EXPLOITING LOW-RANK OBJECTIVE STRUCTURE IN DISCRETE QUADRATIC OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

We study the maximization of a positive semidefinite Hermitian quadratic form over vectors whose entries are K-th roots of unity, for any $K \geq 2$, a problem that contains MAX-CUT ($K = 2$), MAX-3-CUT ($K = 3$), and K-ary phase detection. When the objective matrix has rank r, the problem reduces to a search over directions on the unit sphere of $\mathbb{C}^r$: rounding each coordinate of $\mathbf{Vc}$, where $\mathbf{V}$ is the matrix factor of the given Hermitian matrix, to the nearest root turns a direction $\mathbf{c}$ into a candidate, and some direction yields an optimum. Exact search enumerates $O(rn^{2r-1})$ candidates.

In this paper, we analyze the simplest alternative, which rounds S uniformly random directions and keeps the best. One random direction is $(1-\varepsilon)$-optimal with probability at least $(\varepsilon/2)^{r-1}\min\{1, \frac{K}{\pi}\sqrt{\varepsilon/2}\}$, for every n, so $S = O\big((2/\varepsilon)^{r-1/2}\log(1/\delta)\big)$ samples and $O(Snr)$ time give a $(1-\varepsilon)$-optimal solution with probability $1-\delta$. For noisy objectives $\mathbf{Q}^\star + \mathbf{H}$, rounding a rank-$r$ truncation loses at most $2n(\lambda^\star_{r+1} + 2\|\mathbf{H}\|_2)$, with no eigengap condition. In MAX-3-CUT experiments, the number of samples needed for a near-best cut does not grow with n on random regular graphs with up to 10^6 vertices and on tori with up to 10^5, per our settings; on graphs small enough to solve exactly, 100 samples reach 99.8% of the rank-2 optimum on average; and on planted three-community graphs, whose objective is low rank plus noise, the bound for noisy objectives becomes informative as n grows.

1 INTRODUCTION

Given a Hermitian positive semidefinite (PSD) matrix $\mathbf{Q} \in \mathbb{C}^{n\times n}$ and an integer $K \geq 2$, we study:

$$\mathrm{OPT}(\mathbf{Q}) := \max_{\mathbf{z}\in\mathcal{A}_K^n} \mathbf{z}^\dagger\mathbf{Q}\mathbf{z}, \qquad \mathcal{A}_K := \big\{e^{2\pi\hat{\imath}k/K} : k = 0, 1, \ldots, K-1\big\}, \quad \hat{\imath} := \sqrt{-1}. \tag{1}$$

With a graph Laplacian $\mathbf{Q}$, problem (1) is MAX-CUT for $K = 2$ and MAX-3-CUT for $K = 3$; with a complex $\mathbf{Q}$ it describes maximum-likelihood detection of K-ary phase-shift-keyed signals (Kyrillidis & Karystinos, 2014) and phase synchronization (Singer, 2011). The problem is NP-hard in its general setting. Semidefinite relaxations give the best worst-case ratios, 0.878 for MAX-CUT (Goemans & Williamson, 1995) and 0.836 for MAX-3-CUT (Goemans & Williamson, 2004), but cost $O(n^{3.5})$; heuristics and learned solvers scale further, often without guarantees (Ma & Hao, 2017; Gui et al., 2019; Schuetz et al., 2022; Tönshoff et al., 2023; Qiu et al., 2025).

This paper asks what low-rankness in the objective matrix buys. If $\mathbf{Q} = \mathbf{V}\mathbf{V}^\dagger$ with $\mathbf{V} \in \mathbb{C}^{n\times r}$, then $\mathbf{z}^\dagger\mathbf{Q}\mathbf{z} = \|\mathbf{V}^\dagger\mathbf{z}\|_2^2$, and Cauchy–Schwarz turns the search over K^n discrete vectors into a search over directions $\mathbf{c}$ on the unit sphere of $\mathbb{C}^r$. Rounding each coordinate of $\mathbf{Vc}$ to the nearest root of unity turns a direction into a candidate solution, and some direction yields an optimum (Section 2). Exactly low-rank objectives arise in limited-feedback MIMO beamforming (Zheng et al., 2007; Leung et al., 2010) and noncoherent sequence detection (Papailiopoulos & Karystinos, 2010), and more generally whenever a few eigenvectors carry the structure of $\mathbf{Q}$. Our guarantees for approximately low-rank matrices pay for such a spectral tail explicitly (Section 4).

There are two ways to search the sphere. *Enumeration* splits it into the cells of an arrangement of $O(nK)$ hyperplanes and visits every cell. It is exact for every K, with n candidates at $r = 1$ found

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

by one sort (Mackenthun, 1994; Sweldens, 2001) and $O(rn^{2r-1})$ candidates in general (Kyrillidis & Karystinos, 2014), which is already cubic in n at $r = 2$. *Sampling* draws directions uniformly at random, rounds them, and keeps the best, at $O(nr)$ time per draw. Sampling is what one runs at scale, so the question is how many draws it needs. For low-rank sparse PCA and densest subgraph, random directions in the rank-d eigenspace give a $(1-\varepsilon)$-approximation of the rank-d optimum with probability $1-\delta$ using $O(\varepsilon^{-d}\log(1/(\varepsilon\delta)))$ draws, independent of n (Asteris et al., 2014; Papailiopoulos et al., 2014). On the sphere of $\mathbb{C}^r \cong \mathbb{R}^{2r}$, the same argument implies roughly ε^{-2r} draws are needed.

For low-rank quadratic form maximization, we compare each draw with the discrete optimum and its K rotations. The loss is then quadratic in the angle to the nearest rotation, the same effect that gives a $(1/\varepsilon)^{(d-1)/2}$ dependence for norm maximization over zonotopes in $\mathbb{R}^d$ (Cao et al., 2026), and for roots of unity the angle has an explicit law. We prove that a uniformly random direction is $(1-\varepsilon)$-optimal with probability at least $(\varepsilon/2)^{r-1}\min\{1, \frac{K}{\pi}\sqrt{\varepsilon/2}\}$ for every n and every $\mathbf{V}$ (Theorem 1), so about $\varepsilon^{-(r-1/2)}$ draws suffice. On a family of instances with one heavy row and many isotropic rows, the exponent $r-1/2$ is attained numerically (Appendix B.4).

Contributions. We study randomized algorithms for quadratic form maximization over roots of unity in their simplest form: round random directions and keep the best. We reach the guarantees in three steps: compare any draw with the optimum (Lemma 3.1), find the exact law of that comparison for a random direction (Lemma 3.2), and transfer the result to approximately low-rank objectives (Proposition 4.1). We give experiments testing each result on MAX-3-CUT.

