \text{by (19)} \\
&\leq \frac{1}{\alpha_{\text{GW}}} \frac{1}{1-\gamma} (\eta^\circ(G, \hat{z}) + \sigma \|\hat{z}\|_\infty) && \text{by (30)} \\
&\leq \frac{1}{\alpha_{\text{GW}}} \frac{1}{1-\gamma} (\eta^\circ(G, z + \tfrac{1}{2}\varepsilon \|\hat{z}\|_\infty \mathbb{1}) + \sigma \|\hat{z}\|_\infty) && \text{by (23), as } \hat{z} \leq z + \tfrac{1}{2}\varepsilon \|\hat{z}\|_\infty \mathbb{1} \\
&\leq \frac{1}{\alpha_{\text{GW}}} \frac{1}{1-\gamma} (\eta^\circ(G, z) + \tfrac{1}{2}\varepsilon \|\hat{z}\|_\infty \eta^\circ(G, \mathbb{1}) + \sigma \|\hat{z}\|_\infty) && \text{by (24)} \\
&\leq \frac{1}{\alpha_{\text{GW}}} \frac{1}{1-\gamma} (\eta^\circ(G, z) + \varepsilon \|\hat{z}\|_\infty + \sigma \|\hat{z}\|_\infty) && \text{by (31)} \\
&= \frac{1}{\alpha_{\text{GW}}} \frac{1}{1-\gamma} (\eta^\circ(G, z) + (\varepsilon + \sigma) \|\hat{z}\|_\infty) && \text{since } \|z\|_\infty = \|\hat{z}\|_\infty \\
&\leq \frac{1}{\alpha_{\text{GW}}} \frac{1+\varepsilon+\sigma}{1-\gamma} \eta^\circ(G, z) && \text{by (25)} \\
&\leq \frac{1}{\beta} \eta^\circ(G, z) && \text{by (29).}
\end{aligned}$$

□

3. A PRIMAL-DUAL EXTENSION OF THE GOEMANS–WILLIAMSON ALGORITHM WITH CERTIFICATES OF APPROXIMATE OPTIMALITY

Section 2 describes an approximation algorithm for the fractional cut-covering problem. The work of Goemans and Williamson is so ubiquitous in our reasoning that one may claim Algorithm 1 to be “dual” to the algorithm described in [19]. This language suggests a primal-dual approach, where cuts and fractional cut covers simultaneously certify each other’s (approximate) optimality via a suitable notion of “weak duality”. This section provides randomized polynomial-time algorithms exploiting this idea. Fix a desired approximation factor $\beta \in (0, \alpha_{\text{GW}})$. Given a fractional cut-covering instance, we produce a fractional cut cover whose $(1/\beta)$ -approximate optimality is certified by a maximum cut instance with one of its β -approximately optimal solutions. Symmetrically, the input may be a maximum cut instance, and the algorithm then produces a β -approximately optimal cut and certifies it via a fractional cut-covering instance with one of its $(1/\beta)$ -approximately optimal solutions. The alignment of subtopics between this section and Section 2 are highlighted by Figure 1.

3.1. Gauge Duality. This subsection presents the gauge duality theory that permeates and forms the foundational basis for our results. In this manner, this subsection places our work within the literature and equips readers with a theoretical framework that can lead to new results.

Proposition 7. Let $G = (V, E)$ be a graph. The functions $\text{mc}(G, \cdot)$ and $\text{fcc}(G, \cdot)$ satisfy

$$(32a) \quad \text{mc}(G, w) = \max\{w^\top z : z \in \mathbb{R}_+^E, \text{fcc}(G, z) \leq 1\} \quad \text{for every } w \in \mathbb{R}_+^E,$$

$$(32b) \quad \text{fcc}(G, z) = \max\{z^\top w : w \in \mathbb{R}_+^E, \text{mc}(G, w) \leq 1\} \quad \text{for every } z \in \mathbb{R}_+^E,$$

$$(32c) \quad w^\top z \leq \text{mc}(G, w) \text{fcc}(G, z) \quad \text{for every } w, z \in \mathbb{R}_+^E.$$

The functions $\eta(G, \cdot)$ and $\eta^\circ(G, \cdot)$ satisfy

$$(33a) \quad \eta(G, w) = \max\{w^\top z : z \in \mathbb{R}_+^E, \eta^\circ(G, z) \leq 1\}, \quad \text{for every } w \in \mathbb{R}_+^E,$$

$$(33b) \quad \eta^\circ(G, z) = \max\{z^\top w : w \in \mathbb{R}_+^E, \eta(G, w) \leq 1\}, \quad \text{for every } z \in \mathbb{R}_+^E,$$

$$(33c) \quad w^\top z \leq \eta(G, w)\eta^\circ(G, z) \quad \text{for every } w, z \in \mathbb{R}_+^E.$$

Remark. The striking similarities between (32) and (33) underscore the existence of a theoretical framework that explains this phenomenon, rather than being a fortunate coincidence. It turns out that the functions $\text{mc}(G, \cdot)$, $\text{fcc}(G, \cdot)$, $\eta(G, \cdot)$, and $\eta^\circ(G, \cdot)$ are positive definite monotone gauges, which we shall define presently. Furthermore, $\text{mc}(G, \cdot)$ and $\text{fcc}(G, \cdot)$ form a dual pair, as do $\eta(G, \cdot)$ and $\eta^\circ(G, \cdot)$.

Proof of Proposition 7. Equation (32b) follows directly from Linear Programming Strong Duality, as

$$\begin{aligned} \text{fcc}(G, z) &= \min\left\{\mathbb{1}^\top y : y \in \mathbb{R}_+^{\mathcal{P}(V)}, \sum_{S \subseteq V} y_S \mathbb{1}_{\delta(S)} \geq z\right\} \\ &= \max\{z^\top w : w \in \mathbb{R}_+^E, \forall S \subseteq V, w^\top \mathbb{1}_{\delta(S)} \leq 1\} \\ &= \max\{z^\top w : w \in \mathbb{R}_+^E, \text{mc}(G, w) \leq 1\}, \end{aligned}$$

and (32c) is then straightforward. Next, we show that (32a) holds. We have that

$$\begin{aligned} \text{mc}(G, w) &= \max\{w^\top \mathbb{1}_{\delta(S)} : S \subseteq V\} \\ &\leq \max\left\{w^\top z : z \in \mathbb{R}_+^E, y \in \mathbb{R}_+^{\mathcal{P}(V)}, \mathbb{1}^\top y \leq 1, z \leq \sum_{S \subseteq V} y_S \mathbb{1}_{\delta(S)}\right\} \quad \text{take } (z, y) := (\mathbb{1}_{\delta(S)}, e_S) \\ &= \max\{w^\top z : z \in \mathbb{R}_+^E, \text{fcc}(G, z) \leq 1\} \quad \text{by (1)} \\ &\leq \text{mc}(G, w) \quad \text{by (32c).} \end{aligned}$$

The proof of (33) follows a similar structure. For every $w \in \mathbb{R}_+^E$, equation (33b) follows from SDP Strong Duality via (4) and (26b). The Cauchy-Schwarz inequality in (33c) then follows from (33b). Finally, (33a) follows, since

$$\begin{aligned} \eta(G, w) &= \max\{w^\top (\tfrac{1}{4}\mathcal{L}_G^*(Y)) : Y \in \mathbb{S}_+^V, \text{diag}(Y) = \mathbb{1}\} \quad \text{by (4a)} \\ &\leq \max\{w^\top z : z \in \mathbb{R}_+^E, z \leq \tfrac{1}{4}\mathcal{L}_G^*(Y), Y \in \mathbb{S}_+^V, \text{diag}(Y) = \mathbb{1}\} \end{aligned}$$

$$\begin{aligned}
&= \max\{w^\top z : z \in \mathbb{R}_+^E, \eta^\circ(G, z) \leq 1\} && \text{by (11)} \\
&\leq \eta(G, w) && \text{by (33c).} \quad \square
\end{aligned}$$

In the following discussion, we will elaborate on the above concepts along with their associated implications. Let E be a finite set, and let $\phi: \mathbb{R}_+^E \rightarrow \mathbb{R}_+$ be a function such that $\phi(0) = 0$. We say that

- ϕ is *positive definite* if $\phi(x) > 0$ for every nonzero $x \in \mathbb{R}_+^E$;
- ϕ is *monotone* whenever $x \leq y$ implies $\phi(x) \leq \phi(y)$ for every $x, y \in \mathbb{R}_+^E$;
- ϕ is *positively homogeneous* if $\phi(\alpha x) = \alpha \phi(x)$ for every $x \in \mathbb{R}_+^E$ and $\alpha \in \mathbb{R}_{++}$;
- ϕ is a *gauge* if it is convex and positively homogeneous.

If $\phi: \mathbb{R}_+^E \rightarrow \mathbb{R}_+$ is a positive definite monotone gauge, we define its *dual* $\phi^\circ: \mathbb{R}_+^E \rightarrow \mathbb{R}_+$ by

$$(34) \quad \phi^\circ(z) := \max\{z^\top w : w \in \mathbb{R}_+^E, \phi(w) \leq 1\} \quad \text{for every } z \in \mathbb{R}_+^E.$$

It is routine to check that ϕ° is a positive definite monotone gauge. One can also exploit a hyperplane separation theorem to show that $\phi^{\circ\circ} = \phi$.

Corollary 8. Let $G = (V, E)$ be a graph. Then the functions $\text{mc}(G, \cdot)$ and $\eta(G, \cdot)$ are positive definite monotone gauges, and their duals are $\text{fcc}(G, \cdot)$ and $\eta^\circ(G, \cdot)$, respectively.

Proof. This follows directly from Proposition 7. \square

We have already exploited the properties of positive definite monotone gauges throughout our work: recall (23) and (24), for example, which state monotonicity and convexity of $\eta^\circ(G, \cdot)$. More importantly, gauge duality immediately establishes a *bound conversion* procedure [6, Section 6] on which this work is based. Recall (5) and (13):

$$(5) \quad \alpha_{\text{GW}} \eta(G, w) \leq \text{mc}(G, w) \leq \eta(G, w) \quad \text{for each } w \in \mathbb{R}_+^E,$$

$$(13) \quad \eta^\circ(G, z) \leq \text{fcc}(G, z) \leq \frac{1}{\alpha_{\text{GW}}} \eta^\circ(G, z) \quad \text{for each } z \in \mathbb{R}_+^E.$$

The relationship between these inequalities are instances of the following result.

Proposition 9. Let E be a finite set, and let $\phi, \psi: \mathbb{R}_+^E \rightarrow \mathbb{R}_+$ be positive definite monotone gauges. Let $\alpha, \beta \in \mathbb{R}_{++}$. Then the following are equivalent:

$$\alpha \psi(w) \leq \phi(w) \leq \beta \psi(w) \quad \text{for every } w \in \mathbb{R}_+^E;$$

$$\frac{1}{\beta} \psi^\circ(z) \leq \phi^\circ(z) \leq \frac{1}{\alpha} \psi^\circ(z) \quad \text{for every } z \in \mathbb{R}_+^E.$$

Proof. Since $\phi^\circ, \psi^\circ: \mathbb{R}_+^E \rightarrow \mathbb{R}_+$ are positive definite monotone gauges and $\phi^{\circ\circ} = \phi$ and $\psi^{\circ\circ} = \psi$, it suffices to show that

$$(35) \quad \text{if } \phi(w) \leq \beta \psi(w) \text{ for every } w \in \mathbb{R}_+^E, \text{ then } \frac{1}{\beta} \psi^\circ(z) \leq \phi^\circ(z) \text{ for every } z \in \mathbb{R}_+^E.$$

This follows from the fact that, for every $z \in \mathbb{R}_+^E$,

$$\phi^\circ(z) = \max\{z^\top w : w \in \mathbb{R}_+^E, \phi(w) \leq 1\} \text{ and } \frac{1}{\beta} \psi^\circ(z) = \max\{z^\top w : w \in \mathbb{R}_+^E, \beta \psi(w) \leq 1\}. \quad \square$$

From a geometric viewpoint, positive definite monotone gauges are deeply related to convex corners, which are the underlying convex sets of many classical optimization problems. Let E be a finite set. The *lower-comprehensive hull* of $\mathcal{C} \subseteq \mathbb{R}_+^E$ is defined by $\text{lc}(\mathcal{C}) := \{z \in \mathbb{R}_+^E : \exists x \in \mathcal{C}, z \leq x\}$, and \mathcal{C} is *lower comprehensive* if $\text{lc}(\mathcal{C}) = \mathcal{C}$. A *convex corner* is a lower-comprehensive compact convex set $\mathcal{C} \subseteq \mathbb{R}_+^E$ with nonempty interior.

Every positive definite monotone gauge $\phi: \mathbb{R}_+^E \rightarrow \mathbb{R}_+$ is associated to a convex corner

$$\mathcal{C}_\phi := \{x \in \mathbb{R}_+^E : \phi^\circ(x) \leq 1\}$$

which satisfies

$$\phi(w) = \max\{w^\top x : x \in \mathcal{C}_\phi\} \quad \text{for each } w \in \mathbb{R}_+^E.$$

The *antiblocker* of $\mathcal{C} \subseteq \mathbb{R}_+^E$ is defined by $\text{abl}(\mathcal{C}) := \{y \in \mathbb{R}_+^E : \forall x \in \mathcal{C}, y^\top x \leq 1\}$. From these definitions and the fact that $\phi^{\circ\circ} = \phi$, it is clear that $\mathcal{C}_{\phi^\circ} = \text{abl}(\mathcal{C}_\phi)$. This correspondence allows one to recast gauge duality results in terms of *antiblocking duality* [16, 17].

Let $G = (V, E)$ be a graph. Define the *cut polytope* of G as

$$(36a) \quad \text{CUT}(G) := \text{conv}\{\mathbb{1}_{\delta(S)} : S \subseteq V\} \subseteq \mathbb{R}^E,$$

where $\text{conv}(\cdot)$ denotes the convex hull, and define its semidefinite relaxation as

$$(36b) \quad \text{CUT}_{\text{SDP}}(G) := \{\tfrac{1}{4}\mathcal{L}_G^*(Y) : Y \in \mathcal{E}^V\} \subseteq \mathbb{R}^E.$$

The convex corners associated to $\text{mc}(G, \cdot)$, $\text{fcc}(G, \cdot)$, $\eta(G, \cdot)$, and $\eta^\circ(G, \cdot)$ are the following:

$$(37) \quad \mathcal{C}_{\text{mc}(G, \cdot)} = \text{lc}(\text{CUT}(G)) \quad \text{and} \quad \mathcal{C}_{\text{fcc}(G, \cdot)} = \text{abl}(\text{CUT}(G)),$$

$$(38) \quad \mathcal{C}_{\eta(G, \cdot)} = \text{lc}(\text{CUT}_{\text{SDP}}(G)) \quad \text{and} \quad \mathcal{C}_{\eta^\circ(G, \cdot)} = \text{abl}(\text{CUT}_{\text{SDP}}(G)).$$

Basic convex analysis shows that studying $\text{fcc}(G, \cdot)$ for *all* weights $z \in \mathbb{R}_+^E$ corresponds to studying the *whole* boundary structure of $\mathcal{C}_{\text{fcc}(G, \cdot)}$, not just in the direction $z = \mathbb{1}$. The set $\text{lc}(\text{CUT}(G))$ above has appeared previously in the literature as the bipartite subgraph polytope of G ; see, e.g., [22].

Moreover, the inequalities in (5) and (13) can be interpreted as the following set inclusions:

$$(39a) \quad \alpha_{\text{GW}} \text{lc}(\text{CUT}_{\text{SDP}}(G)) \subseteq \text{lc}(\text{CUT}(G)) \subseteq \text{lc}(\text{CUT}_{\text{SDP}}(G)),$$

$$(39b) \quad \text{abl}(\text{CUT}_{\text{SDP}}(G)) \subseteq \text{abl}(\text{CUT}(G)) \subseteq \frac{1}{\alpha_{\text{GW}}} \text{abl}(\text{CUT}_{\text{SDP}}(G)).$$

We refer the reader to [6, Sections 2–7] or [5, Sections 4.1–4.3] for an in-depth discussion about gauge duality, with elementary proofs of the aforementioned results. One may regard gauge duality as a manifestation of convex duality. For example, Freund [14] formulates pairs of primal and dual gauge optimization problems, and proves a strong duality result using a hyperplane separation theorem. In this form, gauge duality has received a lot of attention in the optimization community recently; see [2, 15]. We remark that the work of Grötschel, Lovász, and Schrijver [21, Corollary 3.5], together with the remark that (4a) defines a positive definite monotone gauge, immediately implies that one can approximate the optimal value of (11) to any given precision in polynomial time. The algorithm in Section 2 refines this by showing that, beyond the polynomial-time computable lower bound to the value of the fractional cut-covering number, one has a suitable approximation algorithm leveraging the work of Goemans and Williamson [19] that actually constructs an approximately optimal fractional cut cover.

Remark 10. The work [21] is relevant to our discussion for another reason. Let $\beta \in (\alpha_{\text{GW}}, 1)$. Assume that for every graph G , there exists a positive definite monotone gauge $\psi_G : \mathbb{R}_+^E \rightarrow \mathbb{R}_+$ such that

- (1) there exists a polynomial time algorithm that takes as input an instance (G, z) of the fractional cut-covering problem and $\varepsilon > 0$, and outputs $q \in \mathbb{Q}$ such that $|q - \psi_G(z)| < \varepsilon$; and
- (2) $\psi_G(z) \leq \text{fcc}(G, z) \leq \frac{1}{\beta} \psi_G(z)$ for every $z \in \mathbb{R}_+^E$.

We argue that, under these assumptions, the Unique Games Conjecture [27] is false. Indeed, Grötschel, Lovász, and Schrijver [21] showed that under the first hypothesis, one can approximate the number $\psi_G^\circ(w) \in \mathbb{R}$ for every input $w \in \mathbb{Q}_+^E$ in polynomial time. Proposition 9 and item (2) above imply that $\beta \psi_G^\circ(w) \leq \text{mc}(G, w) \leq \psi_G^\circ(w)$. However, by [27, Theorem 1], this contradicts the Unique Games Conjecture.

3.2. β -pairings. Let $G = (V, E)$ be a graph. For every vector $w \in \mathbb{R}_+^E$, one has an instance (G, w) of the maximum cut problem. Similarly, for every vector $z \in \mathbb{R}_+^E$, one has an instance (G, z) of the fractional cut-covering problem. From a computational complexity point of view, it is remarkable how (32c) relates the optimal values $\text{mc}(G, w)$ and $\text{fcc}(G, z)$. It is natural then to try to find pairs $(w, z) \in \mathbb{R}_+^E \times \mathbb{R}_+^E$ which mutually certify the optimality of each other. That such pairs exist is a consequence of (32). As we are interested in approximation algorithms, we then parameterize this relationship between instances by a real number $\beta \in (0, 1]$, interpreted as an approximation factor.