- *Sampling guarantee* (Theorem 1). For every rank-r objective, $S = O\big((2/\varepsilon)^{r-1/2}\log(1/\delta)\big)$ random directions give a $(1-\varepsilon)$-optimal solution with probability $1-\delta$ in $O(Snr)$ time, and one draw attains $\kappa_K \mathrm{OPT}_r/r$ in expectation, with $\kappa_K = \frac{1}{2} + \frac{K}{4\pi}\sin\frac{2\pi}{K}$ ($\kappa_3 \approx 0.71$). The proof needs no incoherence, eigengap, or general-position assumption, and the exponent $r-1/2$ is attained numerically on explicit instances. The proof bounds every draw from below with Cauchy–Schwarz and coordinatewise rounding, losing only quadratically in the angle to the nearest optimal rotation, and uses the exact $\mathrm{Beta}(1, r-1)$ law of the overlap and the uniform law of the phase.
- *Noisy objectives* (Proposition 4.1). For $\mathbf{Q} = \mathbf{Q}^\star + \mathbf{H}$ with arbitrary $\mathbf{H}$, sampling on a rank-r truncation of the Hermitian part of $\mathbf{Q}$ returns $\hat{\mathbf{z}}$ with $\hat{\mathbf{z}}^\dagger \mathbf{Q}^\star \hat{\mathbf{z}} \geq (1-\varepsilon)\mathrm{OPT}(\mathbf{Q}^\star) - 2n(\lambda^\star_{r+1} + 2\|\mathbf{H}\|_2)$. The proof compares objective values, not eigenvectors: Weyl's and the triangle inequality give $\|\mathbf{Q}^\star - \mathbf{Q}_r\|_2 \leq \lambda^\star_{r+1} + 2\|\mathbf{H}\|_2$, and $\|\mathbf{z}\|_2^2 = n$ on $\mathcal{A}_K^n$.
- *Experiments* (Section 5). On MAX-3-CUT, 10 draws come within 0.1% of the best cut found (median seed) on random regular graphs up to $n = 10^6$ and tori up to $n = 10^5$. On Erdős–Rényi graphs the draws needed to come within 1% grow from 10 to 10^2–10^3 at the largest sizes. On graphs small enough to solve exactly, 100 draws reach 99.8% of the rank-2 optimum on average, far above the guarantee. On planted three-community graphs, Proposition 4.1 becomes informative as n grows.

2 FROM DISCRETE VECTORS TO DIRECTIONS

Notation. Vectors are bold lowercase and matrices bold uppercase; $(\cdot)^\dagger$ is the conjugate transpose, Re and Im are real and imaginary parts, $\|\cdot\|_2$ is the vector Euclidean norm and the spectral norm of a matrix, and $[n] = \{1, \ldots, n\}$. We write $\mathbb{S}^r_{\mathbb{C}} = \{\mathbf{c} \in \mathbb{C}^r : \|\mathbf{c}\|_2 = 1\}$ and $\omega = e^{2\pi i/K}$. Throughout, $\mathbf{Q}_r = \mathbf{V}\mathbf{V}^\dagger$ with $\mathbf{V} \in \mathbb{C}^{n\times r}$, $f(\mathbf{z}) = \mathbf{z}^\dagger \mathbf{Q}_r \mathbf{z} = \|\mathbf{V}^\dagger \mathbf{z}\|_2^2$, and $\mathrm{OPT}_r = \max_{\mathbf{z}\in\mathcal{A}_K^n} f(\mathbf{z})$.

Directions and rounding. By Cauchy–Schwarz, $\|\mathbf{V}^\dagger \mathbf{z}\|_2 = \max_{\mathbf{c}\in\mathbb{S}^r_{\mathbb{C}}} \mathrm{Re}(\mathbf{z}^\dagger \mathbf{V}\mathbf{c})$. Exchanging the two maximizations gives:

$$\sqrt{\mathrm{OPT}_r} = \max_{\mathbf{c}\in\mathbb{S}^r_{\mathbb{C}}} \sum_{i=1}^{n} \max_{z_i \in \mathcal{A}_K} \mathrm{Re}\big(\bar{z}_i (\mathbf{V}\mathbf{c})_i\big). \tag{2}$$

For a fixed direction $\mathbf{c}$, the inner maximization separates across coordinates and is solved by rounding each coordinate to the nearest root of unity:

$$z_i(\mathbf{c}) = e^{2\pi i k_i/K}, \qquad k_i \in \operatorname*{argmin}_{k\in\{0,\ldots,K-1\}} \left| \frac{2\pi k}{K} - \arg(\mathbf{V}\mathbf{c})_i \right|_{\bmod 2\pi}. \tag{3}$$

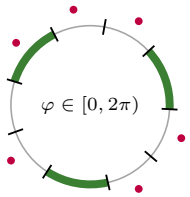


Figure 1: Rank one, $K = 3$, $n = 3$: the direction is $\mathbf{c} = e^{\hat{\imath}\varphi}$ and the sphere is a circle. Each coordinate contributes K boundary points (ticks); the rounding is constant between ticks. The optimal cell and its two rotations by $2\pi/3$ (green) all attain OPT_1 (computed for $q = (e^{0.3\hat{\imath}}, 0.8e^{1.1\hat{\imath}}, 1.2e^{1.75\hat{\imath}})$). Enumeration evaluates one direction per cell; sampling draws random directions (dots) and succeeds when one lands in a green arc.

We write $f(\mathbf{c}) := f(\mathbf{z}(\mathbf{c}))$. Every direction yields a feasible candidate, and if $\mathbf{z}^\star$ is optimal then the direction $\mathbf{c}^\star = \mathbf{V}^\dagger \mathbf{z}^\star / \left\|\mathbf{V}^\dagger \mathbf{z}^\star\right\|_2$ yields an optimal candidate (Lemma 3.1 at $\mathbf{c} = \mathbf{c}^\star$). Rotating a direction by ω relabels every coordinate, $\mathbf{z}(\omega\mathbf{c}) = \omega\mathbf{z}(\mathbf{c})$, so $f(\omega\mathbf{c}) = f(\mathbf{c})$.

Cells. The rounding of coordinate i changes only when $\arg(\mathbf{V}\mathbf{c})_i$ crosses one of the K boundary angles $\vartheta_m = \pi(2m+1)/K$. Identifying $\mathbb{C}^r$ with $\mathbb{R}^{2r}$, each condition $\mathrm{Im}\big(e^{-\hat{\imath}\vartheta_m}(\mathbf{V}\mathbf{c})_i\big) = 0$ is a hyperplane through the origin; boundaries ϑ_m and $\vartheta_m + \pi$ give the same hyperplane, so coordinate i contributes $B_K = K/2$ hyperplanes for even K and $B_K = K$ for odd K. The nB_K hyperplanes cut the sphere into *cells*, and $\mathbf{z}(\mathbf{c})$ is constant on each cell (Figure 1). Enumeration visits every cell; sampling throws random directions at the sphere (Algorithm 1).