Definition 11. Let $G = (V, E)$ be a graph and let $\beta \in (0, 1]$. A β -pairing on G is a pair $(w, z) \in \mathbb{R}_+^E \times \mathbb{R}_+^E$ such that there exist $\rho, \mu \in \mathbb{R}_+$, such that $\rho = 0 = \mu$ if and only if $w = 0 = z$, and

$$(40) \quad w^\top z \stackrel{(40a)}{=} \rho\mu \quad \text{and} \quad \beta\rho\mu \stackrel{(40b)}{\leq} \text{mc}(G, w)\mu \stackrel{(40c)}{\leq} \rho\mu \stackrel{(40d)}{\leq} \rho \text{fcc}(G, z) \stackrel{(40e)}{\leq} \frac{1}{\beta} \rho\mu.$$

We define an *exact pairing* on G to be a 1-pairing on G .

When $\rho > 0$ and $\mu > 0$, which we regard as the “typical” case, we may restate (40) as

$$w^\top z \stackrel{(40a)}{=} \rho\mu, \quad \beta\rho \stackrel{(40b)}{\leq} \text{mc}(G, w) \stackrel{(40c)}{\leq} \rho, \quad \text{and} \quad \mu \stackrel{(40d)}{\leq} \text{fcc}(G, z) \stackrel{(40e)}{\leq} \frac{1}{\beta}\mu.$$

The definition is made to accommodate the case $0 \in \{\rho, \mu\}$. The nonzero conditions on ρ and μ are meant only to avoid “spurious” β -pairings. In fact, one can easily check that, in Definition 11,

$$\rho = 0 \text{ if and only if } w = 0 \quad \text{and} \quad \mu = 0 \text{ if and only if } z = 0.$$

Let $G = (V, E)$ be a graph, and let $w, z \in \mathbb{R}_+^E$. If we take $\rho := \eta(G, w)$ and $\mu := \eta^\circ(G, z)$, Proposition 9 states that (40b) and (40c) are the “dual inequalities” to (40e) and (40d), respectively. Definition 11 abstracts some of the concepts from positive definite monotone gauges mentioned in Section 3.1, while preserving the relevant duality. Definition 11 also shifts the focus to the relation established between w and z , which is crucial to our certification approach. Note that for an exact pairing (w, z) , the above solutions are optimal with a precise relationship:

$$w^\top z = \text{mc}(G, w) \text{fcc}(G, z), \quad \text{mc}(G, w) = \rho, \quad \text{and} \quad \text{fcc}(G, z) = \mu.$$

While the definition of a β -pairing captures the notion of simultaneous β -approximations of the numbers $\text{mc}(G, w)$ and $\text{fcc}(G, z)$, it is important to consider which objects might *certify* that a pair $(w, z) \in \mathbb{R}_+^E \times \mathbb{R}_+^E$ is indeed a β -pairing for a fixed $\beta \in (0, 1]$. The lower bound (40b) on the maximum cut value and the upper bound (40e) on the fractional cut-covering value in (40) can be naturally certified by (the shore of) a cut and a fractional cut cover, respectively. However, certifying the upper bound (40c) on $\text{mc}(G, w)$ and the lower bound (40d) on $\text{fcc}(G, z)$, i.e., $\text{mc}(G, w)\mu \leq \rho\mu \leq \rho \text{fcc}(G, z)$, poses a more complex question. We achieve this certification by using semidefinite programming weak duality. Since $\text{mc}(G, w) \leq \eta(G, w)$, we have that

$$(41) \quad \text{mc}(G, w) \leq \rho \text{ for each feasible solution } (\rho, x) \text{ of (4b).}$$

From the viewpoint of $\eta(G, w)$ as a semidefinite relaxation for $\text{mc}(G, w)$, it is very natural to regard x as (the key part of) a feasible solution for its dual SDP (4b). However, the vector $x \in \mathbb{R}^V$ also has a combinatorial interpretation. Direct computation using (12) shows that

$$(42a) \quad \mathcal{L}_G^*(\mathbb{1}_S \mathbb{1}_S^\top) = \mathbb{1}_{\delta(S)} \quad \text{for each } S \subseteq V,$$

$$(42b) \quad \mathcal{L}_G^*(\mathbb{1} h^\top) = \mathcal{L}_G^*(h \mathbb{1}^\top) = 0 \quad \text{for each } h \in \mathbb{R}^V.$$

Thus, the inequality $\frac{1}{4}\mathcal{L}_G(w) \preceq \text{Diag}(x)$ from (4b) implies the middle inequality in:

$$(43) \quad w^\top \mathbb{1}_{\delta(S)} = \langle \frac{1}{4}\mathcal{L}_G(w), (\mathbb{1} - 2\mathbb{1}_S)(\mathbb{1} - 2\mathbb{1}_S)^\top \rangle \leq \langle \text{Diag}(x), (\mathbb{1} - 2\mathbb{1}_S)(\mathbb{1} - 2\mathbb{1}_S)^\top \rangle = \mathbb{1}^\top x \leq \rho.$$

This shows that, while the semidefinite inequality $\frac{1}{4}\mathcal{L}_G(w) \preceq \text{Diag}(x)$ just used may look at a first glance not quite combinatorial, the only property used in the proof of (43) is the more combinatorial-looking

$$(44) \quad \langle \frac{1}{4}\mathcal{L}_G(w), hh^\top \rangle \leq \langle \text{Diag}(x), hh^\top \rangle \quad \text{for each } h \in \{\pm 1\}^V.$$

Hence, any certificate for the inequality $\frac{1}{4}\mathcal{L}_G(w) \preceq \text{Diag}(x)$ certifies the inequality in (44) for arbitrary $h \in \mathbb{R}^V$, and in particular for each $h \in \{\pm 1\}^V$, i.e., for each $h \in \mathbb{R}^V$ of the form $\mathbb{1} - 2\mathbb{1}_S$ for some $S \subseteq V$. Certifying a richer family of inequalities can be seen as dual to solving a relaxation of $\text{mc}(G, w)$. Finally, the semidefinite inequality $\frac{1}{4}\mathcal{L}_G(w) \preceq \text{Diag}(x)$ can be certified by providing an LDL^\top factorization (that is, a square-root-free Cholesky decomposition) of the positive semidefinite matrix $\text{Diag}(x) - \frac{1}{4}\mathcal{L}_G(w)$. With (43) we have showed how to certify $\text{mc}(G, w) \leq \rho$; it remains to discuss certification of $\text{fcc}(G, z) \geq \mu$. However, as we shall prove in the upcoming results, the latter inequality can be certified by the *same* objects that certify the former inequality for appropriately paired edge weights w and z . Such a *simultaneous certification* is a key aspect of our work. We have now gathered all the ingredients we need to define certificates for β -pairings.

Definition 12. Let $G = (V, E)$ be a graph and let $\beta \in (0, 1]$. Let $(w, z) \in \mathbb{R}_+^E \times \mathbb{R}_+^E$. A β -certificate for (w, z) is a tuple (ρ, μ, S, y, x) such that $\rho = 0 = \mu$ if and only if $w = 0 = z$, and

$$(45.i) \quad \rho, \mu \in \mathbb{R}_+ \text{ are such that } \rho\mu = w^\top z,$$

$$(45.ii) \quad S \subseteq V \text{ is such that } w^\top \mathbb{1}_{\delta(S)} \geq \beta\rho,$$

$$(45.iii) \quad y \in \mathbb{R}_+^{\mathcal{P}(V)} \text{ is such that } \sum_{U \subseteq V} y_U \mathbb{1}_{\delta(U)} \geq z \text{ and } \mathbb{1}^\top y \leq \frac{1}{\beta}\mu, \text{ and}$$

$$(45.iv) \quad x \in \mathbb{R}^V \text{ is such that } \rho \geq \mathbb{1}^\top x \text{ and } \text{Diag}(x) \succeq \frac{1}{4}\mathcal{L}_G(w).$$

Remark 13. Recalling the discussion preceding Definition 11, note how items (45.i), (45.ii), and (45.iii) are the natural certificates for the inequalities (40a), (40b), and (40e), respectively. Whereas we work with an SDP certificate in (45.iv), this setup opens the possibility of using other techniques which upper bound the maximum value of the weighted maximum cut problem. Concretely, one could substitute (45.iv) by appropriate certificates arising from [23, 41], for example.

Next, we prove that a β -certificate for (w, z) does indeed certify that (w, z) is a β -pairing.

Proposition 14. Let $G = (V, E)$ be a graph and let $\beta \in (0, 1]$. Let $(w, z) \in \mathbb{R}_+^E \times \mathbb{R}_+^E$. If there exists a β -certificate for (w, z) , then (w, z) is a β -pairing.

Proof. Let (ρ, μ, S, y, x) be a β -certificate for (w, z) . Item (45.i) proves (40a). One has

$$\beta \rho \mu \leq \mu w^\top \mathbb{1}_{\delta(S)} \leq \mu \text{mc}(G, w)$$

by (45.ii). This proves (40b). Similarly,

$$\rho \text{fcc}(G, z) \leq \rho \mathbb{1}^\top y \leq \frac{1}{\beta} \rho \mu$$

by (45.iii). This proves (40e).

So far we have only used feasible solutions for $\text{mc}(G, w)$ and $\text{fcc}(G, z)$ to obtain bounds for the optimal values. Finally, (45.iv) shows that (41) applies, so

$$\text{mc}(G, w) \mu \stackrel{(41)}{\leq} \rho \mu \stackrel{(45.i)}{=} w^\top z \stackrel{(32c)}{\leq} \text{mc}(G, w) \text{fcc}(G, z) \stackrel{(41)}{\leq} \rho \text{fcc}(G, z),$$

thus proving (40c) and (40d). \square

3.3. Existence of α_{GW} -Certificates. Having motivated the definition of β -certificates as objects simultaneously proving approximate optimality for the maximum cut and fractional cut-covering problems, a next step would be to determine conditions on a β -pairing $(w, z) \in \mathbb{R}_+^E \times \mathbb{R}_+^E$ that guarantee the existence of a β -certificate for (w, z) . Set

$$(46) \quad \text{H}(G) := \{ (w, z) \in \mathbb{R}_+^E \times \mathbb{R}_+^E : w^\top z = \eta(G, w) \eta^\circ(G, z) \}.$$

We now expand (46) into a more convenient characterization. Recalling (4b) and (11),

$$(4b) \quad \eta(G, w) = \min \{ \rho : \rho \in \mathbb{R}_+, x \in \mathbb{R}^V, \rho \geq \mathbb{1}^\top x, \text{Diag}(x) \succeq \frac{1}{4} \mathcal{L}_G(w) \},$$

$$(11) \quad \eta^\circ(G, z) = \min \{ \mu : \mu \in \mathbb{R}_+, Y \in \mathbb{S}_+^V, \text{diag}(Y) = \mu \mathbb{1}, \frac{1}{4} \mathcal{L}_G^*(Y) \succeq z \},$$

we claim that

$$(47) \quad \text{H}(G) = \left\{ (w, z) \in \mathbb{R}_+^E \times \mathbb{R}_+^E : \begin{array}{l} \exists (\rho, x) \text{ feasible for (4b) for } (G, w), \\ \exists (\mu, Y) \text{ feasible for (11) for } (G, z), \\ \text{and } w^\top z \geq \rho \mu \end{array} \right\}.$$

For each $(w, z) \in \text{H}(G)$, we shall refer to pairs (ρ, x) and (μ, Y) assumed to exist as in the RHS of (47), as *witnesses of the membership* $(w, z) \in \text{H}(G)$. To prove ‘ \subseteq ’ in (47), it suffices to choose as witnesses an optimal solution (ρ, x) for (4b) and an optimal solution (μ, Y) for (11). Next we prove ‘ \supseteq ’. Note that,

$$(48) \quad \text{if } (\rho, x) \text{ and } (\mu, Y) \text{ witness the membership } (w, z) \in \text{H}(G), \text{ then } \rho \mu = w^\top z = \langle \frac{1}{4} \mathcal{L}_G(w), Y \rangle,$$

since $\rho \mu \leq w^\top z \leq w^\top (\frac{1}{4} \mathcal{L}_G^*(Y)) = \langle \frac{1}{4} \mathcal{L}_G(w), Y \rangle \leq \langle \text{Diag}(x), Y \rangle = \mu \mathbb{1}^\top x \leq \rho \mu$, so equality holds throughout. In particular, we further have that $\eta(G, w) \leq \rho$ and $\eta^\circ(G, z) \leq \mu$, so equality holds in both cases by (33c). Thus, ‘ \supseteq ’ is proved in (47). To see the connection between (47) and Definition 12, note that in the Goemans-Williamson approximation algorithm for the maximum cut problem and in our approximation algorithm for the fractional cut-covering problem, a crucial step is obtaining a matrix $Y \in \mathbb{S}_+^V$ that is feasible in the semidefinite relaxation, which is then used in the sampling of shores. In this way, the matrix Y encodes both the shore (45.ii) and the fractional cut cover (45.iii) featured in β -certificates.

In this subsection, we present two main results. The first result (stated in Proposition 15) is that there is an α_{GW} -certificate for every nonzero $(w, z) \in \text{H}(G)$; in particular, by Proposition 14, every nonzero $(w, z) \in \text{H}(G)$ is an α_{GW} -pairing. The second result (stated in Proposition 16) shows that, given an instance (G, w) of the maximum cut problem, we have that $(w, z) \in \text{H}(G)$ for $z \in \mathbb{R}_+^E$ if and only if z is in the convex cone generated by the optimal solutions of a certain formulation for $\eta(G, w)$. Symmetrically, given an instance (G, z) of the

fractional cut-covering problem, we have that $(w, z) \in H(G)$ for $w \in \mathbb{R}_+^E$ if and only if w is in the convex cone generated by the optimal solutions of a certain formulation for $\eta^\circ(G, z)$.

Let $G = (V, E)$ be a graph and let $z \in \mathbb{R}_+^E$. Goemans and Williamson's analysis [19] implies that

$$(49) \quad \text{if } (\mu, Y) \text{ is feasible for (11) and } \mu > 0, \text{ then } \mathbb{E}(w^\top \mathbb{1}_{\delta(GW(Y))}) \geq \frac{\alpha_{\text{GW}}}{\mu} \langle \frac{1}{4} \mathcal{L}_G(w), Y \rangle.$$

Now we show the existence of an α_{GW} -certificate for every nonzero pair $(w, z) \in H(G)$.

Proposition 15. Let $G = (V, E)$ be a graph. For every $(w, z) \in H(G)$ such that $w \neq 0 \neq z$, there exists an α_{GW} -certificate (ρ, μ, S, y, x) for (w, z) such that $\rho \neq 0 \neq \mu$. In particular, (w, z) is an α_{GW} -pairing.

Proof. Let $(w, z) \in H(G)$ be such that $w \neq 0 \neq z$. Let (ρ, x) be an optimal solution for (4b) and let (μ, Y) be an optimal solution for (11), so that $\rho = \eta(G, w) > 0$ and $\mu = \eta^\circ(G, z) > 0$. Thus, $w^\top z = \eta(G, w)\eta^\circ(G, z) = \rho\mu$, so (45.i) holds. Equation (45.iv) holds by the feasibility of (ρ, x) for (4b). Define $y \in \mathbb{R}^{\mathcal{P}(V)}$ as in (16). Equation (45.iii) follows from Proposition 1. Since

$$\mathbb{E}(w^\top \mathbb{1}_{\delta(GW(Y))}) \geq \frac{\alpha_{\text{GW}}}{\mu} \rho\mu = \alpha_{\text{GW}}\rho \geq \frac{\alpha_{\text{GW}}}{\mu} \langle \frac{1}{4} \mathcal{L}_G(w), Y \rangle \geq \frac{\alpha_{\text{GW}}}{\mu} \langle \text{Diag}(x), Y \rangle = \alpha_{\text{GW}} \mathbb{1}^\top x = \alpha_{\text{GW}}\rho,$$

there exists $S \subseteq V$ as in (45.ii). Thus, (w, z) is an α_{GW} -pairing by Proposition 14. \square

The set $H(G)$ defines a relation between maximum cut and fractional cut-covering instances. Suppose the starting point is an instance of one of these problems and one builds an instance for the other one so that the pair is in $H(G)$. This motivates the following definitions. Define

$$(50) \quad \mathfrak{z}_G(w) := \{z \in \mathbb{R}_+^E : (w, z) \in H(G)\} \quad \text{for every } w \in \mathbb{R}_+^E,$$

$$(51) \quad \mathfrak{w}_G(z) := \{w \in \mathbb{R}_+^E : (w, z) \in H(G)\} \quad \text{for every } z \in \mathbb{R}_+^E.$$

The upcoming Proposition 16, which describes the encoding of the optimal solutions to our SDPs in $H(G)$, is more conveniently stated using a slight variant of (4):

$$(52a) \quad \eta(G, w) = \max\{w^\top z : z \in \mathbb{R}_+^E, Y \in \mathbb{S}_+^V, \text{diag}(Y) = \mathbb{1}, z \leq \frac{1}{4} \mathcal{L}_G^*(Y)\}$$

$$(52b) \quad = \min\{\rho : \rho \in \mathbb{R}, x \in \mathbb{R}^V, \text{Diag}(x) \succeq \frac{1}{4} \mathcal{L}_G(w), \rho \geq \mathbb{1}^\top x\}.$$

Recall (26a) and (26b):

$$(26a) \quad \eta^\circ(G, z) = \min\{\mu : \mu \in \mathbb{R}_+, Y \in \mathbb{S}_+^V, \text{diag}(Y) = \mu \mathbb{1}, \frac{1}{4} \mathcal{L}_G^*(Y) \geq z\}$$

$$(26b) \quad = \max\{z^\top w : w \in \mathbb{R}_+^E, x \in \mathbb{R}^V, \frac{1}{4} \mathcal{L}_G(w) \preceq \text{Diag}(x), \mathbb{1}^\top x \leq 1\}.$$

Note the various symmetries relating these SDPs. The constraint ' $\text{Diag}(x) \succeq \frac{1}{4} \mathcal{L}_G(w)$ ' occurs in (52b) and (26b) and ' $\frac{1}{4} \mathcal{L}_G^*(Y) \geq z$ ' occurs in (52a) and (26a), although in each constraint one of w or z is a variable in one but not in the other SDP. Additionally, the constraint ' $\text{diag}(Y) = \mathbb{1}$ ' from (52a) appears homogenized in (26a) with the variable μ , whereas the constraint ' $\mathbb{1}^\top x \leq 1$ ' from (26b) appears homogenized in (52b) with the variable ρ .

Let $G = (V, E)$ be a graph. We now relate \mathfrak{z}_G and \mathfrak{w}_G to optimal solutions of (52a) and of (26b), resp. For any finite set U and for any set $S \subseteq \mathbb{R}^U$, we denote by $\text{cone}(S)$ the convex cone generated by S , i.e., the smallest convex cone containing S and the origin.