Algorithm 1 Low-rank rounding by enumeration or sampling

Input: factor $\mathbf{V} \in \mathbb{C}^{n\times r}$ of $\mathbf{Q}_r = \mathbf{V}\mathbf{V}^\dagger$; alphabet size K; mode ENUMERATE or SAMPLE; budget S
if mode = ENUMERATE **then**
  $\mathcal{C} \leftarrow$ one direction per cell: arrangement vertices, adjacent cells, and lower-rank faces (Appendix A)
else
  $\mathcal{C} \leftarrow \{\mathbf{g}_s / \|\mathbf{g}_s\|_2\}_{s=1}^{S}$, with $\mathbf{g}_s \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_r)$ i.i.d.
end if
for $\mathbf{c} \in \mathcal{C}$ **do**
  $\mathbf{y} \leftarrow \mathbf{V}\mathbf{c}$; $O(nr)$
  $z_i \leftarrow$ root of unity nearest to $\arg y_i$, for $i \in [n]$
  score $f(\mathbf{z}) = \left\|\mathbf{V}^\dagger \mathbf{z}\right\|_2^2$ $O(nr)$
end for
return the best $\mathbf{z}$ found

Proposition 2.1 (Enumeration; Kyrillidis & Karystinos, 2014; Mackenthun, 1994)**.** *Let $K \geq 2$ and let the rows of $\mathbf{V}$ be in general position (Assumption 1). Over all directions, the rounding (3) takes $O(rn^{2r-1})$ distinct values (explicit bound in Theorem 3), and* ENUMERATE *lists all of them in $O(rn^{2r+1})$ time; the candidates are independent, so P processors divide this time by P. For $r = 1$ there are n candidates up to rotation, listed by sorting the n boundary phases and flipping one coordinate per boundary, in $O(n \log n)$ time with incremental scoring.*

The rank-one sweep is due to Mackenthun (1994) and was rediscovered by Sweldens (2001) and, for discrete-phase beamforming, by Ren et al. (2023). The rank-r construction is that of Kyrillidis & Karystinos (2014), which covers even K. Appendix A restates the construction for every K; for odd K the count formula there is an upper bound, not an exact count. We use enumeration to certify sampling on small instances; for $r \geq 2$, its $\Theta(n^{2r-1})$ candidates rule it out at scale.

3 How many random directions?

SAMPLE draws $\mathbf{c}$ uniformly from $\mathbb{S}_{\mathbb{C}}^r$ (a normalized standard complex Gaussian). Its analysis rests on one deterministic inequality and one distributional fact. For $\mathbf{V} \neq \mathbf{0}$ we have $\mathrm{OPT}_r > 0$: a $\mathbf{z}$ with independent uniform phases has $\mathbf{E}\big[\mathbf{z}^\dagger \mathbf{Q}_r \mathbf{z}\big] = \mathrm{tr}\,\mathbf{Q}_r = \|\mathbf{V}\|_F^2 > 0$, and some labeling attains the mean. Fix an optimum $\mathbf{z}^\star$, let $\mathbf{u} = \mathbf{V}^\dagger \mathbf{z}^\star / \left\|\mathbf{V}^\dagger \mathbf{z}^\star\right\|_2$ (so $\mathbf{u} = \mathbf{c}^\star$), and let $\psi(\mathbf{c}) \in [0, \pi/K]$ be the distance from $\arg(\mathbf{u}^\dagger \mathbf{c})$ to the nearest multiple of $2\pi/K$ (with $\psi(\mathbf{c}) = 0$ if $\mathbf{u}^\dagger \mathbf{c} = 0$).

Lemma 3.1 (Direct comparison)**.** *For every $\mathbf{V} \in \mathbb{C}^{n\times r}$ with $\mathbf{V} \neq \mathbf{0}$ and every $\mathbf{c} \in \mathbb{S}_{\mathbb{C}}^r$,*

$$f(\mathbf{c}) \geq \mathrm{OPT}_r \cdot |\mathbf{u}^\dagger \mathbf{c}|^2 \cdot \cos^2 \psi(\mathbf{c}). \tag{4}$$

Proof. Fix $k \in \{0, \ldots, K-1\}$. The vector $\omega^k \mathbf{z}^\star$ is feasible, and rounding solves the inner problem in (2), so Cauchy–Schwarz ($\|\mathbf{c}\|_2 = 1$) and optimality of rounding give

$$\begin{aligned}\sqrt{f(\mathbf{c})} = \left\|\mathbf{V}^\dagger \mathbf{z}(\mathbf{c})\right\|_2 &\geq \mathrm{Re}\big(\mathbf{z}(\mathbf{c})^\dagger \mathbf{V}\mathbf{c}\big) \geq \mathrm{Re}\big((\omega^k \mathbf{z}^\star)^\dagger \mathbf{V}\mathbf{c}\big) \\ &= \sqrt{\mathrm{OPT}_r}\,|\mathbf{u}^\dagger \mathbf{c}| \cos\big(\arg(\mathbf{u}^\dagger \mathbf{c}) - 2\pi k/K\big).\end{aligned}$$

Choosing k that minimizes the angle makes the cosine equal to $\cos\psi(\mathbf{c}) \geq \cos(\pi/K) \geq 0$. Both sides are nonnegative, so squaring proves (4). □

What Lemma 3.1 gives. A draw is good when it points close to the optimal direction or to one of its K rotations: its rounding keeps the fraction $|\mathbf{u}^\dagger\mathbf{c}|^2\cos^2\psi(\mathbf{c})$ of OPT_r, and both factors are squared cosines, so the loss grows only quadratically with the miss. The bound is deterministic and uses nothing about $\mathbf{V}$ or n; the rest of the analysis needs only the laws of two angles.

Lemma 3.2 (Modulus and phase of a sphere marginal)**.** *If $\mathbf{c}$ is uniform on $\mathbb{S}^r_{\mathbb{C}}$ and $\mathbf{u} \in \mathbb{S}^r_{\mathbb{C}}$ is fixed, then $X = |\mathbf{u}^\dagger\mathbf{c}|^2$ follows $\mathrm{Beta}(1, r-1)$ ($X = 1$ when $r = 1$), the phase $\arg(\mathbf{u}^\dagger\mathbf{c})$ is uniform on $[0, 2\pi)$ and independent of X, and $\psi(\mathbf{c})$ is uniform on $[0, \pi/K]$.*

What Lemma 3.2 gives. The exact law of both angles for a random direction: the overlap has the tail $\mathbf{Pr}\big[X \geq 1-t\big] = t^{r-1}$, where the rank enters, and the phase is uniform and independent of it, where the alphabet enters. Neither depends on n or $\mathbf{V}$; with Lemma 3.1, every part of Theorem 1 is a short computation (proof in Appendix B).