Proposition 16. Let $G = (V, E)$ be a graph. For every $w \in \mathbb{R}_+^E$,

$$(53) \quad \mathfrak{z}_G(w) = \text{cone}(\{z \in \mathbb{R}_+^E : \exists Y \in \mathbb{S}_+^V \text{ s.t. } (z, Y) \text{ is optimal in (52a) for } \eta(G, w)\}).$$

Similarly, for every $z \in \mathbb{R}_+^E$,

$$(54) \quad \mathfrak{w}_G(z) = \text{cone}(\{w \in \mathbb{R}_+^E : \exists x \in \mathbb{R}^V \text{ s.t. } (w, x) \text{ is optimal in (26b) for } \eta^\circ(G, z)\}).$$

Proof. We first prove (53). Let $w \in \mathbb{R}_+^E$. Note that

$$(55) \quad \{z \in \mathbb{R}_+^E : w^\top z = \eta(G, w)\eta^\circ(G, z)\} = \text{cone}(\{z \in \mathbb{R}_+^E : \eta^\circ(G, z) \leq 1, w^\top z = \eta(G, w)\}).$$

Indeed, ' \subseteq ' holds in (55) since η° is positively homogeneous. For the proof of ' \supseteq ', first note that the LHS is clearly closed under positive scalar multiplication since $\eta^\circ(G, \cdot)$ is positively homogeneous. To see that the LHS is a convex cone, let z_1, z_2 be elements of the LHS. Sublinearity of $\eta^\circ(G, \cdot)$ implies that $w^\top(z_1 + z_2) = \eta(G, w)(\eta^\circ(G, z_1) + \eta^\circ(G, z_2)) \geq \eta(G, w)\eta^\circ(G, z_1 + z_2)$, whence equality holds by (33c), which

proves that $z_1 + z_2$ lies in the LHS. Since $w^\top z = \eta(G, w) \geq \eta(G, w)\eta^\circ(G, z) \geq w^\top z$ for every $z \in \mathbb{R}_+^E$ in the RHS of (55), by (33c), this concludes the proof of (55). We can now prove (53):

$$\begin{aligned}
\mathfrak{z}_G(w) &= \{z \in \mathbb{R}_+^E : (w, z) \in H(G)\} && \text{by (50)} \\
&= \{z \in \mathbb{R}_+^E : w^\top z = \eta(G, w)\eta^\circ(G, z)\} && \text{by (46)} \\
&= \text{cone}(\{z \in \mathbb{R}_+^E : \eta^\circ(G, z) \leq 1, w^\top z = \eta(G, w)\}) && \text{by (55)} \\
&= \text{cone}(\{z \in \mathbb{R}_+^E : \exists Y \in \mathbb{S}_+^V, \text{diag}(Y) = \mathbb{1}, \tfrac{1}{4}\mathcal{L}_G^*(Y) \geq z, w^\top z = \eta(G, w)\}) && \text{by (26a)} \\
&= \text{cone}(\{z \in \mathbb{R}_+^E : \exists Y \in \mathbb{S}_+^V \text{ s.t. } (z, Y) \text{ is optimal in (52a) for } \eta(G, w)\}).
\end{aligned}$$

For (54), let $z \in \mathbb{R}_+^E$. One can prove, analogously to (55), that

$$(56) \quad \{w \in \mathbb{R}_+^E : w^\top z = \eta(G, w)\eta^\circ(G, z)\} = \text{cone}(\{w \in \mathbb{R}_+^E : \eta(G, w) \leq 1, w^\top z = \eta^\circ(G, z)\}).$$

Then

$$\begin{aligned}
\mathfrak{w}_G(z) &= \{w \in \mathbb{R}_+^E : (w, z) \in H(G)\} && \text{by (51)} \\
&= \{w \in \mathbb{R}_+^E : w^\top z = \eta(G, w)\eta^\circ(G, z)\} && \text{by (46)} \\
&= \text{cone}(\{w \in \mathbb{R}_+^E : \eta(G, w) \leq 1, w^\top z = \eta^\circ(G, z)\}) && \text{by (56)} \\
&= \text{cone}(\{w \in \mathbb{R}_+^E : \exists x \in \mathbb{R}^V, \mathbb{1} \geq \mathbb{1}^\top x, \text{Diag}(x) \succeq \tfrac{1}{4}\mathcal{L}_G(w), w^\top z = \eta^\circ(G, z)\}) && \text{by (4b)} \\
&= \text{cone}(\{w \in \mathbb{R}_+^E : \exists x \in \mathbb{R}^V \text{ s.t. } (w, x) \text{ is optimal in (26b) for } \eta(G, w)\}). \quad \square
\end{aligned}$$

Remark 17. Let $G = (V, E)$ be a graph. Proposition 16 establishes a stronger result than nonemptiness of $H(G)$, by showing that the projections $\mathfrak{z}_G(w)$ and $\mathfrak{w}_G(z)$ of $H(G)$ are nontrivial convex cones for every $w, z \in \mathbb{R}_+^E$. These cones are tightly connected to normal cones of the relevant convex corners. Recall the sets defined in (38). Fix a nonzero $w \in \mathbb{R}_+^E$, and set $\rho := \eta(G, w) > 0$. Then $\eta(G, \rho^{-1}w) \leq 1$, so $\rho^{-1}w \in \mathcal{C}_{\eta^\circ(G, \cdot)} = \text{abl}(\mathcal{C}_{\eta(G, \cdot)})$. Then one can show that

$$\begin{aligned}
\mathfrak{z}_G(w) &= \{z \in \mathbb{R}_+^E : w^\top z = \eta(G, w)\eta^\circ(G, z)\} \\
&= \{z \in \mathbb{R}_+^E : (\rho^{-1}w)^\top z = \eta^\circ(G, z)\} \\
&= \{z \in \mathbb{R}_+^E : z \in \text{Normal}(\mathcal{C}_{\eta^\circ(G, \cdot)}, \rho^{-1}w)\},
\end{aligned}$$

where $\text{Normal}(S, \bar{x}) := \{c \in \mathbb{R}^E : \forall x \in S, c^\top \bar{x} \geq c^\top x\}$ denotes the normal cone of $S \subseteq \mathbb{R}^E$ at $\bar{x} \in S$. Thus

$$\mathfrak{z}_G(w) = \mathbb{R}_+^E \cap \text{Normal}(\mathcal{C}_{\eta^\circ(G, \cdot)}, \eta(G, w)^{-1}w) \quad \text{for every } w \in \mathbb{R}_+^E \setminus \{0\}.$$

Dually,

$$\mathfrak{w}_G(z) = \mathbb{R}_+^E \cap \text{Normal}(\mathcal{C}_{\eta(G, \cdot)}, \eta^\circ(G, z)^{-1}z) \quad \text{for every } z \in \mathbb{R}_+^E \setminus \{0\}.$$

By combining Propositions 15 and 16, we show how to obtain α_{GW} -certificates when the starting point is either a maximum cut instance or a fractional cut-covering instance. In Section 3.4, we present algorithmic versions of these results: from nearly optimal SDPs solutions and a randomized sampling procedure, we produce a heavy cut and a light fractional cut cover.

Theorem 18. Let $G = (V, E)$ be a graph and let $w \in \mathbb{R}_+^E$ be nonzero. Then there exist a nonzero $z \in \mathfrak{z}_G(w)$ and an α_{GW} -certificate for (w, z) .

Proof. There exists an optimal solution (\bar{z}, \bar{Y}) for (52a). Thus, $(w, \bar{z}) \in H(G)$ by Proposition 16. As $w \neq 0$, we have that $\bar{z} \neq 0$. By Proposition 15, there exists an α_{GW} -certificate for (w, \bar{z}) . \square

Theorem 19. Let $G = (V, E)$ be a graph and let $z \in \mathbb{R}_+^E$ be nonzero. Then there exist a nonzero $w \in \mathfrak{w}_G(z)$ and an α_{GW} -certificate for (w, z) .

Proof. There exists an optimal solution (\bar{w}, \bar{x}) for (26b). Thus, $(\bar{w}, z) \in H(G)$ by Proposition 16. As $z \neq 0$, we have that $\bar{w} \neq 0$. By Proposition 15, there exists an α_{GW} -certificate for (\bar{w}, z) . \square

3.4. Algorithmically Obtaining β -Certificates. Theorems 18 and 19 are pleasantly symmetric, both in statement and in proof. However, they are not directly suitable for algorithmic use. The first issue is in the definition of $H(G)$ itself: as Proposition 16 states, its elements arise as optimal solutions to SDPs, whereas in algorithms one must work with nearly optimal solutions. The second issue is that, in Proposition 15, the cut $\delta(S)$ and the fractional cut cover y obtained have significant caveats: the shore S is not explicitly constructed and y may have exponential support size. In particular, the same issues arising at Section 2 resurface here again, requiring one to “thicken” edges and to settle for sparse surrogates of the probability distribution described in Proposition 1.

Rather than parameterizing our rounding procedure as we have done in Section 2, we parameterize *our optimization problems*, thus obtaining a family of geometric objects to study. Let $G = (V, E)$ be a graph and let $\varepsilon \in [0, 1)$. For each $w \in \mathbb{R}_+^E$, define

$$(57a) \quad \eta_\varepsilon(G, w) := (1 - \varepsilon)\eta(G, w) + \frac{\varepsilon}{2}\|w\|_1$$

$$(57b) \quad = \min\left\{\rho : \rho \in \mathbb{R}_+, x \in \mathbb{R}^V, \rho \geq (1 - \varepsilon)\mathbb{1}^\top x + \frac{\varepsilon}{2}\mathbb{1}^\top w, \text{Diag}(x) \succeq \frac{1}{4}\mathcal{L}_G(w)\right\}.$$

Then $\eta_\varepsilon(G, \cdot) : \mathbb{R}_+^E \rightarrow \mathbb{R}_+$ is a positive definite monotone gauge, and by (34) its gauge dual can be written as

$$(58a) \quad \eta_\varepsilon^\circ(G, z) = \max\left\{z^\top w : x \in \mathbb{R}^V, w \in \mathbb{R}_+^E, \frac{1}{4}\mathcal{L}_G(w) \preceq \text{Diag}(x), (1 - \varepsilon)\mathbb{1}^\top x + \frac{\varepsilon}{2}\mathbb{1}^\top w \leq 1\right\}$$

$$(58b) \quad = \min\left\{\mu : \mu \in \mathbb{R}_+, Y \in \mathbb{S}^V, Y \succeq \varepsilon\mu I, \frac{1}{4}\mathcal{L}_G^*(Y) \geq z, \text{diag}(Y) = \mu\mathbb{1}\right\}$$

for every $z \in \mathbb{R}_+^E$. The equality in (58b) follows from SDP Strong Duality, as the relaxed Slater points we exhibited in (26) remain relaxed Slater points in (58). Let $\sigma \in (0, 1)$ and set

$$(59) \quad H_{\varepsilon, \sigma}(G) := \left\{ \begin{array}{l} \exists(\rho, x) \text{ feasible for (57b) for } (G, w), \\ (w, z) \in \mathbb{R}_+^E \times \mathbb{R}_+^E : \exists(\mu, Y) \text{ feasible for (58b) for } (G, z), \\ \text{and } w^\top z \geq (1 - \sigma)\rho\mu \end{array} \right\}.$$

As for $H(G)$, for each $(w, z) \in H_{\varepsilon, \sigma}(G)$, we refer to pairs (ρ, x) and (μ, Y) assumed to exist as in the RHS of (59), as *witnesses of the membership* $(w, z) \in H_{\varepsilon, \sigma}(G)$. Note the following approximate version of (48):

$$(60) \quad \text{if } (\rho, x) \text{ and } (\mu, Y) \text{ witness the membership } (w, z) \in H_{\varepsilon, \sigma}(G), \text{ then } (1 - \sigma)\rho\mu \leq w^\top z \leq \rho\mu.$$

This holds since

$$\begin{aligned} w^\top z &\leq w^\top \left(\frac{1}{4}\mathcal{L}_G^*(Y)\right) = \langle \frac{1}{4}\mathcal{L}_G(w), Y - \varepsilon\mu I \rangle + \langle \frac{1}{4}\mathcal{L}_G(w), \varepsilon\mu I \rangle \\ &\leq \langle \text{Diag}(x), Y - \varepsilon\mu I \rangle + \frac{\varepsilon\mu}{2}\mathbb{1}^\top w = ((1 - \varepsilon)\mathbb{1}^\top x + \frac{\varepsilon}{2}\mathbb{1}^\top w)\mu \leq \rho\mu. \end{aligned}$$

The analogue of the expression we took as definition of $H(G)$ in (46) is

$$H_{\varepsilon, \sigma}(G) = \left\{ (w, z) \in \mathbb{R}_+^E \times \mathbb{R}_+^E : \left| \frac{\eta_\varepsilon(G, w)\eta_\varepsilon^\circ(G, z) - w^\top z}{\eta_\varepsilon(G, w)\eta_\varepsilon^\circ(G, z)} \right| \leq \sigma \right\},$$

where the expression inside the absolute value is taken to be zero whenever $0 \in \{w, z\}$.

Let $G = (V, E)$ be a graph and let $\varepsilon \in [0, 1)$. We show in Theorem 20 below that $\eta_\varepsilon(G, w)$ and $\eta_\varepsilon^\circ(G, z)$ are approximations for $\eta(G, w)$ and $\eta^\circ(G, z)$, respectively. Before that, we state two important monotonicity properties of the Laplacian of a graph G :

$$(61a) \quad v \leq w \text{ implies } \mathcal{L}_G(v) \preceq \mathcal{L}_G(w), \quad \text{for every } v, w \in \mathbb{R}^E,$$

$$(61b) \quad A \preceq B \text{ implies } \mathcal{L}_G^*(A) \leq \mathcal{L}_G^*(B), \quad \text{for every } A, B \in \mathbb{S}^V.$$

Both follow from the fact that $\mathcal{L}_G(e_{ij}) \succeq 0$ for every $ij \in E$. This is immediate for (61a) by using the definition of \mathcal{L}_G . For (61b), one has

$$(\mathcal{L}_G^*(A))_{ij} = \langle \mathcal{L}_G^*(A), e_{ij} \rangle = \langle A, \mathcal{L}_G(e_{ij}) \rangle \leq \langle B, \mathcal{L}_G(e_{ij}) \rangle = \langle \mathcal{L}_G^*(B), e_{ij} \rangle = (\mathcal{L}_G^*(B))_{ij}$$

for every $A, B \in \mathbb{S}^V$ such that $A \preceq B$, and for every $ij \in E$. These results imply one of the motivating properties of η_ε° :

$$(62) \quad \text{if } (\mu, Y) \text{ is feasible for (58b) and } \mu > 0, \text{ then } \mathbb{P}(ij \in \text{GW}(Y)) \geq \frac{\sqrt{2\varepsilon}}{\pi} \text{ for every } ij \in E.$$

Let (μ, Y) be feasible in (58b), and let $ij \in E$. Then $\mu - Y_{ij} = \frac{1}{2}\mathcal{L}_G^*(Y)_{ij} \geq \frac{1}{2}\mathcal{L}_G^*(\varepsilon\mu I)_{ij} = \varepsilon\mu$ by (61b). Thus $\mu(1 - \varepsilon) \geq Y_{ij}$. Hence (15) and (20) imply that

$$\mathbb{P}(ij \in \text{GW}(Y)) = \frac{\arccos(\mu^{-1}Y_{ij})}{\pi} \geq \frac{\sqrt{2\varepsilon}}{\pi},$$

so (62) holds.

Theorem 20. Let $G = (V, E)$ be a graph, and let $\varepsilon \in [0, 1)$. Then, for each $w, z \in \mathbb{R}_+^E$,

$$(63a) \quad (1 - \varepsilon)\eta(G, w) \leq \eta_\varepsilon(G, w) \leq \eta(G, w), \quad \text{for each } w \in \mathbb{R}_+^E,$$

$$(63b) \quad \eta^\circ(G, z) \leq \eta_\varepsilon^\circ(G, z) \leq \frac{1}{1 - \varepsilon}\eta^\circ(G, z), \quad \text{for each } z \in \mathbb{R}_+^E.$$

Proof. As the identity matrix is feasible in (4a), the rightmost inequality below holds:

$$(1 - \varepsilon)\eta(G, w) \leq (1 - \varepsilon)\eta(G, w) + \frac{\varepsilon}{2}\|w\|_1 = (1 - \varepsilon)\eta(G, w) + \varepsilon\langle \frac{1}{4}\mathcal{L}_G(w), I \rangle \leq \eta(G, w).$$

Thus (63a) follows from (57a). Now (63b) follows from (63a) and Proposition 9. Alternatively, one can prove (63b) directly. As the feasible region of (11) contains the feasible region of (58b), the first inequality in (63b) holds. Let (μ, Y) be feasible for (11). Set $\mu_\varepsilon := \mu/(1 - \varepsilon)$ and $Y_\varepsilon := Y + \varepsilon\mu_\varepsilon I \succeq \varepsilon\mu_\varepsilon I$. Then $\text{diag}(Y_\varepsilon) = \mu\mathbb{1} + \varepsilon\mu_\varepsilon\mathbb{1} = \mu_\varepsilon\mathbb{1}$. Moreover, from (61b) and $Y_\varepsilon \succeq Y$ we obtain $\frac{1}{4}\mathcal{L}_G^*(Y_\varepsilon) \geq \frac{1}{4}\mathcal{L}_G^*(Y) \geq z$. Thus $(\mu_\varepsilon, Y_\varepsilon)$ is feasible for (58b) with objective value μ_ε . Hence (63b) holds. \square

Let $T \in \mathbb{N} \setminus \{0\}$ and let $\gamma \in (0, 1)$. We use the shore sampling procedure (18) defined in Section 2. Let $G = (V, E)$ be a graph, let $\mu \in \mathbb{R}_+$, and let $Y \in \mu\mathcal{E}_V$. Let $S_1, \dots, S_T \subseteq V$ be independent identically-distributed random shores sampled by $\text{GW}(Y)$. Recall the definition in (18):

$$(18) \quad \mathcal{A}_{T,\gamma}(G, Y) = (\mathcal{F}, y), \text{ where } \mathcal{F} = \{S_1, \dots, S_T\} \subseteq \mathcal{P}(V) \text{ and } y = \frac{\mu}{(1 - \gamma)\alpha_{\text{GW}}} \frac{1}{T} \sum_{t=1}^T e_{S_t} \in \mathbb{R}_+^{\mathcal{F}}.$$

A shore of a sampled cut with largest weight is $\arg \max \{w^\top \mathbb{1}_{\delta(S)} : S \in \mathcal{F}\}$. These objects give rise to the sampling procedure in Algorithm 2, which we analyze next. Similar to Algorithm 1, the pseudocode of Algorithm 2 abstracts away important implementation choices, including the choice of data structures.

Proposition 21. Let $\varepsilon, \sigma, \gamma \in (0, 1)$. Let $G = (V, E)$ be a graph on n vertices and let $(w, z) \in H_{\varepsilon, \sigma}(G)$ be such that $w \neq 0 \neq z$. Set $\beta := \alpha_{\text{GW}}(1 - \gamma)(1 - \sigma)(1 - \varepsilon)$. Let $(\bar{\rho}, x)$ and $(\bar{\mu}, Y)$ witness the membership $(w, z) \in H_{\varepsilon, \sigma}(G)$. For each integer

$$T \geq \left\lceil \frac{6\pi}{(\alpha_{\text{GW}}\gamma(1 - \sigma)(1 - \varepsilon))^2\varepsilon} \ln(n) \right\rceil,$$

the randomized polynomial-time procedure $\mathcal{A}_{T,\gamma}(G, Y)$ satisfies the following: with probability at least $1 - 2/n$, we have that $(\rho, \mu, S^{\max}, y, x)$ is a β -certificate for (w, z) , where

$$(64) \quad \rho := (1 - \varepsilon)^{-1}\bar{\rho}, \quad \mu := \rho^{-1}w^\top z, \\ (\mathcal{F}, y) := \mathcal{A}_{T,\gamma}(G, Y), \quad \text{and} \quad S^{\max} := \arg \max \{w^\top \mathbb{1}_{\delta(S)} : S \in \mathcal{F}\}.$$

In particular, (w, z) is a β -pairing.