Theorem 1 (Sampling guarantee)**.** *Let $K \geq 2$, $r \geq 1$, $\mathbf{V} \in \mathbb{C}^{n\times r}$ nonzero, $\mathbf{c}$ uniform on $\mathbb{S}^r_{\mathbb{C}}$, and $\kappa_K = \frac{1}{2} + \frac{K}{4\pi}\sin\frac{2\pi}{K}$. Then:*

(i) $\mathbf{E}\big[f(\mathbf{c})\big] \geq \dfrac{\kappa_K}{r}\mathrm{OPT}_r$*;*

(ii) $\mathbf{Pr}\big[f(\mathbf{c}) \geq \cos^2(\pi/K)(1-\varepsilon)\mathrm{OPT}_r\big] \geq \varepsilon^{r-1}$ *for every $\varepsilon \in (0,1)$;*

(iii) $\mathbf{Pr}\big[f(\mathbf{c}) \geq (1-\varepsilon)\mathrm{OPT}_r\big] \geq p_{K,r}(\varepsilon)$ *for every $\varepsilon \in (0,1)$, where:*

$$p_{K,r}(\varepsilon) = \left(\tfrac{\varepsilon}{2}\right)^{r-1}\min\left\{1, \tfrac{K}{\pi}\arcsin\sqrt{\varepsilon/2}\right\}$$

for $r \geq 2$ and $p_{K,1}(\varepsilon) = \min\left\{1, \frac{K}{\pi}\arcsin\sqrt{\varepsilon}\right\}$;

(iv) with $S \geq \log(1/\delta)/p_{K,r}(\varepsilon)$ independent directions, the best rounding is $(1-\varepsilon)$-optimal with probability at least $1-\delta$, in $O(Snr)$ time.

Proof sketch. The tools are Cauchy–Schwarz and the exactness of coordinatewise rounding (Lemma 3.1), the K-fold rotational symmetry of optima, the Beta and uniform laws of a sphere marginal and their independence (Lemma 3.2), and $(1-p)^S \leq e^{-pS}$ across independent draws. By Lemmas 3.1 and 3.2, $f(\mathbf{c}) \geq \mathrm{OPT}_r X\cos^2\psi$ with $X \sim \mathrm{Beta}(1, r-1)$ and $\psi \sim \mathrm{Unif}[0, \pi/K]$ independent. Part (i) takes expectations: $\mathbf{E}\big[X\big] = 1/r$ and $\mathbf{E}\big[\cos^2\psi\big] = \kappa_K$. Part (ii) uses $\cos^2\psi \geq \cos^2(\pi/K)$ and $\mathbf{Pr}\big[X \geq 1-\varepsilon\big] = \varepsilon^{r-1}$. Part (iii) intersects the independent events $\{X \geq 1-\varepsilon/2\}$ and $\{\cos^2\psi \geq 1-\varepsilon/2\}$, on which $X\cos^2\psi \geq (1-\varepsilon/2)^2 \geq 1-\varepsilon$. Part (iv) follows from $(1-p)^S \leq e^{-pS}$; one draw costs $O(nr)$ for $\mathbf{Vc}$, the rounding, and $\big\|\mathbf{V}^\dagger\mathbf{z}\big\|_2^2$. Full proofs are in Appendix B. □

No dependence on n or on the structure of $\mathbf{V}$. The guarantee holds for every factor $\mathbf{V}$. Since $\arcsin x \geq x$, part (iv) asks for at most $\frac{\pi}{K}(2/\varepsilon)^{r-1/2}\log(1/\delta)$ draws whenever $\varepsilon \leq 2\pi^2/K^2$ (every $\varepsilon < 1$ when $K \leq 4$), and for at most $2(2/\varepsilon)^{r-1/2}\log(1/\delta)$ draws for every K. For $K = 3$, $r = 2$, and $\varepsilon = \delta = 0.01$, this is 1.4×10^4 draws for every n, whereas enumeration visits $\Theta(n^3)$ cells. For $r = 1$, 49 draws suffice at the same ε and δ, with no margin condition on the phases.

Where the exponent comes from. The event in (iii) asks a draw to land within an angle of order $\sqrt{\varepsilon}$ of one of the K rotations of $\mathbf{c}^\star$ (the green arcs of Figure 1); on $\mathbb{S}^r_{\mathbb{C}} \cong \mathbb{S}^{2r-1}$ such a cap has measure of order $\varepsilon^{r-1/2}$. The theorem uses nothing else about the instance, which is why it holds for every $\mathbf{V}$ and why it is conservative: experimentally, 100 draws reach $0.998\mathrm{OPT}_r$ on average (Section 5).

Relation to prior analyses. Comparing a draw with an optimal direction is the argument behind random-sampling and ε-net guarantees for low-rank sparse PCA, densest subgraph, bipartite correlation clustering, and sparse CCA (Asteris et al., 2014; Papailiopoulos et al., 2014; Asteris et al., 2015; 2016b;a); there it is applied through a Lipschitz bound, which loses linearly in the distance to the optimal direction and leads to about ε^{-2r} draws. Two features sharpen it here: the K rotations of $\mathbf{z}^\star$ are all optimal, which multiplies the success probability by K, and (4) loses only quadratically in the angle. The quadratic loss is the same effect behind the deterministic $(1/\varepsilon)^{(d-1)/2}$ algorithm

of Cao et al. (2026) for norm maximization over zonotopes in $\mathbb{R}^d$, which covers $K = 2$ and $K = 4$; our contribution is the explicit law of the angle for every K and the resulting constants for random directions. Comparing a draw with the relaxation $\max_{\mathbf{c}} \|\mathbf{V}\mathbf{c}\|_1^2 \geq \mathrm{OPT}_r$ instead loses the rounding factor and yields only part (ii), whose threshold stays at $\cos^2(\pi/K)\mathrm{OPT}_r$ however many draws are taken. Part (i) raises the constant in $\mathbf{E}[f]/\mathrm{OPT}_r$ from $\cos^2(\pi/K)/r$ to κ_K/r (from 0.25 to 0.71 in the numerator for $K = 3$); for $r \geq 2$ it remains below the semidefinite ratio of So et al. (2007) (0.35 vs. 0.54 at $K = 3$, $r = 2$), which it does not aim to match. For large r, a second-moment argument gives a sample size polynomial in r at the weaker threshold $\kappa_K \mathrm{OPT}_r/r$ under an incoherence assumption (Appendix D).

Tightness and the real binary case. The exponent $r - 1/2$ corresponds to guarantees that hold for every $\mathbf{V}$. Consider one heavy row plus many isotropic rows of the same total weight. In the limit of many rows, $f(\mathbf{c}) = \frac{\mathrm{OPT}_r}{2}\big(1 + |\mathbf{u}^\dagger \mathbf{c}| \cos\psi(\mathbf{c})\big)$, whose success probability is of order $\varepsilon^{r-1/2}$; on finite instances of this family with numerically verified optima, the measured success probability decays as $\varepsilon^{0.51}$, $\varepsilon^{1.48}$ to $\varepsilon^{1.50}$, and $\varepsilon^{2.46}$ for $r = 1, 2, 3$ (Appendix B.4). For a fixed instance the probability tends to a positive constant as $\varepsilon \to 0$, since $f = \mathrm{OPT}_r$ on a neighborhood of $\mathbf{c}^\star$. For real $\mathbf{V}$ and $K = 2$, the rounding depends only on $\mathrm{Re}\,\mathbf{c}$, and the same comparison gives a success probability of order $\varepsilon^{(r-1)/2}$ (Appendix B.4).

4 Approximately Low-Rank Objectives

Objectives are rarely exactly low rank. We model the input as $\mathbf{Q} = \mathbf{Q}^\star + \mathbf{H}$, where $\mathbf{Q}^\star \succeq 0$ is nonzero, with eigenvalues $\lambda_1^\star \geq \lambda_2^\star \geq \cdots \geq \lambda_n^\star \geq 0$ (and $\lambda_j^\star := 0$ for $j > n$) and top eigenvector $\mathbf{u}_1^\star$, and $\mathbf{H} \in \mathbb{C}^{n\times n}$ is arbitrary: it need not be Hermitian, PSD, or random.

We measure a solution by the noiseless objective $\mathbf{z}^\dagger \mathbf{Q}^\star \mathbf{z}$. Only the Hermitian part matters, since $\mathrm{Re}(\mathbf{z}^\dagger \mathbf{Q}\mathbf{z}) = \mathbf{z}^\dagger \mathbf{Q}_h \mathbf{z}$ with $\mathbf{Q}_h = (\mathbf{Q} + \mathbf{Q}^\dagger)/2$, so we run Algorithm 1 on $\mathbf{V} = [\sqrt{\lambda_1^+}\mathbf{w}_1, \ldots, \sqrt{\lambda_r^+}\mathbf{w}_r]$, where $(\lambda_j, \mathbf{w}_j)$ are the top r eigenpairs of $\mathbf{Q}_h$ and $\lambda_j^+ = \max\{\lambda_j, 0\}$; then $\mathbf{Q}_r = \mathbf{V}\mathbf{V}^\dagger$.

The guarantee below also holds if the best draw is selected by the full objective $\mathrm{Re}(\mathbf{z}^\dagger \mathbf{Q}\mathbf{z})$ rather than by $\mathbf{z}^\dagger \mathbf{Q}_r \mathbf{z}$ (Appendix C).

Proposition 4.1 (Noisy objectives, no eigengap)**.** *Let* $\mathbf{H}_h = (\mathbf{H}+\mathbf{H}^\dagger)/2$ *and* $\Delta = \lambda_{r+1}^\star + 2\|\mathbf{H}_h\|_2$, *so* $\Delta \leq \lambda_{r+1}^\star + 2\|\mathbf{H}\|_2$. *Then* $\|\mathbf{Q}^\star - \mathbf{Q}_r\|_2 \leq \Delta$, *and every* $\hat{\mathbf{z}}$ *with* $\hat{\mathbf{z}}^\dagger \mathbf{Q}_r \hat{\mathbf{z}} \geq (1-\varepsilon)\mathrm{OPT}(\mathbf{Q}_r)$ *satisfies:*

$$\hat{\mathbf{z}}^\dagger \mathbf{Q}^\star \hat{\mathbf{z}} \geq (1-\varepsilon)\mathrm{OPT}(\mathbf{Q}^\star) - 2n\Delta. \tag{5}$$

If also $\|\mathbf{u}_1^\star\|_\infty \leq \mu/\sqrt{n}$, *then, with* $s_K = \frac{K}{\pi}\sin\frac{\pi}{K}$,

$$\frac{\hat{\mathbf{z}}^\dagger \mathbf{Q}^\star \hat{\mathbf{z}}}{\mathrm{OPT}(\mathbf{Q}^\star)} \geq 1 - \varepsilon - \frac{2\mu^2}{s_K^2} \cdot \frac{\lambda_{r+1}^\star + 2\|\mathbf{H}_h\|_2}{\lambda_1^\star}. \tag{6}$$

Proof sketch. Weyl's inequality bounds the discarded eigenvalues of $\mathbf{Q}_h$ by $\lambda_{r+1}^\star + \|\mathbf{H}\|_2$ from above and by $-\|\mathbf{H}\|_2$ from below, and $\|\mathbf{Q}_h - \mathbf{Q}^\star\|_2 \leq \|\mathbf{H}\|_2$. Since $\|\mathbf{z}\|_2^2 = n$ on $\mathcal{A}_K^n$, every objective value moves by at most $n\Delta$ between $\mathbf{Q}^\star$ and $\mathbf{Q}_r$; applying this at $\hat{\mathbf{z}}$ and at an optimum of $\mathbf{Q}^\star$ gives (5). For (6), rounding $e^{i\hat{\varphi}}\mathbf{u}_1^\star$ with the best global phase φ certifies $\mathrm{OPT}(\mathbf{Q}^\star) \geq s_K^2 \lambda_1^\star \|\mathbf{u}_1^\star\|_1^2 \geq s_K^2 \lambda_1^\star n/\mu^2$. Full proofs are in Appendix C. □

By Theorem 1(iv), sampling on $\mathbf{Q}_r$ supplies such a $\hat{\mathbf{z}}$ with probability $1 - \delta$ after $S = O\big((2/\varepsilon)^{r-1/2}\log(1/\delta)\big)$ draws.

No eigengap. Bounds that go through eigenvector perturbation need a gap between $\lambda_r^\star$ and $\lambda_{r+1}^\star$ and scale as $\lambda_1^\star \|\mathbf{H}\|_2$ /gap (Wedin, 1972; Chen et al., 2021). Proposition 4.1 compares objective values directly, so no gap appears, and a non-Hermitian $\mathbf{H}$ costs no more than its Hermitian part $\mathbf{H}_h$; a skew-Hermitian $\mathbf{H}$ costs nothing. The same comparison shows when the optimizer itself survives truncation: if every non-optimal value of $\mathbf{Q}^\star$ is below $\mathrm{OPT}(\mathbf{Q}^\star)$ by more than $2n\Delta$, the optimum of $\mathbf{Q}_r$ is optimal for $\mathbf{Q}^\star$.

When the bound is informative. Inequality (5) matters when $n\Delta$ is small compared with $\mathrm{OPT}(\mathbf{Q}^\star)$, which (6) makes explicit. The incoherence condition there only certifies that $\mathrm{OPT}(\mathbf{Q}^\star)$ grows like $n\lambda_1^\star$. It holds with $\mu = 1$ for a K-ary synchronization spike $\mathbf{Q}^\star = \frac{\lambda}{n}\mathbf{x}\mathbf{x}^\dagger$ with $\mathbf{x} \in \mathcal{A}_K^n$. Graph Laplacians of expanders sit at the other extreme: their spectra are flat, $\lambda_{r+1}^\star$ is close to $\lambda_1^\star$, and Proposition 4.1 says nothing (Section 5).