Remark. Alternatively, rather than sampling from $\text{GW}(Y)$ in the call to $\mathcal{A}_{T,\gamma}$ in Proposition 21, using the definition of η_ε as a starting point, one may use a perturbed sampling $\text{GW}_\varepsilon(Y)$ obtained by sampling from $\text{GW}(Y)$ with probability $(1 - \varepsilon)$, and by sampling uniformly among all shores with probability ε .

Proof of Proposition 21. Let $(w, z) \in H_{\varepsilon, \sigma}(G)$. Let $(\bar{\rho}, x)$ and $(\bar{\mu}, Y)$ witness the membership $(w, z) \in H_{\varepsilon, \sigma}(G)$. Note that $\bar{\rho}, \bar{\mu} > 0$ as $w \neq 0 \neq z$. Set $\rho := (1 - \varepsilon)^{-1}\bar{\rho}$ and $\mu := (1/\rho)w^\top z$. Item (45.i) in the Definition 12 of β -certificates holds trivially. We also have (45.iv), since $\text{Diag}(x) \succeq \frac{1}{4}\mathcal{L}_G(w)$ and

$$\rho = \frac{\bar{\rho}}{1 - \varepsilon} \geq \frac{1}{1 - \varepsilon} \left((1 - \varepsilon)\mathbb{1}^\top x + \frac{\varepsilon}{2}\|w\|_1 \right) \geq \mathbb{1}^\top x.$$

In particular, (ρ, x) is feasible in (4b) so

$$(65) \quad \text{mc}(G, w) \leq \eta(G, w) \leq \rho.$$

Next we prove (45.ii) and (45.iii).

Algorithm 2 Certification procedure

Parameters: a constant approximation factor $\beta \in (0, \alpha_{\text{GW}})$ parameterizes the algorithm CERTIFY_β . As in (68) and (70), define the following constants in terms of β :

$$\tau := 1 - \frac{\beta}{\alpha_{\text{GW}}} \in (0, 1), \quad \sigma := \frac{2}{3}\tau \in (0, 2/3), \quad \varepsilon := \frac{\tau}{3(3-2\tau)} \in (0, 1/3), \quad \text{and} \quad \gamma := \frac{2\tau}{9-7\tau} \in (0, 1).$$

Input: a graph $G = (V, E)$, a pair $(w, z) \in \mathbb{R}_+^E \times \mathbb{R}_+^E$ of nonzero edge weights, and witnesses $(\bar{\rho}, x)$ and $(\bar{\mu}, Y)$ of the membership of (w, z) in $H_{\varepsilon, \sigma}(G)$.

Output: $\text{CERTIFY}_\beta(G, (w, z), (\bar{\rho}, x), (\bar{\mu}, Y))$ is a β -certificate (ρ, μ, S, y, x) with high probability, where the support of y has size bounded above by $T := \lceil \frac{2187\pi}{2\alpha_{\text{GW}}^2\tau^3} \ln(|V|) \rceil$, as in (69).

```

1 procedure  $\text{CERTIFY}_\beta(G, (w, z), (\bar{\rho}, x), (\bar{\mu}, Y))$ 
2    $\mathcal{F} \leftarrow \emptyset$ 
3    $\bar{y} \leftarrow 0 \in \mathbb{R}_+^{\mathcal{P}(V)}$ 
4   repeat  $T$  times
5      $S \leftarrow \text{GW}(Y)$  ▷ Sample a shore  $S \subseteq V$  via the random hyperplane technique
6      $\mathcal{F} \leftarrow \mathcal{F} \cup \{S\}$ 
7      $\bar{y}_S \leftarrow \bar{y}_S + 1$ 
8   end
9    $S^{\max} \leftarrow \arg \max \{ w^\top \mathbb{1}_{\delta(S)} : S \in \mathcal{F} \}$ 
10   $y \leftarrow \frac{\mu}{(1-\gamma)\alpha_{\text{GW}}} \frac{1}{T} \bar{y}$ 
11   $\rho \leftarrow (1-\varepsilon)^{-1} \bar{\rho}$ 
12   $\mu \leftarrow \rho^{-1} w^\top z$ 
13  return  $(\rho, \mu, S^{\max}, y, x)$ 
14 end procedure

```

Let S_1, \dots, S_T and (\mathcal{F}, y) be defined as in (18), so that $S^{\max} = \arg \max \{ w^\top \mathbb{1}_{\delta(S)} : S \in \mathcal{F} \}$. We will now prove (45.ii) for the shore S^{\max} . More precisely, we show that $w^\top \mathbb{1}_{\delta(S^{\max})} \geq \beta\rho$ with probability at least $1 - 1/n$. As $(\bar{\mu}, Y)$ is feasible in (58b), it is feasible in (11), so from (49) we have that

$$\mathbb{E}(w^\top \mathbb{1}_{\delta(\text{GW}(Y))}) \geq \frac{\alpha_{\text{GW}}}{\bar{\mu}} \langle \frac{1}{4} \mathcal{L}_G(w), Y \rangle \geq \alpha_{\text{GW}} \frac{w^\top z}{\bar{\mu}} \geq \alpha_{\text{GW}}(1-\sigma)\bar{\rho} = \alpha_{\text{GW}}(1-\sigma)(1-\varepsilon)\rho,$$

and by (65), we can bound the range of $w^\top \mathbb{1}_{\delta(\text{GW}(Y))}$ as $0 \leq w^\top \mathbb{1}_{\delta(\text{GW}(Y))} \leq \rho$. Define $X_t := w^\top \mathbb{1}_{\delta(S_t)}$ for every $t \in [T]$. Define $S := \sum_{t=1}^T X_t$. The random variables X_1, \dots, X_T are independent and satisfy $0 \leq X_t \leq \rho$ and $\mathbb{E}(X_t) \geq \alpha_{\text{GW}}(1-\sigma)(1-\varepsilon)\rho$ for each $t \in [T]$. Using Hoeffding's inequality,

$$\begin{aligned} \mathbb{P}(X_t \leq (1-\gamma)\mathbb{E}(X_t) \text{ for all } t \in [T]) &\leq \mathbb{P}(S \leq (1-\gamma)\mathbb{E}(S)) \leq \exp\left(\frac{-\gamma^2 \mathbb{E}(S)^2}{T\rho^2}\right) \\ &\leq \exp\left(\frac{-\gamma^2(\alpha_{\text{GW}}(1-\sigma)(1-\varepsilon)\rho T)^2}{T\rho^2}\right) = \exp\left(-(\alpha_{\text{GW}}\gamma(1-\sigma)(1-\varepsilon))^2 T\right) \leq 1/n, \end{aligned}$$

since

$$T \geq \left\lceil \frac{\ln(n)}{(\alpha_{\text{GW}}\gamma(1-\sigma)(1-\varepsilon))^2} \right\rceil.$$

Thus $w^\top \mathbb{1}_{\delta(S^{\max})} \geq \alpha_{\text{GW}}(1-\gamma)(1-\sigma)(1-\varepsilon)\rho = \beta\rho$ with probability at least $1 - 1/n$, so the proof of (45.ii) is complete.

We now prove (45.iii). By (19),

$$(66) \quad \mathbb{1}^\top y = \frac{1}{(1-\gamma)\alpha_{\text{GW}}} \bar{\mu} = \frac{1}{\beta}(1-\varepsilon)(1-\sigma)\bar{\mu} \leq \frac{1}{\beta}(1-\varepsilon) \frac{w^\top z}{\bar{\rho}} = \frac{1}{\beta} \frac{w^\top z}{\rho} = \frac{1}{\beta} \mu.$$

Next we will use Proposition 4 to show that $\sum_{S \in \mathcal{F}} y_S \mathbb{1}_{\delta(S)} \geq z$ with probability at least $1 - 1/n$. Set $\hat{z} := \frac{1}{4} \mathcal{L}_G^*(Y)$. Set $\kappa := 2/\varepsilon$ and $\xi := \varepsilon/2$. We claim that the hypotheses from Proposition 4 for \hat{z} are met

by $(\bar{\mu}, Y)$ as well as ξ , κ , and γ . It is immediate that $(\bar{\mu}, Y)$ is feasible in (11) for (G, \hat{z}) . We claim that

$$(67) \quad \frac{\varepsilon}{2} \bar{\mu} \leq \hat{z}_{ij} \leq \bar{\mu} \quad \text{for every } ij \in E.$$

By using (61b), (25), (63b), and that feasibility of $(\bar{\mu}, Y)$ in (58b) for (G, \hat{z}) implies that $\eta_\varepsilon^\circ(G, \hat{z}) \leq \bar{\mu}$, we see that

$$\frac{\varepsilon}{2} \bar{\mu} \mathbb{1} = \frac{1}{4} \mathcal{L}_G^*(\varepsilon \bar{\mu} I) \leq \frac{1}{4} \mathcal{L}_G^*(Y) = \hat{z} \leq \|\hat{z}\|_\infty \mathbb{1} \leq \eta^\circ(G, \hat{z}) \mathbb{1} \leq \eta_\varepsilon^\circ(G, \hat{z}) \mathbb{1} \leq \bar{\mu} \mathbb{1}.$$

Thus (67) holds. Condition (21a) of Proposition 4 holds by the first inequality in (67); condition (21b) follows from noting that $\frac{\varepsilon}{2} \|\hat{z}\|_\infty \leq \frac{\varepsilon}{2} \bar{\mu} \leq \hat{z}_{ij}$ for every $ij \in E$. Hence, with probability at least $1 - 1/n$, $(\mathcal{F}, y) := \mathcal{A}_{T, \gamma}(G, Y)$ covers z — by covering $\hat{z} \geq z$. Thus (66) proves (45.iii), as

$$T \geq \left\lceil \frac{6\pi}{\gamma^2 \varepsilon} \ln(n) \right\rceil = \left\lceil \frac{3\pi}{\gamma^2} \left(\frac{\kappa}{\xi} \right)^{1/2} \ln(n) \right\rceil.$$

Hence, the probability that $(\rho, \mu, S^{\max}, y, x)$ is not a β -certificate for (w, z) is at most $2/n$. \square

Proposition 22. Let $\beta \in (0, \alpha_{\text{GW}})$ and set

$$(68) \quad \tau := 1 - \frac{\beta}{\alpha_{\text{GW}}} \in (0, 1), \quad \sigma := \frac{2}{3}\tau \in (0, 2/3), \quad \text{and} \quad \varepsilon := \frac{\tau - \sigma}{3(1 - \sigma)} = \frac{\tau}{3(3 - 2\tau)} \in (0, 1/3).$$

There exists a polynomial-time randomized algorithm that takes as input a graph $G = (V, E)$ on n vertices, a pair $(w, z) \in H_{\varepsilon, \sigma}(G)$, and objects $(\bar{\rho}, x)$ and $(\bar{\mu}, Y)$ witnessing the membership $(w, z) \in H_{\varepsilon, \sigma}(G)$, and, with probability at least $1 - 2/n$, outputs a β -certificate $(\rho, \mu, S^{\max}, y, x)$ for (w, z) , where

$$(69) \quad |\text{supp}(y)| \leq T := \left\lceil \frac{2187\pi}{2\alpha_{\text{GW}}^2 \tau^3} \ln(n) \right\rceil.$$

This algorithm may be implemented so that it makes:

- (1) a single call to an oracle producing a Cholesky factorization of an $n \times n$ matrix;
- (2) Tn calls to an oracle sampling from a standard Gaussian distribution; and
- (3) $O(Tn^2)$ extra work.

Proof. Set

$$(70) \quad \gamma := \frac{2\varepsilon}{1 - \varepsilon} = \frac{2\tau}{9 - 7\tau} \in (0, 1).$$

The constants σ , ε , and γ are defined so that

$$(1 - \gamma)(1 - \sigma)(1 - \varepsilon) = \left(1 - \frac{2\varepsilon}{1 - \varepsilon}\right)(1 - \varepsilon)(1 - \sigma) = (1 - 3\varepsilon)(1 - \sigma) = \left(1 - \frac{\tau - \sigma}{1 - \sigma}\right)(1 - \sigma) = 1 - \tau.$$

Thus $\beta = \alpha_{\text{GW}}(1 - \tau) = \alpha_{\text{GW}}(1 - \gamma)(1 - \sigma)(1 - \varepsilon)$. Moreover,

$$(71) \quad ((1 - \varepsilon)\gamma)^2 \varepsilon = (2\varepsilon)^2 \varepsilon = 4\varepsilon^3 = \frac{4}{27} \left(\frac{\tau - \sigma}{1 - \sigma} \right)^3.$$

The proof is then an application of Proposition 21, since

$$\begin{aligned} \left\lceil \frac{6\pi}{(\alpha_{\text{GW}}\gamma(1 - \sigma)(1 - \varepsilon))^2 \varepsilon} \ln(n) \right\rceil &= \left\lceil \frac{6\pi}{(\alpha_{\text{GW}}(1 - \sigma))^2} \frac{27}{4} \left(\frac{1 - \sigma}{\tau - \sigma} \right)^3 \ln(n) \right\rceil \quad \text{by (71)} \\ &= \left\lceil \frac{81\pi}{2\alpha_{\text{GW}}^2} \frac{(1 - \sigma)^3}{(\tau - \sigma)^3} \ln(n) \right\rceil \\ &\leq \left\lceil \frac{2187\pi}{2\alpha_{\text{GW}}^2} \frac{1}{\tau^3} \ln(n) \right\rceil \quad \text{since } \tau - \sigma = \frac{1}{3}\tau \text{ and } \sigma > 0 \\ &= T. \end{aligned}$$

Let $B \in \mathbb{R}^{[n] \times V}$ be a Cholesky factorization of Y , i.e., such that $Y = B^\top B$. Let $\Omega \in \mathbb{R}^{[T] \times [n]}$ have each entry independently sampled from a standard Gaussian distribution. Algorithm 2 may be easily implemented by computing the matrix product $X := \Omega B$ and checking the signs of the entries of X . This matrix product can be computed in $O(Tn^2)$ time, which dominates the $O(T^2n)$ time of naively computing \mathcal{F} and the $O(Tm)$ time of computing S^{\max} . \square

Algorithm 3 Primal-dual randomized rounding approximation algorithm for mc, building an instance for fcc and certificates

Parameters: a constant approximation factor $\beta \in (0, \alpha_{\text{GW}})$ parameterizes the algorithm APPROXMCFCC_β .

Define the following constants in terms of β :

$$\tau := 1 - \frac{\beta}{\alpha_{\text{GW}}}, \quad \sigma := \frac{2}{3}\tau, \quad \varepsilon := \frac{\tau}{3(3-2\tau)}, \quad \gamma := \frac{2\tau}{9-7\tau}, \quad \text{and} \quad C := \frac{2187\pi}{2\alpha_{\text{GW}}^2\tau^3}.$$

Input: a graph $G = (V, E)$ and nonzero edge weights $w \in \mathbb{R}_+^E$.

Output: $\text{APPROXMCFCC}_\beta(G, w)$ returns a nonzero $z \in \mathbb{R}_+^E$ and a β -certificate (ρ, μ, S, y, x) for (w, z) with probability at least $1 - 2/|V|$, such that $|\text{supp}(y)| \leq \lceil C \ln(|V|) \rceil$. The algorithm runs in strongly polynomial time.

```

1 procedure APPROXMCFCC $_\beta(G, w)$ 
2   Find a feasible pair  $(\tilde{Y}, (\tilde{\rho}, x))$  for (4) such that  $\tilde{\rho} \leq \langle \frac{1}{4}\mathcal{L}_G(w), \tilde{Y} \rangle + \sigma\|w\|_\infty$ .
3    $\bar{\mu} \leftarrow 1$ 
4    $Y \leftarrow (1 - \varepsilon)\tilde{Y} + \varepsilon I$ 
5    $z \leftarrow \frac{1}{4}\mathcal{L}_G^*(Y)$ 
6    $\bar{\rho} \leftarrow (1 - \varepsilon)\mathbb{1}^\top x + \frac{\varepsilon}{2}\|w\|_1$  ▷ Solve  $\eta_\varepsilon(G, w)$  within  $\sigma\|w\|_\infty$ 
7   return  $(z, \text{CERTIFY}_\beta(G, (w, z), (\bar{\rho}, x), (\bar{\mu}, Y)))$ 
8 end procedure

```

Algorithm 4 Primal-dual randomized rounding approximation algorithm for fcc, building an instance for mc and certificates for both instances

Parameters: a constant approximation factor $\beta \in (0, \alpha_{\text{GW}})$ parameterizes the algorithm APPROXFCCMC_β .

Define the following constants in terms of β :

$$\tau := 1 - \frac{\beta}{\alpha_{\text{GW}}}, \quad \sigma := \frac{2}{3}\tau, \quad \varepsilon := \frac{\tau}{3(3-2\tau)}, \quad \gamma := \frac{2\tau}{9-7\tau}, \quad \text{and} \quad C := \frac{2187\pi}{2\alpha_{\text{GW}}^2\tau^3}.$$

Input: a graph $G = (V, E)$ and nonzero edge weights $z \in \mathbb{R}_+^E$.

Output: $\text{APPROXFCCMC}_\beta(G, z)$ returns a nonzero $w \in \mathbb{R}_+^E$ and a β -certificate (ρ, μ, S, y, x) for (w, z) with probability at least $1 - 2/|V|$, such that $|\text{supp}(y)| \leq \lceil C \ln(|V|) \rceil$. The algorithm runs in strongly polynomial time.

```

1 procedure APPROXFCCMC $_\beta(G, z)$ 
2   Find a feasible pair  $(\bar{\rho}, w), (\bar{\mu}, Y)$  for (58) such that  $\bar{\mu} \leq z^\top w + \sigma\|z\|_\infty$ . ▷ Solve  $\eta_\varepsilon^o(G, z)$  within  $\sigma\|z\|_\infty$ 
3   return  $(w, \text{CERTIFY}_\beta(G, (w, z), (\bar{\rho}, x), (\bar{\mu}, Y)))$ 
4 end procedure

```

Theorem 23. Let $\beta \in (0, \alpha_{\text{GW}})$. Set $\tau := 1 - \beta/\alpha_{\text{GW}}$ and $C := 2187\pi/(2\alpha_{\text{GW}}^2\tau^3)$. There exists a randomized polynomial-time algorithm that, given a graph $G = (V, E)$ on n vertices and a nonzero $w \in \mathbb{R}_+^E$, computes with probability at least $1 - 2/n$ a nonzero vector $z \in \mathbb{R}_+^E$ and a β -certificate (ρ, μ, S, y, x) for (w, z) such that the support of the fractional cut cover y has size at most $\lceil C \ln(n) \rceil$.

Proof. Let $w \in \mathbb{R}_+^E$ be nonzero. Define $\sigma := \frac{2}{3}\tau$ and $\varepsilon := \tau/(9 - 6\tau)$ as in Proposition 22. By (nearly) solving the primal-dual SDPs in (4), one can compute in polynomial time a feasible solution \tilde{Y} for (4a) and a feasible solution $(\tilde{\rho}, x)$ for (4b) such that

$$\tilde{\rho} \leq \langle \frac{1}{4}\mathcal{L}_G(w), \tilde{Y} \rangle + \sigma\|w\|_\infty.$$

Set

$$(72) \quad Y := (1 - \varepsilon)\tilde{Y} + \varepsilon I, \quad z := \frac{1}{4}\mathcal{L}_G^*(Y), \quad \bar{\rho} := (1 - \varepsilon)\mathbb{1}^\top x + \frac{\varepsilon}{2}\|w\|_1.$$

We claim that

$$(73) \quad (\bar{\rho}, x) \text{ and } (1, Y) \text{ witness the membership } (w, z) \in H_{\varepsilon, \sigma}(G).$$

It is immediate that $(\bar{\rho}, x)$ is feasible in (57b), whereas $Y \succeq \varepsilon I$ implies $(1, Y)$ is feasible in (58b). Since $\|w\|_\infty \leq \eta(G, w)$, we can use (63a) and (57b) to see

$$(74) \quad (1 - \varepsilon)\|w\|_\infty \leq (1 - \varepsilon)\eta(G, w) \leq \eta_\varepsilon(G, w) \leq \bar{\rho}.$$

Then

$$\begin{aligned} \bar{\rho} &= (1 - \varepsilon)\mathbb{1}^\top x + \frac{\varepsilon}{2}\|w\|_1 && \text{by (72)} \\ &\leq (1 - \varepsilon)(\langle \tfrac{1}{4}\mathcal{L}_G(w), \tilde{Y} \rangle + \sigma\|w\|_\infty) + \frac{\varepsilon}{2}\|w\|_1 && \text{since } \mathbb{1}^\top x \leq \tilde{\rho} \leq \langle \tfrac{1}{4}\mathcal{L}_G(w), \tilde{Y} \rangle + \sigma\|w\|_\infty \\ &= \langle \tfrac{1}{4}\mathcal{L}_G(w), (1 - \varepsilon)\tilde{Y} + \varepsilon I \rangle + (1 - \varepsilon)\sigma\|w\|_\infty && \text{since } \tfrac{\varepsilon}{2}\|w\|_1 = \langle \tfrac{1}{4}\mathcal{L}_G(w), \varepsilon I \rangle \\ &= w^\top z + (1 - \varepsilon)\sigma\|w\|_\infty && \text{by the definitions of } Y \text{ and } z \\ &\leq w^\top z + \sigma\bar{\rho}. && \text{by (74)} \end{aligned}$$

This completes the proof of (73). Moreover, as $w \neq 0$, we get from (74) that $\bar{\rho} > 0$. Hence $w^\top z \geq (1 - \sigma)\bar{\rho} > 0$, as $\sigma < 1$. Thus $z \neq 0$. Proposition 22 then finishes the proof. \square

Theorem 24. Let $\beta \in (0, \alpha_{\text{GW}})$. Set $\tau := 1 - \beta/\alpha_{\text{GW}}$ and $C := 2187\pi/(2\alpha_{\text{GW}}^2\tau^3)$. There exists a randomized polynomial-time algorithm that, given a graph $G = (V, E)$ on n vertices and a nonzero $z \in \mathbb{R}_+^E$, computes with probability at least $1 - 2/n$ a nonzero vector $w \in \mathbb{R}_+^E$ and a β -certificate (ρ, μ, S, y, x) for (w, z) such that the support of the fractional cut cover y has size at most $\lceil C \ln(n) \rceil$.

Proof. Let $z \in \mathbb{R}_+^E$ be nonzero. Define $\sigma := \frac{2}{3}\tau$ and $\varepsilon := \tau/(9 - 6\tau)$, as in Proposition 22. By (nearly) solving the primal-dual SDPs in (58), one can compute in polynomial time a feasible solution (x, w) for (58a) and a feasible solution $(\bar{\mu}, Y)$ for (58b) such that

$$\bar{\mu} \leq w^\top z + \sigma\|z\|_\infty.$$

By combining (25), (63b), and (58b), we see that

$$(75) \quad \|z\|_\infty \leq \eta^\circ(G, z) \leq \eta_\varepsilon^\circ(G, z) \leq \bar{\mu}.$$

We claim that

$$(76) \quad (1, x) \text{ and } (\bar{\mu}, Y) \text{ witness the membership } (w, z) \in H_{\varepsilon, \sigma}(G).$$

Feasibility of $(1, x)$ and $(\bar{\mu}, Y)$ in the appropriate SDPs are easily verified. Moreover, by (75), one has $\bar{\mu} \leq w^\top z + \sigma\|z\|_\infty \leq w^\top z + \sigma\bar{\mu}$. This proves (76). Moreover, as $z \neq 0$, we get from (75) that $\bar{\mu} > 0$. Hence $w^\top z \geq \bar{\mu}(1 - \sigma) > 0$ as $\sigma < 1$. Thus $w \neq 0$. Proposition 22 finishes the proof. \square

Remark 25. Theorem 23 uses that the SDP (4) is nearly solvable in polynomial time. However, Theorem 24 relies on nearly solving the SDP (58), which is introduced in this work. Appendix C proves this can be done in polynomial time.

4. GEOMETRIC REPRESENTATION OF GRAPHS

Definition 26. A *hypersphere representation* of a graph $G = (V, E)$ is a map $u: V \rightarrow \mathbb{R}^d$ for some $d \in \mathbb{N}$ such that the map $i \in V \mapsto \|u_i\| \in \mathbb{R}$ is constant. Such constant is the *radius* of u , denoted by $r(u)$. We denote by $\mathcal{H}(G)$ the set of all hypersphere representations of G .

Let $G = (V, E)$ be a graph, and let $\mu \in \mathbb{R}_+$. Hypersphere representations of G with squared radius μ are directly related to $\mu\mathcal{E}^V$ via their Gram matrices: if $u \in \mathcal{H}(G)$ has squared radius μ and one defines a matrix U with columns $\{u_i\}_{i \in V}$, then $U^\top U \in \mu\mathcal{E}^V$; conversely, if $X \in \mu\mathcal{E}^V$, then the columns of $X^{1/2}$ form a hypersphere representation of G with squared radius μ .

Let $X \in \mu\mathcal{E}^V$ be the Gram matrix corresponding to a hypersphere representation $u \in \mathcal{H}(G)$. Then for every $z \in \mathbb{R}_+^E$,

$$(77) \quad \mathcal{L}_G^*(X) \geq z \text{ if and only if } \|u_i - u_j\|^2 \geq z_{ij} \text{ for every } ij \in E.$$

In this way, the constraint defined by the adjoint of the Laplacian has a natural geometric interpretation.

The study of geometric representations of graphs has been very fruitful [29]. To our knowledge, geometric representations provided the first proof that specializing the SDP in (11) to the case $z = \mathbb{1}$ recovers the relationship with the vector chromatic number. A *unit-distance representation* of a graph $G = (V, E)$ is a

hypersphere representation $u: V \rightarrow \mathbb{R}^d$ of G such that $\|u_i - u_j\|^2 = 1$ for every $ij \in E$. The *hypersphere number* of G , denoted by $t(G)$, is the smallest squared radius of a unit-distance representation of G . Using semidefinite programming, one can write

$$\begin{aligned} t(G) &= \inf \{ r(u)^2 : u \in \mathcal{H}(G) \text{ such that } \|u_i - u_j\|^2 = 1 \text{ for every } ij \in E \} \\ &= \inf \{ \mu : \mu \in \mathbb{R}_+, X \in \mathbb{S}_+^V, \text{diag}(X) = \mu \mathbb{1}, \mathcal{L}_G^*(X) = \mathbb{1} \}. \end{aligned}$$

The only differences between this optimization problem and the problem (26) specialized to $z = \mathbb{1}$ appear in the constraint featuring the adjoint of the Laplacian: the $\frac{1}{4}$ factor is gone, and ‘ \geq ’ was changed to ‘ $=$ ’. Lovász [30, p. 23] proved that

$$(78) \quad 2t(G) + \frac{1}{\vartheta(\overline{G})} = 1,$$

where ϑ denotes the *Lovász theta function* [31]. The similarities between the Lovász theta function ϑ and the vector chromatic number are already discussed in the work introducing χ_{vec} [24]. In fact, $\chi_{\text{vec}}(G) = \vartheta'(\overline{G})$, where \overline{G} denotes the complement of a graph and ϑ' denotes a variant of ϑ introduced independently in [33] and [40]. A natural variation of (78) that involves ϑ' (see, e.g., [8, Sec. 4]) is

$$2t'(G) + \frac{1}{\chi_{\text{vec}}(G)} = 1,$$

where

$$t'(G) := \inf \{ \mu : X \in \mathbb{S}_+^V, \text{diag}(X) = \mu \mathbb{1}, \mathcal{L}_G(X) \geq \mathbb{1} \} = \frac{1}{4} \eta^\circ(G).$$

It is then immediate that

$$(79) \quad \eta^\circ(G, \mathbb{1}) = 2 \left(1 - \frac{1}{\chi_{\text{vec}}(G)} \right).$$

Our introduction presented (10) as a motivating question of our work. Note that (79) and Corollary 2 provide an alternative proof of (10). Equation (79) is another manifestation of the well-known connection between the Lovász theta function (and its variants) with the maximum cut problem — see, e.g., [9, 28].

The optimization problems we have considered optimize different objective functions over certain hypersphere representations of a graph $G = (V, E)$. For each $z \in \mathbb{R}_+^E$, an *optimal fcc representation* of (G, z) is a hypersphere representation $u: V \rightarrow \mathbb{R}^d$ such that $\frac{1}{4} \|u_i - u_j\|^2 \geq z_{ij}$ for every edge $ij \in E$, and with minimal radius among such representations. For each $w \in \mathbb{R}_+^E$, an *optimal mc representation* of (G, w) is a hypersphere representation $u: V \rightarrow \mathbb{R}^d$ with radius 1 which maximizes $\sum_{ij \in E} w_{ij} \|u_i - u_j\|^2$. Define, for every $w, z \in \mathbb{R}_+^E$,

$$\begin{aligned} \text{FCC}_{\mathcal{H}}(G, z) &:= \arg \min \{ r(u)^2 : u \in \mathcal{H}(G) \text{ such that } \frac{1}{4} \|u_i - u_j\|^2 \geq z_{ij} \text{ for every } ij \in E \}, \text{ and} \\ \text{MC}_{\mathcal{H}}(G, w) &:= \arg \max \left\{ \frac{1}{4} \sum_{ij \in E} w_{ij} \|u_i - u_j\|^2 : u \in \mathcal{H}(G), r(u) = 1 \right\}. \end{aligned}$$

The connection between these two sets of optimal geometric representations, which we illustrate with Figure 2, is captured by the following theorem.

Theorem 27. Let $G = (V, E)$ be a graph. If $w \in \mathbb{R}_+^E$ is nonzero, then for every $u \in \text{MC}_{\mathcal{H}}(G, w)$ there exists $z \in \mathfrak{z}_G(w)$ such that

$$(80) \quad (\sqrt{\mu} u_i)_{i \in V} \in \text{FCC}_{\mathcal{H}}(G, z), \text{ where } \mu := \min \{ \rho \in \mathbb{R}_+ : \rho \frac{1}{4} \|u_i - u_j\|^2 \geq z_{ij} \text{ for every } ij \in E \}.$$

Conversely, if $z \in \mathbb{R}_+^E$ is nonzero, then for every $v \in \text{FCC}_{\mathcal{H}}(G, z)$ there exists $w \in \mathfrak{w}_G(z)$ such that

$$(81) \quad (r(v)^{-1} v_i)_{i \in V} \in \text{MC}_{\mathcal{H}}(G, w).$$

Remark. Let $G = (V, E)$ be a graph. Note that (80) describes an algorithm producing an element in $\text{FCC}_{\mathcal{H}}(G, z)$ from inputs $G, w \in \mathbb{R}_+^E, u \in \text{MC}_{\mathcal{H}}(G, w)$, and $z \in \mathfrak{z}_G(w)$. Similarly, (81) describes an algorithm producing an element in $\text{MC}_{\mathcal{H}}(G, w)$ from inputs $G, z \in \mathbb{R}_+^E, v \in \text{FCC}_{\mathcal{H}}(G, z)$, and $w \in \mathfrak{w}_G(z)$. If one is given $w \in \mathbb{R}_+^E$ and an optimal mc representation $u \in \text{MC}_{\mathcal{H}}(G, w)$ and simply wants to obtain an associated optimal fcc representation, there is an easily computable choice of $z \in \mathfrak{z}_G(w)$ — namely, $z_{ij} := \frac{1}{4} \|u_i - u_j\|^2$ for every $ij \in E$.

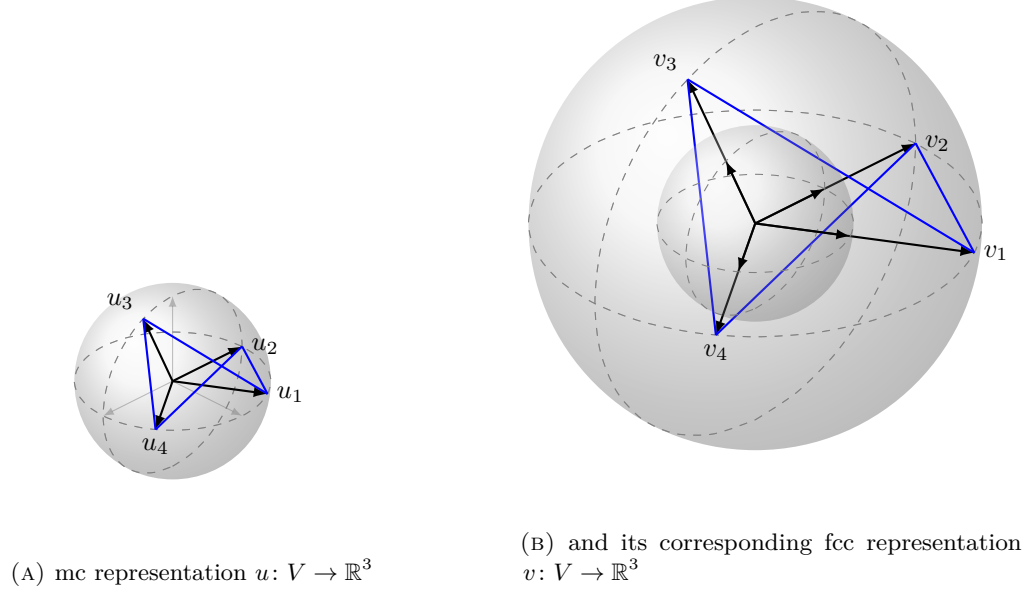


FIGURE 2. Geometric equivalence between hypersphere representations.

Proof. Let $w \in \mathbb{R}_+^E$ be nonzero, and let $u \in \text{MC}_{\mathcal{H}}(G, w)$. Let $Y \in \mathcal{E}^V$ be the Gram matrix corresponding to u , i.e., $Y_{ij} = u_i^\top u_j$ for every $i, j \in V$. Since $u \in \text{MC}_{\mathcal{H}}(G, w)$, there exists $z \in \mathbb{R}_+^E$ and $(\rho, x) \in \mathbb{R}_+ \times \mathbb{R}^V$ such that (z, Y) and (ρ, x) are optimal solutions for the SDPs (52a) and (52b), respectively. We will prove that $v := (\sqrt{\mu} u_i)_{i \in V} \in \text{FCC}_{\mathcal{H}}(G, z)$. From optimality of (z, Y) ,

$$(82) \quad r(v)^2 = \min\{\rho \in \mathbb{R}_+ : \rho \frac{1}{4} \|u_i - u_j\|^2 \geq z_{ij} \text{ for every } ij \in E\} = \min\{\rho \in \mathbb{R}_+ : \frac{1}{4} \mathcal{L}_G^*(\rho Y) \geq z\} = 1.$$

Now let $\tilde{v} \in \mathcal{H}(G)$ be such that $\frac{1}{4} \|\tilde{v}_i - \tilde{v}_j\|^2 \geq z_{ij}$ for every $ij \in E$, and let \tilde{Y} be the Gram matrix of \tilde{v} . Since (ρ, x) is feasible in (52b),

$$\rho r(\tilde{v})^2 \geq \mathbb{1}^\top x r(\tilde{v})^2 = \langle \text{Diag}(x), \tilde{Y} \rangle \geq \langle \frac{1}{4} \mathcal{L}_G(w), \tilde{Y} \rangle = \sum_{ij \in E} w_{ij} \frac{1}{4} \|\tilde{v}_i - \tilde{v}_j\|^2 \geq w^\top x = \rho.$$

As $w \neq 0$, we have that $\rho > 0$, so $r(\tilde{v})^2 \geq 1$. As the latter inequality holds with equality at v by (82), we conclude that $v \in \text{FCC}_{\mathcal{H}}(G, z)$.

Let $z \in \mathbb{R}_+^E$ be nonzero, let $v \in \text{FCC}_{\mathcal{H}}(G, z)$, and let Y be the Gram matrix of v . Since $v \in \text{FCC}_{\mathcal{H}}(G, z)$, there exists $(w, x) \in \mathbb{R}_+^E \times \mathbb{R}^V$ such that $(r(v)^2, Y)$ and (w, x) are optimal solutions to (26a) and (26b), respectively. We will prove that $u := (r(v)^{-1} v_i)_{i \in V} \in \text{MC}_{\mathcal{H}}(G, w)$. We claim that

$$(83) \quad 1 = \frac{1}{4} \sum_{ij \in E} w_{ij} \|u_i - u_j\|^2.$$

Indeed, by optimality of $(r(v)^2, Y)$ and (w, x) , we have that $r(v)^2 \geq r(v)^2 \mathbb{1}^\top x = \langle Y, \text{Diag}(x) \rangle \geq \langle Y, \frac{1}{4} \mathcal{L}_G(w) \rangle \geq z^\top w = r(v)^2$. Thus

$$r(v)^2 = \langle Y, \frac{1}{4} \mathcal{L}_G(w) \rangle = \frac{1}{4} \sum_{ij \in E} w_{ij} \|v_i - v_j\|^2 = r(v)^2 \frac{1}{4} \sum_{ij \in E} w_{ij} \|u_i - u_j\|^2.$$

As $z \neq 0$, we have that $r(v)^2 > 0$, so (83) holds. Now let $\tilde{u} \in \mathcal{H}(G)$ be such that $r(\tilde{u}) = 1$. Let \tilde{Y} be the Gram matrix corresponding to \tilde{u} . Then $u \in \text{MC}_{\mathcal{H}}(G, w)$ follows from (83), as

$$\frac{1}{4} \sum_{ij \in E} w_{ij} \|\tilde{u}_i - \tilde{u}_j\|^2 = \langle \tilde{Y}, \frac{1}{4} \mathcal{L}_G(w) \rangle \leq \langle \tilde{Y}, \text{Diag}(x) \rangle = r(\tilde{u})^2 \mathbb{1}^\top x \leq 1 = \frac{1}{4} \sum_{ij \in E} w_{ij} \|u_i - u_j\|^2. \quad \square$$

5. TIGHTNESS OF OUR RESULTS

This section discusses several aspects in which our algorithms and analyses are best possible. We collect instances of the maximum cut and fractional cut-covering problems that justify several aspects of our algorithms, including

- (i) the need to sparsify the cut cover defined in Proposition 1,
- (ii) the need to either “thicken” edges, as in Section 2, or to perturb the SDP, as in Section 3.4,
- (iii) the asymptotic support size of the fractional cut cover we have obtained,
- (iv) the approximation factor.