5 Experiments

We test four predictions, all at $K = 3$: (P1) the number of draws needed for a near-best cut does not grow with n (Theorem 1); (P2) best-of-S approaches the exact rank-r optimum at least as fast as Theorem 1 allows; (P3) Proposition 4.1 becomes informative when the objective is low rank plus noise; (P4) low-rank rounding is competitive where a few eigenvectors carry the cut structure.

Our guarantees do not cover (P4): Theorem 1 bounds the rank-r objective rather than the cut, and Proposition 4.1 is vacuous for graph Laplacians, whose spectra are flat. We test it anyway, as a stress test on standard Max-3-Cut benchmarks. We use random 5-regular graphs, Erdős–Rényi graphs, two-dimensional tori, planted three-community graphs, Delaunay triangulations, road networks, and the GSet benchmark (Ye, 2003), and compare with complex semidefinite relaxations with randomized rounding (Goemans & Williamson, 2004), k-means on the top eigenvectors (Lei & Rinaldo, 2015), the greedy and genetic algorithms of Gui et al. (2019) and Li & Wang (2016), local search, simulated annealing (Johnson et al., 1991), and published MOH values (Ma & Hao, 2017). Appendix E gives the setup, hardware, and all results.

5.1 Sampling at scale

We run Sample with $r = 2$ on the random 5-regular, Erdős–Rényi, and toroidal families, sweeping n from 10^3 to 10^6 (10^5 for tori) and S up to 10^6 (10^5 for $n \geq 5 \times 10^5$). The exact optimum OPT_r is out of reach at these sizes, where enumeration costs $\mathcal{O}(n^{2r-1})$, so we divide the best cut after S draws by the best cut after 10^5 draws, the largest budget run at every size, and report the median over 3 to 5 seeds.

The budget does not grow with n on two of three families (P1). On 5-regular graphs (n up to 10^6) and toroidal graphs (n up to 10^5), the median seed comes within 0.1% of the reference after 10 draws at every size (Figure 2). On Erdős–Rényi graphs, 10 draws come within 1% on every seed up to $n = 10^5$, but the median seed needs 10^3 draws at $n = 5 \times 10^5$ and 10^2 at $n = 10^6$; the median budget to come within 0.1% ranges from 10 to 3×10^4 with no upward trend in n. Theorem 1 bounds the rank-r objective f, not the cut, so these measurements are consistent with its n-free success probability rather than implied by it.

One draw is already good, and 100 draws nearly reach OPT_r (P2). On 27 instances with $n \leq 100$ (5-regular and Erdős–Rényi graphs at three sizes with three seeds each, and one torus per size with three sampling seeds: 21 distinct graphs) we compute OPT_r exactly by visiting every cell of the rank-2 arrangement (Appendix E.2) and record f, the quantity that Theorem 1 bounds, for 2×10^5 draws per graph. A single draw attains 0.79 to 0.93 of OPT_r in expectation, 2.2 to 2.6 times the guarantee $\kappa_3/2 \approx 0.35$ of part (i), and on every instance $(1-\varepsilon)$-optimal draws are at least 12 ($\varepsilon = 0.3$) to 52 ($\varepsilon = 0.01$) times more frequent than the bound $p_{3,2}(\varepsilon)$ of part (iii) (Figure 3a). The best of 10, 100, and 1000 draws averages 0.977, 0.998, and 0.9999 of OPT_r, and is at least 0.967, 0.994, and 0.999 on every instance, while part (iv) asks for 1.4×10^4 draws to guarantee $0.99\mathrm{OPT}_r$ with probability 0.99: the guarantee holds with a wide margin.

A low-rank-plus-noise objective (P3). Laplacians of our graph families have flat spectra, so Proposition 4.1 says nothing about them. We therefore plant three communities of equal size, with edges inside a community with probability p_{in} and across with $p_{\mathrm{out}} < p_{\mathrm{in}}$, and maximize $\mathrm{Re}(\mathbf{z}^\dagger \mathbf{B}\mathbf{z})$ for the centered adjacency matrix $\mathbf{B} = \mathbf{A} - \hat{p}\mathbf{1}\mathbf{1}^\top$, where $\hat{p}$ is the edge density. Up to its diagonal, $\mathbf{E}[\mathbf{B}]$ is the rank-2 matrix $\mathbf{Q}^\star = \frac{p_{\mathrm{in}}-p_{\mathrm{out}}}{3}(\mathbf{x}\mathbf{x}^\dagger + \bar{\mathbf{x}}\mathbf{x}^\top)$ with $x_i = e^{2\pi \hat{\imath} \sigma_i/3}$ for the community σ_i of vertex i; for $K = 3$ this objective depends only on the partition. With $p_{\mathrm{in}} = 0.3$ and $p_{\mathrm{out}} = 0.1$, the right-hand side of Proposition 4.1 is negative at $n = 1500$ and rises to 0.20, 0.43, and 0.59 at $n = 3000$, 6000, and 12000, always below the attained 1.00 (Figure 3b). On sparse instances with average degree 30 and n up to 3×10^5, the median number of draws needed to come within 1% of the final objective does not grow with n, sampling agrees with k-means on the same eigenvectors (Lei &

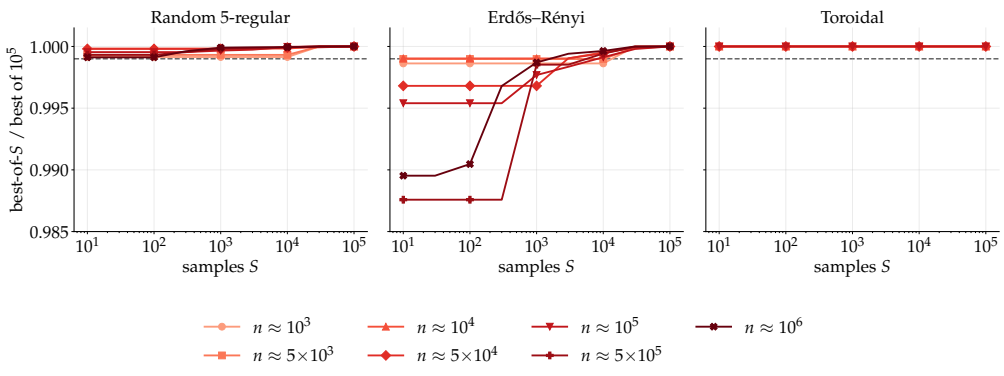


Figure 2: Uniform-sphere sampling on MAX-3-CUT ($K = 3$, $r = 2$); panels, left to right, are random 5-regular, Erdős–Rényi, and toroidal graphs (torus sizes are the nearest squares). Best cut after S draws divided by the best cut after 10^5 draws, median over 3 to 5 seeds, one curve per size n; the dashed line marks 0.999.