These aspects are intertwined. Item (i) motivates the use of repeated sampling. Item (ii) shows that naively sampling cuts until a fractional cut cover is obtained takes too long for some choice of weights, and item (iii) discusses the amount of cuts needed to be sampled. The whole algorithmic set up grounds our analysis and guides our discussion of item (iv).

5.1. Sparsification of Rounded Solution. Let $G = (V, E)$ be a graph, and let $z \in \mathbb{R}_+^E$. Let (μ, Y) be feasible in (11), and set

$$(84) \quad y := \bar{\nu} p, \text{ where } p_S := \mathbb{P}(\text{GW}(Y) = S) \text{ for every } S \subseteq V, \text{ and } \bar{\nu} := \min \left\{ \nu \in \mathbb{R}_+ : \nu \sum_{S \subseteq V} p_S \mathbb{1}_{\delta(S)} \geq z \right\}.$$

Proposition 1 implies that $\bar{\nu} \leq \mu/\alpha_{\text{GW}}$. Via (84) we have a deterministic fractional cut cover $y \in \mathbb{R}_+^{\mathcal{P}(V)}$ for every feasible solution (μ, Y) of (11). It is then natural to question the necessity of the repeated sampling approach. Remark 3 mentions the difficulty of computing the vector $p \in \mathbb{R}_+^{\mathcal{P}(V)}$, and we now exhibit an instance where y has exponential support size. One may check that $(\bar{\mu}, \bar{Y}) := (2 - \frac{2}{n}, 2I - \frac{2}{n} \mathbb{1} \mathbb{1}^\top)$ is an optimal solution to $\eta^\circ(K_n)$ for every nonzero $n \in \mathbb{N}$, certified by the dual optimal solution $(\bar{x}, \bar{w}) := (\frac{1}{n} \mathbb{1}, \frac{4}{n^2} \mathbb{1})$ for (26b). For every $i \in [n]$, let g_i be independently sampled from the standard normal distribution, and set $h := \|g\|^{-1} g$. One can prove — see Appendix B — that

$$(85) \quad \text{supp}(y) = \text{supp}(p) = \{S \subseteq V : \mathbb{P}(\text{GW}(\bar{Y}, h) = S) > 0\} = \{S \subseteq V : S \neq \emptyset, S \neq V\}.$$

Hence, the vector $y \in \mathbb{R}_+^{\mathcal{P}(V)}$ defined in (84) may have exponential support size.

5.2. Edge Thickening or SDP Perturbation. The repeated sampling approach naturally arises as a sparsification of the probability distribution in (84). Let $G = (V, E)$ be a graph, let $z \in \mathbb{R}_+^E$, and let (μ, Y) be feasible in (11). Set $\hat{y}_0 := 0$. For every nonzero $t \in \mathbb{N}$, set

$$(86) \quad \hat{y}_t := \hat{y}_{t-1} + e_{S(t)} \text{ and } \mu_t := \inf \left\{ \nu \in \mathbb{R}_+ : \nu \sum_{S \subseteq V} (\hat{y}_t)_S \mathbb{1}_{\delta(S)} \geq z \right\},$$

where each $S(t)$ is independently sampled from $\text{GW}(Y)$. These objects capture what we mean by “repeated sampling”. We claim that

$$(87) \quad \mathbb{P} \left(\lim_{t \rightarrow \infty} \mu_t \hat{y}_t = y \right) = 1,$$

where $y \in \mathbb{R}_+^{\mathcal{P}(V)}$ is defined as in (84). In words, (87) states that almost surely, the vector y describes the behavior of the repeated sampling approach as more samples are taken. We now prove (87). Let $B \in \mathbb{R}^{E \times \mathcal{P}(V)}$ be the incidence matrix of the cuts of G , so $Bx = \sum_{S \subseteq V} x_S \mathbb{1}_{\delta(S)}$ for every $x \in \mathbb{R}^{\mathcal{P}(V)}$. Set $\mathcal{D} := \{x \in \mathbb{R}_+^{\mathcal{P}(V)} : \text{supp}(Bx) \supseteq \text{supp}(z)\}$, and set $f(x) := \max\{z_{ij}/(Bx)_{ij} : ij \in \text{supp}(z)\}$ for every $x \in \mathcal{D}$. Observe that f is continuous on \mathcal{D} , that $p \in \mathcal{D}$ by Proposition 1 and $f(p) = \bar{\nu}$, and that $f(\alpha x) = \frac{1}{\alpha} x$ for every $\alpha \in \mathbb{R}_{++}$ and $x \in \mathcal{D}$. The Strong Law of Large Numbers implies that $\lim_{t \rightarrow \infty} \frac{1}{t} \hat{y}_t = p$ almost surely. Assume that this event happens. For every sufficiently large $t \in \mathbb{N}$, one has $\hat{y}_t \in \mathcal{D}$, as $\text{supp}(B\hat{y}_t) = \text{supp}(Bp) \supseteq \text{supp}(z)$. Thus (87) holds, as

$$\lim_{t \rightarrow \infty} \mu_t \hat{y}_t = \lim_{t \rightarrow \infty} f(\hat{y}_t) \hat{y}_t = \lim_{t \rightarrow \infty} f(\hat{y}_t) t \frac{1}{t} \hat{y}_t = \lim_{t \rightarrow \infty} f\left(\frac{1}{t} \hat{y}_t\right) \frac{1}{t} \hat{y}_t = f(p) p = \bar{\nu} p = y.$$

Proposition 28. Let $\varepsilon \in (0, 2)$. There exist a graph $G = (V, E)$, vectors $w \in \mathbb{R}_+^E$ and $z \in \mathbb{R}_+^E$, as well as $(1, x)$ and (μ, Y) witnessing $(w, z) \in \text{H}(G)$ such that, for μ_t and \hat{y}_t defined as in (86),

$$\mathbb{E}(\min\{t \in \mathbb{N} : \mu_t < +\infty\}) \geq \left(\frac{2\sqrt{\varepsilon}}{\pi} + O(\varepsilon^{3/2}) \right)^{-1}.$$

Proof. Let $G := K_3$ be the complete graph on three vertices, with $V(K_3) = \{1, 2, 3\}$. Set $(z_{12}, z_{13}, z_{23}) := (1, 1, \varepsilon) \in \mathbb{R}_+^E$. Appendix A shows there exist $w \in \mathbb{R}_+^E$ and $x \in \mathbb{R}^V$ such that $(1, x)$ and (μ, Y) witness the membership $(w, z) \in H(G)$, where $\mu := 4/(4 - \varepsilon)$ and

$$Y := \frac{4}{4 - \varepsilon} \begin{bmatrix} 1 & \varepsilon/2 - 1 & \varepsilon/2 - 1 \\ \varepsilon/2 - 1 & 1 & 1 - 2\varepsilon + \varepsilon^2/2 \\ \varepsilon/2 - 1 & 1 - 2\varepsilon + \varepsilon^2/2 & 1 \end{bmatrix}.$$

If $t \in \mathbb{N}$ is such that $\mu_t < +\infty$, then the edge $23 \in E$ was covered, and hence

$$\mathbb{E}(\min\{t \in \mathbb{N} : \mu_t < +\infty\}) \geq \frac{1}{\mathbb{P}(\{2, 3\} \in \delta(\text{GW}(Y)))} = \frac{\pi}{\arccos(1 - 2\varepsilon + \varepsilon^2/2)} = \left(\frac{2\sqrt{\varepsilon}}{\pi} + O(\varepsilon^{3/2})\right)^{-1}. \quad \square$$

Although (87) ensures repeated sampling converges to the solution in (84) almost surely, algorithmically it is necessary to bound the number of samples one has to take. Remark 5 mentions that edges with relatively small weights can force exponentially many samples to be taken just to enable feasibility. Proposition 28 formalizes that remark: it defines a family of instances where naively sampling from an optimal solution to the SDP (11) may require, in expectation, exponentially (on the size of the input (G, z)) many cuts just to output a feasible fractional cut cover. Proposition 28 also motivates the perturbed SDPs introduced in (57) and (58), as it shows, in particular, that one cannot take $\varepsilon = 0$ in Proposition 22.

5.3. Asymptotic Support Size. Let $G = (V, E)$ be a graph, let $z \in \mathbb{R}_+^E$, and set $n := |V|$. Proposition 4 shows that, by assuming the ratio between the largest and smallest entry of z is bounded, we can produce in polynomial time a fractional cut cover with support size $O(\ln(n))$. Our algorithms then perturb the input so this hypothesis is met. The logarithmic bound may not be asymptotically improved without further assumptions on the input. Assume that $\text{supp}(z) = E$. We claim that

$$(88) \quad \lceil \lg(\chi(G)) \rceil \leq |\text{supp}(y)| \text{ for every fractional cut cover } y \text{ of } (G, z),$$

where $\chi(G)$ denotes the chromatic number of G . Since $\text{supp}(z) = E$, every edge must be in some cut defined by a shore in $\text{supp}(y)$. The minimum number of cuts necessary to cover the edges of a graph is $\lceil \lg(\chi(G)) \rceil$ — see, for example, [12, Section 6]. Thus (88) holds. In particular, the bound on $|\text{supp}(y)|$ given by Theorem 6 or Proposition 22 is asymptotically best possible for graphs such that $\chi(G) = \Theta(|V(G)|)$ — in particular, for complete graphs.

5.4. Approximation Factor. The algorithms we presented work in two steps: first they solve an SDP relaxation, and then they employ a rounding procedure to convert nearly optimal SDP solutions into actual combinatorial solutions — namely, (the shore of) a cut or a fractional cut cover. Each of these two stages impacts the approximation factor. Let $G = (V, E)$ be a graph. The *integrality ratio of a maximum cut instance* (G, w) (with respect to η) is $\text{mc}(G, w)/\eta(G, w)$. The *integrality ratio of a fractional cut-covering instance* (G, z) (with respect to η°) is $\eta^\circ(G, z)/\text{fcc}(G, z)$. In either case, the integrality ratio is a number between 0 and 1 capturing how well the semidefinite programming relaxation approximates the actual optimal value of the problem at hand.

Our theory ties the integrality ratios of both problems. Let $G = (V, E)$ be a graph, and let $w \in \mathbb{R}_+^E$. Then

$$(89) \quad \frac{\eta^\circ(G, z)}{\text{fcc}(G, z)} \leq \frac{\text{mc}(G, w)}{\eta(G, w)} \text{ for every } z \in \mathfrak{z}_G(w).$$

Indeed, this follows from $\eta(G, w)\eta^\circ(G, z) = w^\top z \leq \text{mc}(G, w)\text{fcc}(G, z)$. In this way, the set $\mathfrak{z}_G(w)$ describes instances (G, z) of the fractional cut covering problem with the same or worse integrality ratio. This construction can be made algorithmic via the tools developed in Section 2. If the graph G is edge-transitive, then one can prove that $(\mathbb{1}, \mathbb{1}) \in H(G)$ and $\text{mc}(G)\text{fcc}(G) = |E| = \eta(G)\eta^\circ(G)$. In particular, equality holds in (89), and the cycle on five vertices C_5 is a concrete example with bad integrality ratio for both problems, as

$$0.878 \approx \alpha_{\text{GW}} \leq \frac{\eta^\circ(C_5)}{\text{fcc}(C_5)} = \frac{\text{mc}(C_5)}{\eta(C_5)} \approx 0.884,$$

where unweighted graph parameters are evaluated from their weighted versions with $\mathbb{1}$ as edge weights. The integrality ratio can be arbitrarily close to α_{GW} .

Proposition 29. For every $\varepsilon > 0$, there exists a graph G such that for every $z \in \mathfrak{J}_G(\mathbb{1})$,

$$\alpha_{\text{GW}} \leq \frac{\eta^\circ(G, z)}{\text{fcc}(G, z)} \leq \frac{\text{mc}(G)}{\eta(G)} \leq \alpha_{\text{GW}} + \varepsilon.$$

Proof. Let $\varepsilon > 0$. Feige and Schechtman [13] prove that there exists a graph G such that

$$\frac{\text{mc}(G)}{\eta(G)} \leq \alpha_{\text{GW}} + \varepsilon.$$

The result follows from Corollary 2 and (89). \square

Proposition 29, despite showing that the approximation factor α_{GW} is tight in our analysis, leaves open the possibility that strengthening our semidefinite programming relaxations could lead to better results, even if we keep the same rounding procedure. This is not the case, as there exist graphs in which the relaxation is tight, but the rounding procedure still produces solutions with approximation factor as bad as α_{GW} . We now present such an example. Similar to the proof of Proposition 29, our result builds on what is known in the literature, and it exploits certain simple eigenvalue bounds.

Let $G = (V, E)$ be a graph, and let $L(G) := \mathcal{L}_G(\mathbb{1})$ denote the unweighted Laplacian of G . It holds that

$$(90) \quad \eta(G) \leq \frac{|V|}{4} \lambda_{\max}(L(G)) \text{ and } \eta^\circ(G) \geq \frac{4|E|}{|V|} \frac{1}{\lambda_{\max}(L(G))}.$$

One may easily produce feasible solutions for (4b) and (26b) that prove (90). The first inequality is a well-known bound on the maximum cut value, easily obtained from the characterization of eigenvalues in terms of Rayleigh quotients; see, for example, [18, Lemma 13.7.4]. More generally, it holds that

$$\eta(G, w) = \min \left\{ \frac{|V|}{4} \lambda_{\max}(\mathcal{L}_G(w) + \text{Diag}(u)) : u \in \mathbb{R}^V, u^\top \mathbb{1} = 0 \right\};$$

this formulation was introduced in [11]. The second inequality in (90) is used — implicitly — in [38]. Let $a, b \in \mathbb{N}$ be such that $b \leq a$. Let $\text{Ham}(a, b)$ denote the *Hamming distance graph*, which has as vertex set $\{0, 1\}^a$, with two vertices adjacent if they differ in at least b entries. We denote by $\text{Ham}_=(a, b)$ the *exact Hamming distance graph*, which is the spanning subgraph of $\text{Ham}(a, b)$ where two vertices are adjacent when they differ in exactly b entries. It is known [1, 38] that

$$(91) \quad \text{if } b \leq a < 2b \text{ and } b \text{ is even, then } \lambda_{\min}(A(\text{Ham}_=(a, b))) = \binom{a}{b} \left(1 - \frac{2b}{a}\right),$$

where $A(G) \in \mathbb{S}^V$ denotes the adjacency matrix of a graph $G = (V, E)$. The interest on the smallest eigenvalue of the adjacency matrix in both works [1, 38] stems directly from (90): since $G := \text{Ham}_=(a, b)$ is $\binom{a}{b}$ -regular, it follows that $\lambda_{\max}(L(G)) = \binom{a}{b} - \lambda_{\min}(A(G))$.

Proposition 30. For every real number $\beta > \alpha_{\text{GW}}$ there exists a graph $G = (V, E)$, as well as witnesses (ρ, x) and (μ, Y) of the membership $(\mathbb{1}, \mathbb{1}) \in H(G)$, satisfying the following conditions. One has that $\eta(G) = \text{mc}(G)$ and $\eta^\circ(G) = \text{fcc}(G)$. Furthermore,

$$(92) \quad \mathbb{E}(|\delta(\text{GW}(Y))|) < \beta \text{mc}(G) \text{ and } \frac{1}{\beta} \text{fcc}(G) < y^\top \mathbb{1},$$

where $y \in \mathbb{R}_+^{\mathcal{P}(V)}$ is defined as in (84).

Proof. By (6), there exists a rational number $\zeta \in (-1, 0)$ such that

$$(93) \quad \alpha_{\text{GW}} \leq \frac{2}{1 - \zeta} \frac{\arccos(\zeta)}{\pi} < \beta.$$

Let $a \in \mathbb{N}$ be such that $b := (1 - \zeta)\frac{a}{2}$ is an even natural number, so that

$$\zeta = 1 - 2b/a.$$

We now prove the statement holds for $G := \text{Ham}_=(a, b)$.

Let $U \in \mathbb{R}^{[a] \times V(G)}$ be defined by $Ue_s := 2s - \mathbb{1} \in \mathbb{R}^{[a]}$ for every $s \in V(G) = \{0, 1\}^{[a]}$. We claim that

$$(94) \quad (\rho, x) := \left(\frac{b}{a}|E|, \frac{1}{2}\left(\frac{a-1}{b-1}\right)\mathbb{1}\right) \text{ and } (\mu, Y) := \left(\frac{a}{b}, \frac{1}{b}U^\top U\right) \text{ witness the membership } (\mathbb{1}, \mathbb{1}) \in H(G).$$

Feasibility of $(\frac{a}{b}, \frac{1}{b}U^\top U)$ in (11) with $z := \mathbb{1}$ follows directly from the definition of U and G , since

$$(95) \quad Y_{ij} = \frac{a}{b}\zeta \quad \text{for every } ij \in E,$$

whereas feasibility of (ρ, x) in (4b) with $w := \mathbb{1}$ follows from (91), since $L(G) = \binom{a}{b}I - A(G)$ as G is $\binom{a}{b}$ -regular. It is immediate that $\rho\mu = |E|$. Thus, (94) is proved. We claim that

$$(96) \quad (\mathbb{1}, \mathbb{1}) \text{ is an exact pairing.}$$

Note that, in particular, (96) ensures that $\eta(G) = \text{mc}(G)$ and $\eta^\circ(G) = \text{fcc}(G)$. For every $i \in [a]$, set $S_i := \{s \in \{0, 1\}^a : s_i = 1\}$. One may easily check that $|\delta(S_i)| = \binom{a-1}{b-1}2^{a-1}$ for every $i \in [a]$. By (94) and (5), all such cuts are maximum, since

$$\eta(G) = \frac{b}{a}|E| = \frac{b}{a}\binom{a}{b}2^{a-1} = \binom{a-1}{b-1}2^{a-1} = |\delta(S_i)| \leq \text{mc}(G) \leq \eta(G).$$

Now consider the fractional cut cover $\bar{y} := \frac{1}{b} \sum_{i \in [a]} e_{S_i}$. By definition of $\text{Ham}_=(a, b)$, each edge belongs to b of the cuts in $\{\delta(S_i) : i \in [a]\}$, so $\sum_{S \subseteq V} \bar{y}_S \mathbb{1}_{\delta(S)} = \mathbb{1}$. Hence (94) and (13) imply that \bar{y} is an optimal fractional cut cover, as

$$\frac{a}{b} = \eta^\circ(G) \leq \text{fcc}(G) \leq \mathbb{1}^\top \bar{y} = \frac{a}{b}.$$

This concludes the proof of (96). By (15), (95), and linearity of expectation,

$$\mathbb{E}(|\delta(\text{GW}(Y))|) = \sum_{ij \in E} \mathbb{P}(ij \in \delta(\text{GW}(Y))) = \frac{\arccos(\zeta)}{\pi} |E|.$$

Let $y, p \in \mathbb{R}_+^{\mathcal{P}(V)}$ and $\bar{\nu} \in \mathbb{R}_+$ be defined from (μ, Y) as in (84). Then for every $ij \in E$,

$$e_{ij}^\top \left(\sum_{S \subseteq V} p_S \mathbb{1}_{\delta(S)} \right) = \mathbb{P}(ij \in \delta(\text{GW}(Y))) = \frac{\arccos(\zeta)}{\pi}$$

by (15). Hence $\bar{\nu} = \pi / \arccos(\zeta)$. Since $a/b = 2/(1 - \zeta)$, we have that

$$\frac{\mathbb{E}(|\delta(\text{GW}(Y))|)}{\text{mc}(G)} = \frac{2}{1 - \zeta} \frac{\arccos(\zeta)}{\pi} = \frac{\text{fcc}(G)}{\mathbb{1}^\top y} < \beta.$$

and this number is $< \beta$ by (93). \square

Karloff [25] studied a family of graphs closely related to the one featured in the proof of Proposition 30, using it to bound the quality of the approximation factor obtained by Goemans and Williamson [19]. The specific construction in our proof is a simplification of this work due to Alon and Sudakov [1]. The relevance of Proposition 30 to our algorithms inherits several aspects of the relevance of these examples to the maximum cut setting. Let β be a real number such that $\beta > \alpha_{\text{GW}}$. Proposition 30 presents an obstruction for the use of the randomized hyperplane technique to produce β -certificates. Let $G = (V, E)$ be a graph whose existence is ensured by the proposition. Then G defines instances where the SDP relaxations (4a) and (11) are tight — i.e., the integrality ratio is one —, but where the rounding itself can be responsible for the approximation factor α_{GW} of the final algorithm. The first inequality in (92) states that the expected value of a cut produced by the random hyperplane technique will be too small for the desired approximation factor β . The second inequality in (92) does not directly translate to the setting of our algorithm, which takes finitely many samples and outputs a surrogate for the $y \in \mathbb{R}_+^{\mathcal{P}(V)}$ defined in (84). Proposition 30 and (87) do imply that, almost surely, the objective value μ_t/t obtained from (86) deteriorates above $\frac{1}{\beta} \text{fcc}(G)$ for sufficiently large values of $t \in \mathbb{N}$. Note, however, that similar to how (92) allows for a finite number of samples to define a fractional cut cover with objective value better than $\frac{1}{\beta} \text{fcc}(G)$, it also allows for a single cut to have objective value better than $\beta \text{mc}(G)$.

Let $G = (V, E)$ be a graph and let $w \in \mathbb{R}_+^E$. Propositions 29 and 30 describe limitations of the approach we chose for producing approximation algorithms for the maximum cut and fractional cut-covering problems. The *Unique Games Conjecture* [27] implies that it is NP-hard to compute, given an instance (G, w) of the maximum cut problem as input, an upper bound to $\text{mc}(G, w)$ with a better approximation guarantee than (5). Recall from Remark 10 that this conjectured optimality extends to the SDP in (11) and the fractional cut-covering problem: obtaining any better approximation factor on the objective value for Algorithm 1 would disprove the conjecture.

6. CONCLUDING REMARKS

We summarize the main contributions of this paper and discuss future directions. Key to our contributions is precisely describing the relationship between the (weighted) maximum cut and the weighted fractional cut-covering problems through the lens of gauge duality: the functions $\text{mc}(G, \cdot)$ and $\text{fcc}(G, \cdot)$ form a gauge dual pair. Crucially, the SDP relaxation for the maximum cut problem, utilized by Goemans and Williamson [19], and the SDP relaxation we provided for the fractional cut-covering problem share the same property: $\eta(G, \cdot)$ and $\eta^\circ(G, \cdot)$ form a gauge dual pair (see Proposition 7). This explicit connection establishes the background and foundation for the development of our algorithms. Gauge duality promptly yields a bound conversion procedure, enabling us to extend the α_{GW} approximation ratio between $\text{mc}(G, \cdot)$ and $\eta(G, \cdot)$ to a $1/\alpha_{\text{GW}}$ approximation ratio between $\text{fcc}(G, \cdot)$ and $\eta^\circ(G, \cdot)$ (see Proposition 9).

The understanding of this connection organizes our efforts in algorithmic design:

- (i) Optimal solutions $Y \in \mathbb{S}_+^V$ of the SDP relaxation (4a) for maximum cut are employed in [19] to sample (the shore of) a cut by the randomized hyperplane technique, with Y as the generating parameter. We extend this technique to the fractional cut-covering problem in two stages: initially by showing that the marginal probabilities for the shores provide a fractional cut cover with the same approximation quality, and subsequently by sparsifying our cover through a polynomial-time randomized procedure (see Propositions 1 and 4 and Theorem 6).
- (ii) We address the problem of simultaneously obtaining approximate solutions for the maximum cut and the weighted fractional cut-covering problems with certificates of approximate optimality. This inspires our definition of the set $H(G)$, which precisely links tightness of $w^\top z \leq \eta(G, w)\eta^\circ(G, z)$ in gauge duality with the optimality of solutions for the SDP relaxations. This directly motivates the definition of β -pairings and β -certificates, which are essential to our simultaneous certification process (see Proposition 15, and Theorems 18 and 19).
- (iii) Given the strong algorithmic focus of our work, it is crucial to deal with the fact that exact optimal solutions for SDPs may not be computable in polynomial time. We combine the strategies from the previous items: in Section 3.4, we introduce the set $H_{\epsilon, \sigma}(G)$, which is a perturbed version of $H(G)$, and then develop a randomized approximation algorithm which requires only nearly optimal solutions for the SDPs.
- (iv) When run on a connected graph with n vertices and m edges, our algorithms rely on obtaining a nearly optimal solution Y to an instance of the relevant SDP problems, which can be done in $O(m^4)$ time, followed by the rounding procedure on Y , which involves a single $n \times n$ Cholesky factorization, $O(n \ln(n))$ samples from a standard Gaussian distributions, and $O(n^2 \ln(n))$ extra work.
- (v) Section 5 addresses many aspects of our approach that are optimal. Many of them are mirror images under gauge duality of corresponding properties for the Goemans and Williamson's algorithm for maximum cut (see Proposition 30).

Throughout our work we have assumed the real-number machine model [7] with two extra oracles: one sampling from a standard normal distribution, and one computing Cholesky factorizations. This second assumption is equivalent to assuming exact computation of square roots of real positive numbers. Despite these assumptions, our results build towards an implementation in the Turing machine model. Explicitly, the Slater points for (58) used in Appendix C may allow one to utilize the work of de Klerk and Vallentin [10] to conclude that both η_ϵ and η_ϵ° may be computed in polynomial time on a Turing machine. Proposition 22 identifies the main computational steps required to round these optimal solutions, and it may be extended to the probabilistic Turing machine model with an appropriate analysis of approximate square root computations.

The interplay of duality with probabilistic aspects that permeates this work suggests interesting avenues for future research. Note that, essential to our approach, Goemans and Williamson's randomized hyperplane technique casts an optimal solution Y for (4a) as a *distribution* of cuts. This enables edges that do not occur in any heavy cut to be probably covered after appropriate thickening. The SDP perturbation (58b) makes this even more robust, since for any feasible solution (μ, Y) , sampling from $\text{GW}(Y)$ covers every edge with probability at least $\sqrt{2\epsilon}/\pi$ (by (62)) and Y has full rank. The analysis in [19] for maximum cut relies on expected values, whereas for fractional cut covers we rely on concentration inequalities (see Propositions 4 and 21). In this respect, it would be interesting to find a gauge dual analogue of the analysis in [19, Sec. 3.1], which improves the approximation factor α_{GW} when the maximum cut is large. A similar question can be

made about a gauge dual analogue of low-rank SDP solutions with improved approximation factors, as in [4]. One may also ask whether the techniques from [32] can be applied to derandomize our algorithms.

Further natural research directions include the use of duality in close relatives of the maximum cut problem, such as maximum bisection problem (see, e.g., [3, 37]) and Nesterov’s generalization to quadratic optimization problems [35]. Beyond the maximum cut and fractional cut-covering problems, it is natural to search for and study other pairs of combinatorial optimization problems that are linked by gauge duality. For example, the literature has both time-honored [20] and more recent [6] publications leveraging the fact that the stability number and the fractional chromatic number of a graph define gauges dual to each other. Specially in cases where an approximation algorithm rounds a solution to a convex relaxation of a combinatorial problem, the (gauge) dual combinatorial problem may be approximated utilizing the ideas within this work. On a broader note, it seems interesting to look for further pair of problems which can be “simultaneously approximated”. Definition 12 provides a convenient formalization of simultaneous approximation for this work, but even within the maximum cut and fractional cut-covering context there are potentially sensible alternative definitions, as mentioned in Remark 13.

REFERENCES

- [1] N. Alon and B. Sudakov. “Bipartite subgraphs and the smallest eigenvalue”. In: *Combin. Probab. Comput.* 9.1 (2000), pages 1–12.
- [2] A. Y. Aravkin, J. V. Burke, D. Drusvyatskiy, M. P. Friedlander, and K. J. MacPhee. “Foundations of gauge and perspective duality”. In: *SIAM J. Optim.* 28.3 (2018), pages 2406–2434.
- [3] P. Austrin, S. Benabbas, and K. Georgiou. “Better balance by being biased: a 0.8776-approximation for Max Bisection”. In: *ACM Trans. Algorithms* 13.1 (2016), Art. 2, 27.
- [4] A. Avidor and U. Zwick. “Rounding two and three dimensional solutions of the SDP relaxation of MAX CUT”. In: *International Workshop on Approximation Algorithms for Combinatorial Optimization*. Springer, 2005, pages 14–25.
- [5] N. Benedetto Proença. “Combinatorial and geometric dualities in graph homomorphism optimization problems”. Master’s thesis. Universidade de São Paulo, July 2021. xii+173 pages. URL: <https://www.teses.usp.br/teses/disponiveis/45/45134/tde-23092021-183204/> (visited on 06/27/2023).
- [6] N. Benedetto Proença, M. K. de Carli Silva, and G. Coutinho. “Dual Hoffman bounds for the stability and chromatic numbers based on semidefinite programming”. In: *SIAM J. Discrete Math.* 35.4 (2021), pages 2880–2907.
- [7] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and Real Computation*. New York, NY: Springer, 1998. xvi+453 pages.
- [8] M. K. de Carli Silva and L. Tunçel. “Optimization problems over unit-distance representations of graphs”. In: *Electron. J. Combin.* 20.1 (2013), Paper 43, 20pp.
- [9] M. K. de Carli Silva and L. Tunçel. “Vertices of spectrahedra arising from the elliptope, the theta body, and their relatives”. In: *SIAM J. Optim.* 25.1 (2015), pages 295–316.
- [10] E. de Klerk and F. Vallentin. “On the Turing model complexity of interior point methods for semidefinite programming”. In: *SIAM J. Optim.* 26.3 (2016), pages 1944–1961.
- [11] C. Delorme and S. Poljak. “Laplacian eigenvalues and the maximum cut problem”. In: *Math. Programming* 62.3, Ser. A (1993), pages 557–574.
- [12] M. DeVos, J. Nešetřil, and A. Raspaud. “On edge-maps whose inverse preserves flows or tensions”. In: *Graph theory in Paris*. Trends Math. Birkhäuser, Basel, 2007, pages 109–138.
- [13] U. Feige and G. Schechtman. “On the optimality of the random hyperplane rounding technique for MAX CUT”. In: *Random Structures & Algorithms* 20.3 (2002), pages 403–440.
- [14] R. M. Freund. “Dual gauge programs, with applications to quadratic programming and the minimum-norm problem”. In: *Math. Programming* 38.1 (1987), pages 47–67.
- [15] M. P. Friedlander, I. Macêdo, and T. K. Pong. “Gauge optimization and duality”. In: *SIAM J. Optim.* 24.4 (2014), pages 1999–2022.
- [16] D. R. Fulkerson. “Anti-blocking polyhedra”. In: *J. Combinatorial Theory Ser. B* 12 (1972), pages 50–71.
- [17] D. R. Fulkerson. “Blocking and anti-blocking pairs of polyhedra”. In: *Math. Programming* 1 (1971), pages 168–194.
- [18] C. Godsil and G. Royle. *Algebraic graph theory*. Volume 207. Graduate Texts in Mathematics. Springer-Verlag, New York, 2001. xx+439 pages.

- [19] M. X. Goemans and D. P. Williamson. “Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming”. In: *J. ACM* 42.6 (1995), pages 1115–1145.
- [20] M. Grötschel, L. Lovász, and A. Schrijver. “Relaxations of vertex packing”. In: *J. Combinatorial Theory Ser. B* 40.3 (1986), pages 330–343.
- [21] M. Grötschel, L. Lovász, and A. Schrijver. “The ellipsoid method and its consequences in combinatorial optimization”. In: *Combinatorica* 1.2 (1981), pages 169–197.
- [22] M. Grötschel and W. R. Pulleyblank. “Weakly bipartite graphs and the max-cut problem”. In: *Oper. Res. Lett.* 1.1 (1981), pages 23–27.
- [23] S. B. Hopkins, T. Schramm, and L. Trevisan. “Subexponential LPs approximate Max-Cut”. In: *2020 IEEE 61st Annual Symposium on Foundations of Computer Science*. IEEE Computer Soc., Los Alamitos, CA, 2020, pages 943–953.
- [24] D. Karger, R. Motwani, and M. Sudan. “Approximate graph coloring by semidefinite programming”. In: *J. ACM* 45.2 (1998), pages 246–265.
- [25] H. Karloff. “How good is the Goemans-Williamson MAX CUT algorithm?” In: *SIAM J. Comput.* 29.1 (1999), pages 336–350.
- [26] R. M. Karp. “Reducibility among combinatorial problems”. In: *Complexity of computer computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972)*. Plenum, New York, 1972, pages 85–103.
- [27] S. Khot, G. Kindler, E. Mossel, and R. O’Donnell. “Optimal inapproximability results for MAX-CUT and other 2-variable CSPs?” In: *SIAM J. Comput.* 37.1 (2007), pages 319–357.
- [28] M. Laurent, S. Poljak, and F. Rendl. “Connections between semidefinite relaxations of the max-cut and stable set problems”. In: *Math. Programming* 77.2, Ser. B (1997). Semidefinite programming, pages 225–246.
- [29] L. Lovász. *Graphs and Geometry*. Colloquium Publications. American Mathematical Society, 2019. 444 pages.
- [30] L. Lovász. “Semidefinite programs and combinatorial optimization”. In: *Recent advances in algorithms and combinatorics*. Volume 11. CMS Books Math./Ouvrages Math. SMC. Springer, New York, 2003, pages 137–194.
- [31] L. Lovász. “On the Shannon capacity of a graph”. In: *IEEE Trans. Inform. Theory* 25.1 (1979), pages 1–7.
- [32] S. Mahajan and H. Ramesh. “Derandomizing approximation algorithms based on semidefinite programming”. In: *SIAM J. Comput.* 28.5 (1999), pages 1641–1663.
- [33] R. J. McEliece, E. R. Rodemich, and H. C. Rumsey Jr. “The Lovász bound and some generalizations”. In: *Journal of Combinatorics, Information & System Sciences* 3.3 (1978), pages 134–152.
- [34] A. Nemirovski. *Introduction to Linear Optimization*. 2012. URL: https://www2.isye.gatech.edu/~nemirovs/OPTI_LectureNotes.pdf (visited on 10/10/2019).
- [35] Y. Nesterov. “Semidefinite relaxation and nonconvex quadratic optimization”. In: *Optim. Methods Softw.* 9.1-3 (1998), pages 141–160.
- [36] J. Neto and W. Ben-Ameur. “On fractional cut covers”. In: *Discrete Appl. Math.* 265 (2019), pages 168–181.
- [37] S. Poljak and F. Rendl. “Nonpolyhedral relaxations of graph-bisection problems”. In: *SIAM J. Optim.* 5.3 (1995), pages 467–487.
- [38] R. Šámal. “Cubical coloring—fractional covering by cuts and semidefinite programming”. In: *Discrete Math. Theor. Comput. Sci.* 17.2 (2015), pages 251–266.
- [39] R. Šámal. “On XY Mappings”. PhD thesis. Charles University, 2006. vi+146 pages. URL: <https://dSPACE.cuni.cz/handle/20.500.11956/6160> (visited on 06/27/2023).
- [40] A. Schrijver. “A comparison of the Delsarte and Lovász bounds”. In: *Institute of Electrical and Electronics Engineers. Transactions on Information Theory* 25.4 (1979), pages 425–429.
- [41] L. Trevisan. “Max cut and the smallest eigenvalue”. In: *SIAM J. Comput.* 41.6 (2012), pages 1769–1786.
- [42] L. Tunçel. *Polyhedral and semidefinite programming methods in combinatorial optimization*. Volume 27. Fields Institute Monographs. Providence, RI: American Mathematical Society, 2010. x+219 pages.

APPENDIX A. EDGES WITH ARBITRARILY SMALL PROBABILITY OF BEING COVERED

Proposition 31. Let $\varepsilon \in (0, 2)$. Set $G := K_3$, set $\beta := 4/(4 - \varepsilon)^2$, and set

$$(97a) \quad w := (2 - \varepsilon)\beta(e_{12} + e_{13}) + \beta e_{23} \in \mathbb{R}_+^{E(G)},$$

$$(97b) \quad z := e_{12} + e_{13} + \varepsilon e_{23} \in \mathbb{R}_+^{E(G)}.$$

Then

$$(98) \quad (1, \bar{x}) \text{ and } (\bar{\mu}, \bar{Y}) \text{ witness the membership } (w, z) \in H(G),$$

where

$$\bar{x} := \frac{1}{4 - \varepsilon}((2 - \varepsilon)e_1 + e_2 + e_3), \quad \bar{\mu} := \frac{4}{4 - \varepsilon}, \quad \text{and} \quad \bar{Y} := \bar{\mu} \begin{bmatrix} 1 & \varepsilon/2 - 1 & \varepsilon/2 - 1 \\ \varepsilon/2 - 1 & 1 & 1 - 2\varepsilon + \varepsilon^2/2 \\ \varepsilon/2 - 1 & 1 - 2\varepsilon + \varepsilon^2/2 & 1 \end{bmatrix}.$$

In particular, $(\bar{\mu}, \bar{Y})$ is an optimal solution for the SDP (11).