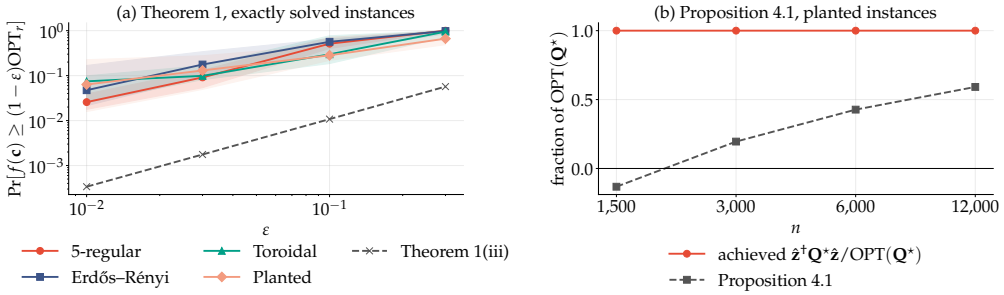


Figure 3: $K = 3$, $r = 2$: (a) Frequency of $(1-\varepsilon)$-optimal draws, $\mathbf{Pr}[f(\mathbf{c}) \geq (1-\varepsilon)\mathrm{OPT}_r]$, over 2×10^5 draws on every instance with exactly computed OPT_r (graphs of P2 and planted instances of P3; median over instances, band from minimum to maximum), and the lower bound of Theorem 1(iii). (b) Planted instances with $p_{\mathrm{in}} = 0.3$ and $p_{\mathrm{out}} = 0.1$: the attained $\hat{\mathbf{z}}^\dagger \mathbf{Q}^\star \hat{\mathbf{z}}/\mathrm{OPT}(\mathbf{Q}^\star)$ after 10^4 draws and the right-hand side of Proposition 4.1 at $\delta = 0.01$, median over five seeds.

Rinaldo, 2015) to within 0.02 in overlap with the planted partition, and on exactly solved instances $(1-\varepsilon)$-optimal draws are at least $8\times$ more frequent than Theorem 1(iii) guarantees (Appendix E.3).

5.2 LOW-RANK ROUNDING AS A MAX-3-CUT SOLVER

We compare low-rank rounding with standard heuristics on random graphs with $n \leq 100$ (20 seeds per setting), on the 71 GSet graphs (Ye, 2003) with $800 \leq n \leq 20{,}000$, and on graphs with up to 1.4×10^6 vertices (Appendix E.1). The baselines are the complex semidefinite relaxation of Goemans & Williamson (2004) with its randomized rounding, the greedy algorithm of Gui et al. (2019) and the genetic algorithm of Li & Wang (2016) as implemented by Qiu et al. (2025), 1-opt local search (LS) from a random start, simulated annealing (Johnson et al., 1991), and the best of 20 runs of MOH (Ma & Hao, 2017). The rank-1 sweep enumerates the cells of the rank-one surrogate $\mathbf{v}_1/\|\mathbf{v}_1\|_2 + \hat{\imath}\mathbf{v}_2/\|\mathbf{v}_2\|_2$ of the top two Laplacian eigenvectors.

On small graphs (Figure 4; means over the 20 seeds), rank-3 sampling with 10^5 draws is within 0.83% of the best cut found on 5-regular graphs, ahead of the SDP at every size and of greedy for $n \leq 50$ (at $n = 100$ it leads greedy by 0.12 points, less than the variation over seeds); on Erdős–Rényi graphs it trails greedy by 0.4 to 3.3 points, and the exact rank-2 optimum trails both greedy and the SDP at every density. The rank-1 sweep followed by LS is within 1.65% of the best cut in every setting, and within 0.71% on Erdős–Rényi graphs. On GSet (Section 5.2; low-rank values are medians over five eigensolver start vectors), the sweep alone cuts all 6000 edges of the unit-weight tori G48 and G49, as greedy does, but averages 0.68 of the MOH value, and 0.51 on the graphs with ± 1 weights (0.28 to 0.67 per graph); followed by LS it reaches 0.92 of MOH on the 59 graphs where greedy finishes within 30 minutes (greedy: 0.94) and 0.91 on all 71. On large graphs

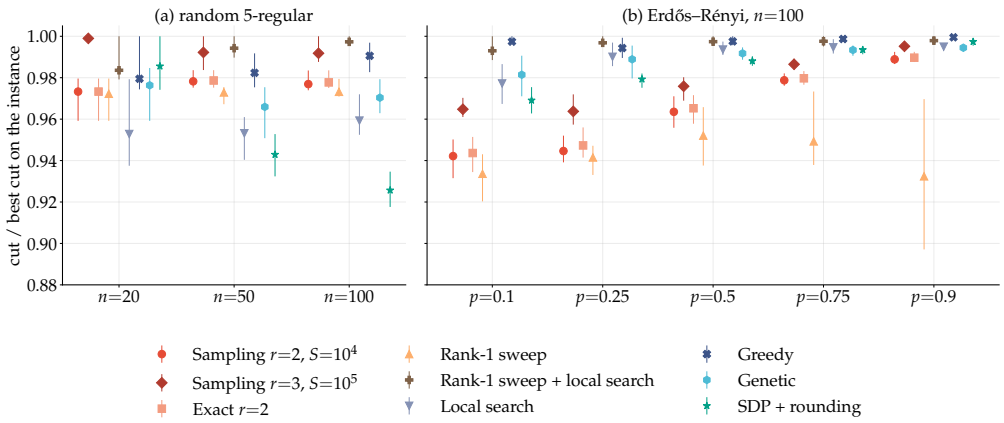


Figure 4: MAX-3-CUT on small graphs: cut divided by the best cut found by any method on the instance, mean over 20 seeds (bars: interquartile range). (a) Random 5-regular graphs; (b) Erdős–Rényi graphs with $n = 100$. Sampling rounds uniform directions on the top r eigenvectors, Exact r=2 is the best cut that rounding a rank-2 direction can give, LS is 1-opt local search, and Greedy and Genetic are the implementations of Qiu et al. (2025).

Graph	n	Rank-1	r=2, S=10^4	r=3, S=10^4	Rank-1+LS	LS	Greedy	Genetic	MOH
G1	800	13713	13955	14072	14858	14824	14829	14075	15165
G14	800	3084	3329	3352	3860	3818	3916	3679	4012
G22	2000	15147	15220	15341	16613	16501	16568	–	17167
G43	1000	7576	7639	7713	8282	8228	8276	7624	8573
G48	3000	6000	6000	6000	6000	5770	6000	–	6000
G49	3000	6000	6000	6000	6000	5786	6000	–	6000
G55	5000	10777	10851	10873	11830	11635	12047	–	12429
G67	10000	3000	3098	2950	6691	6614	–	–	8086
G81	20000	4604	5399	3969	13601	13437	–	–	16321

Table 1: Selected GSet graphs: cut values of the rank-1 sweep, sampling with 10^4 draws ($r = 2, 3$), the sweep followed by local search (Rank-1+LS), local search from a random start (LS), and Greedy and Genetic (implementations of Qiu et al. (2025); – means not finished within 30 minutes or stopped by the method's own time rule; Genetic was run only for $n \leq 1000$), and the best of 20 MOH runs reported by Ma & Hao (2017). Low-rank values are medians over five eigensolver start vectors. All 71 graphs with runtimes are in Appendix E.6.