Proof. It is immediate that $\text{diag}(\bar{Y}) = \bar{\mu} \mathbf{1}$. The direct computation

$$\frac{(4 - \varepsilon)}{4} \bar{Y} = \begin{bmatrix} 1 \\ \varepsilon/2 - 1 \\ \varepsilon/2 - 1 \end{bmatrix} \begin{bmatrix} 1 & \varepsilon/2 - 1 & \varepsilon/2 - 1 \end{bmatrix} + (\varepsilon - \varepsilon^2/4) \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}$$

shows that $\bar{Y} \in \mathbb{S}_+^V$. Moreover, from (12) we get that

$$(\frac{1}{4} \mathcal{L}_G^*(\bar{Y}))_{13} = (\frac{1}{4} \mathcal{L}_G^*(\bar{Y}))_{12} = \frac{2}{4 - \varepsilon} - \frac{\varepsilon - 2}{4 - \varepsilon} = 1 = z_{12} = z_{13}.$$

Furthermore

$$(\frac{1}{4} \mathcal{L}_G^*(\bar{Y}))_{23} = \frac{2}{4 - \varepsilon} - \frac{2(1 - 2\varepsilon + \varepsilon^2/2)}{4 - \varepsilon} = \frac{4\varepsilon - \varepsilon^2}{4 - \varepsilon} = \varepsilon = z_{23}.$$

Thus, $(\bar{\mu}, \bar{Y})$ is feasible in (11) for (G, z) . Set $u := (2 - \varepsilon)e_1 + e_2 + e_3$. Moreover

$$\begin{aligned} \frac{1}{4} \mathcal{L}_G(w) &= \frac{1}{4}(2 - \varepsilon)\beta \mathcal{L}_G(e_{12} + e_{13}) + \frac{1}{4}\beta \mathcal{L}_G(e_{23}) && \text{by (97a)} \\ &= \frac{1}{4}\beta \left((2 - \varepsilon) \begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \right) \\ &= \frac{1}{4}\beta \begin{bmatrix} 4 - 2\varepsilon & \varepsilon - 2 & \varepsilon - 2 \\ \varepsilon - 2 & 3 - \varepsilon & -1 \\ \varepsilon - 2 & -1 & 3 - \varepsilon \end{bmatrix} \\ &= \frac{1}{4}\beta \left((4 - \varepsilon) \begin{bmatrix} 2 - \varepsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} (2 - \varepsilon)^2 & 2 - \varepsilon & 2 - \varepsilon \\ 2 - \varepsilon & 1 & 1 \\ 2 - \varepsilon & 1 & 1 \end{bmatrix} \right) \\ &= \frac{1}{4}\beta((4 - \varepsilon) \text{Diag}(u) - uu^\top). && \text{by the definition of } u \\ &\preceq \frac{1}{4}\beta(4 - \varepsilon) \text{Diag}(u) \\ &= \text{Diag}(\bar{x}) && \text{since } 4\bar{x} = \beta(4 - \varepsilon)u. \end{aligned}$$

Thus $(1, \bar{x})$ is feasible in (4b) for (G, w) as $\mathbf{1}^\top \bar{x} = 1$. Since

$$w^\top z = 2(2 - \varepsilon)\beta + \varepsilon\beta = \beta(4 - \varepsilon) = \frac{4}{4 - \varepsilon} = \bar{\mu},$$

we obtain (98) from (47). Note that since $(1, \bar{x})$ is feasible in (4b), then (w, \bar{x}) is feasible in (26b). Since (11) and (26b) form a primal-dual pair of SDPs, $\bar{\mu} = w^\top z$ implies optimality of $(\bar{\mu}, \bar{Y})$ in (11). \square

APPENDIX B. SDP SOLUTIONS DEFINING FRACTIONAL CUT COVERS WITH EXPONENTIAL SUPPORT

Theorem 32 below presents a rigorous statement and proof regarding (85), i.e., about the exponential support of a fractional cut cover derived from Proposition 1.

Theorem 32. Let $n \geq 3$ be an integer. Set $\bar{Y} := 2I - \frac{2}{n}\mathbb{1}\mathbb{1}^\top$. For every $i \in [n]$, let g_i be independently sampled from the standard normal distribution, and set $h := \|g\|^{-1}g$. Then, for each nonempty $S \subsetneq V$,

$$\mathbb{P}(\text{GW}(\bar{Y}, h) = S) > 0$$

Proof. Let $S \subsetneq V$ be nonempty. Set

$$\beta := \min \left\{ \frac{2n - |S| - 1}{|S| - 1}, \frac{n + |S| - 1}{n - |S| - 1} \right\},$$

where we consider $1/0$ to be $+\infty$. It is straightforward to check that $\beta > 1$. For each $v \in V$, let A_v be the event that $1 < g_v < \beta$ and let B_v be the event that $-\beta < g_v < -1$. Since g_v is sampled from the standard normal distribution, we have that $\mathbb{P}(A_v) = \mathbb{P}(B_v) =: p$ is a positive constant depending only on β .

Set

$$D := \left(\bigcap_{s \in S} A_s \right) \cap \left(\bigcap_{t \in V \setminus S} B_t \right).$$

Note that, by the independence of $(g_v)_{v \in V}$, we have that $\mathbb{P}(D) = p^n > 0$. We complete the proof by showing that D implies that

$$(99) \quad \text{GW}(\bar{Y}, h) = S.$$

We will use the fact that the sampled shore is

$$\begin{aligned} \text{GW}(\bar{Y}, h) &= \{i \in V : e_i^\top \bar{Y}^{1/2} h \geq 0\} = \{i \in V : e_i^\top (I - \tfrac{1}{n}\mathbb{1}\mathbb{1}^\top) h \geq 0\} \\ (100) \quad &= \left\{ i \in V : h_i \geq \frac{\mathbb{1}^\top h}{n} \right\} = \left\{ i \in V : g_i \geq \frac{\mathbb{1}^\top g}{n} \right\} \\ &= \left\{ i \in V : (n-1)g_i \geq \sum_{v \in V \setminus \{i\}} g_v \right\}, \end{aligned}$$

Assume that D holds. We first prove ‘ \supseteq ’ in (99). Let $s \in S$. Then

$$\frac{1}{n-1} \sum_{v \in V \setminus \{s\}} g_v = \frac{1}{n-1} \left(\sum_{s' \in S \setminus \{s\}} g_{s'} + \sum_{t \in V \setminus S} g_t \right) < \frac{1}{n-1} (\beta(|S| - 1) - (n - |S|)) \leq 1 < g_s,$$

since our choice of β ensures that $\beta(|S| - 1) \leq 2n - |S| - 1$. Hence, $s \in \text{GW}(Y^*, h)$ by (100).

We now prove ‘ \subseteq ’ in (99). Let $t \in V \setminus S$. Then

$$\frac{1}{n-1} \sum_{v \in V \setminus \{t\}} g_v = \frac{1}{n-1} \left(\sum_{t' \in V \setminus (S \cup \{t\})} g_{t'} + \sum_{s \in S} g_s \right) > \frac{1}{n-1} (-\beta(n - |S| - 1) + |S|) \geq -1 > g_t,$$

since our choice of β ensures that $\beta(n - |S| - 1) \leq n + |S| - 1$. Thus, $t \notin \text{GW}(Y^*, h)$ by (100). \square

APPENDIX C. SDP SOLVERS

In this section, we analyze the running time of an interior-point method (IPM) to solve the SDP (58b) to near optimality:

$$(58b) \quad \eta_\varepsilon^\circ(G, z) = \min \{ \mu : \mu \in \mathbb{R}_+, Y \in \mathbb{S}^V, Y \succeq \mu \varepsilon I, \tfrac{1}{4} \mathcal{L}_G^*(Y) \geq z, \text{diag}(Y) = \mu \mathbb{1} \}.$$

Algorithms 1 and 4 in Theorems 6 and 24, resp., rely on obtaining such nearly optimal solutions for (58b), where the parameter ε lies in $[0, 1)$.

For the purposes of stating the running time of IPMs, we will consider an arbitrary primal SDP in the format

$$\begin{aligned} (101a) \quad & \text{Minimize} \quad \langle c, x \rangle \\ (101b) \quad & \text{subject to} \quad \mathcal{A}(x) \succeq_{\mathbb{L}^*} b, \\ (101c) \quad & x \in \mathbb{K}, \end{aligned}$$

and its (syntactically symmetric) dual,

$$\begin{aligned}
 (102a) \quad & \text{Maximize} \quad \langle b, y \rangle \\
 (102b) \quad & \text{subject to} \quad y \in \mathbb{L}, \\
 (102c) \quad & \mathcal{A}^*(y) \preceq_{\mathbb{K}^*} c,
 \end{aligned}$$

where

$$(103a) \quad \mathbb{K} := \mathbb{S}_+^{n_1} \oplus \mathbb{R}_+^{n_2} \oplus \mathbb{R}^{n_3} \subseteq \mathbb{S}^{n_1} \oplus \mathbb{R}^{n_2} \oplus \mathbb{R}^{n_3} =: \mathbb{X}$$

and

$$(103b) \quad \mathbb{L} := \mathbb{S}_+^{m_1} \oplus \mathbb{R}_+^{m_2} \oplus \mathbb{R}^{m_3} \subseteq \mathbb{S}^{m_1} \oplus \mathbb{R}^{m_2} \oplus \mathbb{R}^{m_3} =: \mathbb{Y}$$

for some integers $n_1, n_2, n_3, m_1, m_2, m_3 \in \mathbb{N}$, the map $\mathcal{A}: \mathbb{X} \rightarrow \mathbb{Y}$ is linear, $b \in \mathbb{Y}$, and $c \in \mathbb{X}$. Here, the *dual cone* of a cone C in Euclidean space \mathbb{E} is $C^* := \{y \in \mathbb{E} : \forall x \in C, \langle y, x \rangle \geq 0\}$, and each of the notations $a \succeq_C b$ and $b \preceq_C a$, with $a, b \in \mathbb{E}$, means that $a - b \in C$.

Note that (101c) and (103a) allow for variables composed of positive semidefinite matrices, nonnegative vectors, and free scalar variables. Similarly, (101b) and (103b) enables one to require affine functions of the variables to be positive semidefinite, nonnegative, or *equal* to zero. As an example, the SDP (58b) can be cast in the format (101) by setting

$$\begin{aligned}
 \mathbb{K} &:= \mathbb{S}_+^0 \oplus \mathbb{R}_+^1 \oplus \mathbb{S}^V, & \mathbb{L} &:= \mathbb{S}_+^V \oplus \mathbb{R}_+^E \oplus \mathbb{R}^V, \\
 \mathcal{A}: \mu \oplus Y \in \mathbb{R}^1 \oplus \mathbb{S}^V &\mapsto (Y - \mu \varepsilon I) \oplus \left(\frac{1}{4} \mathcal{L}_G^*(Y)\right) \oplus (\mu \mathbb{1} - \text{diag}(Y)), \\
 b &:= 0 \oplus z \oplus 0, & c &:= 1 \oplus 0.
 \end{aligned}$$

Some IPMs proceed by producing a sequence of iterates tracking the so-called central path. We will encode the procedure that updates an iterate to the next one by a function Ξ , so that from an iterate (x_t, y_t) of primal-dual solutions, the next iterate will be $\Xi(x_t, y_t) =: (x_{t+1}, y_{t+1})$.

The number of iterations of IPMs so that the duality gap for a pair of solutions (x_t, y_t) is a δ -fraction of the initial duality gap can be bounded by a function ψ on the initial Slater points $(x_0, y_0) =: (\hat{x}, \hat{y})$. Set

$$(104) \quad N := n_1 + n_2 + m_1 + m_2.$$

For each pair (x, y) of primal-dual feasible solutions, where

$$\begin{aligned}
 x &:= X_1 \oplus x_2 \oplus x_3 \in \mathbb{X}, \\
 y &:= Y_1 \oplus y_2 \oplus y_3 \in \mathbb{Y},
 \end{aligned}$$

and with corresponding slacks

$$\begin{aligned}
 \mathcal{A}(x) - b &:= U_1 \oplus u_2 \oplus u_3 \in \mathbb{Y}, \\
 c - \mathcal{A}^*(y) &:= V_1 \oplus v_2 \oplus v_3 \in \mathbb{X},
 \end{aligned}$$

define

$$\begin{aligned}
 (105) \quad \psi(x, y) &:= N \ln \left(\frac{1}{N} \left\langle X_1 \oplus x_2 \oplus Y_1 \oplus y_2, V_1 \oplus v_2 \oplus U_1 \oplus u_2 \right\rangle \right) \\
 &\quad - \ln \left(\det(X_1) \det(V_1) \det(Y_1) \det(U_1) \left(\prod x_2 \right) \left(\prod v_2 \right) \left(\prod y_2 \right) \left(\prod u_2 \right) \right),
 \end{aligned}$$

where, for each vector a , we denote $\prod a := \det(\text{Diag}(a))$. The function ψ takes into account two important factors that affect the number of required iterations: the first term in the RHS of (105) depends on the initial duality gap $\langle X_1 \oplus x_2 \oplus Y_1 \oplus y_2, V_1 \oplus v_2 \oplus U_1 \oplus u_2 \rangle$, whereas the second term is related to the centrality of the initial Slater point. Indeed, it is intuitive that in IPMs, a good initial point should have a reasonably good duality gap and at the same time not being too close to the boundary of the feasible region.

The next result is adapted from [42, Theorem 4.5] for our format of SDPs:

Theorem 33. Let $\delta \in (0, 1)$ and let (x_0, y_0) be a primal-dual pair of feasible solutions for (101) and (102), respectively, such that

$$\psi(x_0, y_0) \leq \sqrt{N} \ln(1/\delta),$$

where N is defined as in (104) and ψ as in (105). Define the sequence $(x_t, y_t)_{t=0}^\infty$ by $(x_{t+1}, y_{t+1}) := \Xi(x_t, y_t)$ for each $t \in \mathbb{N}$. Define the sequence $(u_t, v_t)_{t=0}^\infty$ by $u_t := \mathcal{A}(x_t) - b$ and $v_t := c - \mathcal{A}^*(y_t)$ for each $t \in \mathbb{N}$. Then

$$\langle x_t \oplus u_t, v_t \oplus y_t \rangle \leq \delta \langle x_0 \oplus u_0, v_0 \oplus y_0 \rangle \quad \text{for each } t \geq \bar{t} := 24\sqrt{N} \ln(1/\delta).$$

Proposition 34. Let $\sigma \in (0, 2/3)$. There exists a polynomial-time algorithm that, given a graph $G = (V, E)$ with n vertices and m edges, a nonzero vector $z \in \mathbb{R}_+^E$, and a number $\varepsilon \in [0, 1/3]$, computes feasible solutions (x, w) for (58a) and (μ, Y) for (58b) such that $\mu \leq z^\top w + \sigma \|z\|_\infty$. The algorithm consists of applying an interior-point method for $T := 24(n + m + 1) \ln(8/\sigma)$ iterations; each iteration encoding one application of the function Ξ can be made to run in time $O((n + m)^3)$.

Remark. Note that, since σ is constant, the SDPs (58) can be nearly solved (in the sense of the $\sigma \|z\|_\infty$ additive error) in strongly polynomial time.

Proof. First we write (58b) and its dual. We scale the cost function of (58b) by $N := n + m + 1$, and we normalize the edge weights by setting $\bar{z} := z / \|z\|_\infty$:

$$\begin{array}{ll} \text{(P)} & \inf \quad N\mu \\ & \text{subject to} \quad Y - \mu \varepsilon I \succeq 0, \\ & \quad \frac{1}{4} \mathcal{L}_G^*(Y) \geq \bar{z}, \\ & \quad -\text{diag}(Y) + \mu \mathbb{1} = 0, \\ & \quad Y \in \mathbb{S}^V, \\ & \quad \mu \in \mathbb{R}_+, \\ \text{(D)} & \sup \quad \bar{z}^\top w \\ & \text{subject to} \quad S \in \mathbb{S}_+^V, \\ & \quad w \in \mathbb{R}_+^E, \\ & \quad x \in \mathbb{R}^V, \\ & \quad S + \frac{1}{4} \mathcal{L}_G(w) - \text{Diag}(x) = 0, \\ & \quad \mathbb{1}^\top x - \varepsilon \text{Tr}(S) \leq N. \end{array}$$

Set

$$\begin{aligned} \dot{\mu} &:= 4, & \dot{Y} &:= 4I, \\ \dot{w} &:= \mathbb{1}, & \dot{x} &:= \frac{1}{2} \deg_G + \mathbb{1}, & \dot{S} &:= I + \frac{1}{4} D(G) + \frac{1}{4} A(G), \end{aligned}$$

where $D(G) := \text{Diag}(\deg_G)$. It is straightforward to check that $\dot{Y} \oplus \dot{\mu}$ and $\dot{S} \oplus \dot{w} \oplus \dot{x}$ are Slater points for (P) and (D), resp., with corresponding slacks

$$\begin{aligned} \dot{U} &:= \dot{Y} - \dot{\mu} \varepsilon I = 4(1 - \varepsilon)I, \\ \dot{u} &:= \frac{1}{4} \mathcal{L}_G^*(\dot{Y}) - \bar{z} = 2\mathbb{1} - \bar{z}, \\ \dot{v} &:= N - \mathbb{1}^\top \dot{x} + \varepsilon \text{Tr}(\dot{S}) = 1 + \frac{\varepsilon}{2}m + \varepsilon n. \end{aligned}$$

The duality gap between $\dot{Y} \oplus \dot{\mu}$ and $\dot{S} \oplus \dot{w} \oplus \dot{x}$ is

$$\begin{aligned} \dot{\mu} \dot{v} + \langle \dot{U}, \dot{S} \rangle + \dot{u}^\top \dot{w} &= 4(1 + \frac{\varepsilon}{2}m + \varepsilon n) + 4(1 - \varepsilon) \text{Tr}(I + \frac{1}{4} D(G) + \frac{1}{4} A(G)) + 2\mathbb{1}^\top \mathbb{1} - \mathbb{1}^\top \bar{z} \\ (106) \quad &\leq 4(1 + \frac{\varepsilon}{2}m + \varepsilon n) + 4n + 4m \leq 4(1 + m/2 + n) + 4n + 4m \\ &\leq 8(1 + m + n) = 8N. \end{aligned}$$

We will now compute an upper bound for the value of the function ψ for $\dot{Y} \oplus \dot{\mu}$ and $\dot{S} \oplus \dot{w} \oplus \dot{x}$. Since we already computed an upper bound for the duality gap in (106), we will now lower bound the determinants:

$$\begin{aligned} \dot{\mu} \cdot \det(\dot{U}) \cdot \left(\prod \dot{u} \right) \cdot \dot{v} \cdot \det(\dot{S}) \cdot \left(\prod \dot{w} \right) \\ = 4 \cdot \det(4(1 - \varepsilon)I) \cdot \prod (2\mathbb{1} - \bar{z}) \cdot (1 + \frac{\varepsilon}{2}m + \varepsilon n) \cdot \det(I + \frac{1}{4} D(G) + \frac{1}{4} A(G)) \cdot \prod \mathbb{1} \\ \geq 4 \det(4(1 - \varepsilon)I) \det(I) \geq 4^n (1 - \varepsilon)^n, \end{aligned}$$

where in the first inequality we used that $\|\bar{z}\|_\infty = 1$ and $I + \frac{1}{4} D(G) + \frac{1}{4} A(G) \succeq I$. Thus,

$$\psi(\dot{\mu} \oplus \dot{Y}, \dot{S} \oplus \dot{w} \oplus \dot{x}) \leq N \ln \left(\frac{1}{N} (8N) \right) - \ln(4^n (1 - \varepsilon)^n) \leq N \ln(8).$$

Set $\delta := (8/\sigma)^{\sqrt{N}} \leq \min\{\sigma/8, 8^{-\sqrt{N}}\}$. By Theorem 33, after $24\sqrt{N} \ln(1/\delta) = O(N)$ iterations of Ξ , the duality gap is at most $8\delta N \leq \sigma N$. That is, we obtain $\tilde{Y} \oplus \tilde{\mu}$ and $\tilde{S} \oplus \tilde{w} \oplus \tilde{x}$, feasible for (P) and (D), resp., such that $N\tilde{\mu} \leq N\tilde{z}^\top \tilde{w} + N\sigma$, that is, $\tilde{\mu} \leq \tilde{z}^\top \tilde{w} + \sigma$. Hence, $\|z\|_\infty \cdot (\tilde{Y} \oplus \tilde{\mu})$ is feasible for (58b) with objective value $\|z\|_\infty \tilde{\mu} \leq z^\top \tilde{w} + \sigma \|z\|_\infty$. \square