(Appendix E.5), the sweep followed by LS improves LS from a random start on all 45 runs (35 distinct graphs), by 0.4 to 3.9%, and beats simulated annealing on every torus; simulated annealing is better on the 5-regular and Erdős–Rényi graphs, Delaunay triangulations, and road networks.

What the bounds do and do not explain (P4). The Laplacian spectra of our graph families are flat: on every instance of the large-graph study the top two eigenvalues carry at most 0.44% of the trace (Appendix E.5). Proposition 4.1 is vacuous for them, and it does not explain why low-rank rounding does well on unit-weight tori and on small 5-regular graphs; a plausible reason is that their top eigenvectors align with near-optimal 3-cuts, but we leave such study and extensions for future work. Where this fails, low rank loses: on the unit-weight Erdős–Rényi GSet graphs the sweep reaches 91.5% of greedy, and on the two road networks it is 4.0% below and 0.4% above the best of 1000 random labelings. Theorem 1 is not affected by flat spectra: it concerns the rank-r problem.

6 RELATED WORK

Semidefinite relaxations. Goemans & Williamson (1995) round an SDP relaxation of MAX-CUT to a 0.878 ratio, optimal under the Unique Games Conjecture (Khot et al., 2007); Frieze & Jerrum (1997) and Goemans & Williamson (2004) extend the approach to MAX-K-CUT, and So et al. (2007) to Hermitian PSD objectives over roots of unity, with ratio $(K \sin(\pi/K))^2/(4\pi)$ (see also Zhang & Huang, 2006; Newman, 2018). Burer–Monteiro factorizations reduce the cost of these relaxations by constraining the rank of the SDP variable (Burer & Monteiro, 2003; Park et al., 2017). Our low-rank assumption is on the objective, not on the relaxation.

Low-rank objectives. For fixed rank, quadratic maximization over binary and phase alphabets is polynomial. At rank one, a sort of the phase boundaries solves K-ary phase detection in $O(n \log n)$ time (Mackenthun, 1994; Sweldens, 2001), an algorithm rediscovered for discrete-phase beamforming (Ren et al., 2023). For rank r, arrangement and zonotope enumeration give $n^{O(r)}$ algorithms for binary vectors (Allemand et al., 2001; Ferrez et al., 2005; Karystinos & Liavas, 2010; Markopoulos et al., 2014), for phase alphabets (Motedayen-Aval et al., 2007; Leung et al., 2010; Kyrillidis & Karystinos, 2011; 2014), and for noncoherent detection (Papailiopoulos & Karystinos, 2010). Concurrent work (Soltanalian & Mousavi, 2026) develops a deterministic rank-collapse principle for such problems, including finite-phase vectors, with stability and approximate-rank certificates; it does not analyze random directions. Analyses of MAX-CUT on graphs of low threshold rank (Barak et al., 2011; Arora et al., 2015; Oveis Gharan & Trevisan, 2013) use a spectral premise on the adjacency matrix and aim at approximation schemes.

Low-rank approximation with random directions. The spannogram framework solves the rank-r surrogate of sparse PCA, densest subgraph, correlation clustering, and sparse CCA, either exactly or with ε-nets and random directions that give $(1-\varepsilon)$ guarantees with an n-independent number of directions (Papailiopoulos et al., 2013; Asteris et al., 2014; Papailiopoulos et al., 2014; Asteris et al., 2015; 2016b;a). Theorem 1 applies the same comparison to roots of unity. The quadratic loss near the optimum, which lowers the exponent from about $2r$ to $r - 1/2$, also underlies the concurrent deterministic $(1-\varepsilon)$-approximation of Cao et al. (2026) for norm maximization over zonotopes, alongside their W[1]-hardness result.

Heuristics and learned solvers. Multiple-search-operator, greedy, and genetic heuristics scale to large graphs without guarantees (Ma & Hao, 2017; Gui et al., 2019; Li & Wang, 2016), and graph neural networks and reinforcement learning are benchmarked on MAX-CUT (Schuetz et al., 2022; Barrett et al., 2022; Tönshoff et al., 2023; Qiu et al., 2025).

7 CONCLUSION AND LIMITATIONS

Low rankness turns discrete quadratic optimization over roots of unity into a search over directions, and a uniform random search needs a number of directions that depends on the rank, the alphabet, and the accuracy, but not on n. The guarantee assumes nothing about the factor $\mathbf{V}$, holds for every $K \geq 2$, and has a half-page proof.

Two limitations remain. First, the sample size grows as $\varepsilon^{-(r-1/2)}$, so the guarantee is useful for small ranks; our experiments suggest that this dependence is pessimistic, and an instance-dependent bound on the measure of near-optimal cells is open. Second, many objectives of interest, including Laplacians of expander-like graphs, are far from low rank; there Proposition 4.1 is vacuous, and the empirical success of low-rank rounding on unit-weight tori and small regular graphs still lacks an explanation. Finally, all our experiments use $K = 3$.

AI USE STATEMENT

The exact enumeration algorithms and their proofs were developed by the authors without AI assistance. An AI assistant (Anthropic's Claude) helped with the rest of the analysis, with the experiments, where it designed, implemented, and ran synthetic experiments reported in Section 5 and Appendix E; and with literature search, reference verification, and language editing. The authors checked every AI-assisted text line by line, independent verification runs reproduced every experimental number from the raw result files, and authors take responsibility for the content of this paper.

REPRODUCIBILITY STATEMENT

Complete proofs appear in the appendix: exact enumeration for every K in Appendix A, the sampling guarantee in Appendix B, the bounds for noisy objectives in Appendix C, and a sample size polynomial in the rank in Appendix D. The graph families, baselines, hardware, and evaluation metrics are described in Appendix E, and the complete GSet results are in Appendix E.6. The numerical check of Lemma 3.1 and Theorem 1 against exact optima on small instances is described in Appendix B.

REFERENCES

Kim Allemand, Komei Fukuda, Thomas M. Liebling, and Erich Steiner. A polynomial case of unconstrained zero-one quadratic optimization. *Mathematical Programming*, 91(1):49–52, 2001. doi: 10.1007/s101070100233.

Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential algorithms for Unique Games and related problems. *Journal of the ACM (JACM)*, 62(5):1–25, 2015.

Megasthenis Asteris, Dimitris Papailiopoulos, and Alexandros Dimakis. Nonnegative sparse PCA with provable guarantees. In *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *PMLR*, pp. 1728–1736, 2014.

Megasthenis Asteris, Dimitris Papailiopoulos, Anastasios Kyrillidis, and Alexandros G. Dimakis. Sparse PCA via bipartite matchings. In *Advances in Neural Information Processing Systems*, volume 28, 2015.

Megasthenis Asteris, Anastasios Kyrillidis, Oluwasanmi Koyejo, and Russell Poldrack. A simple and provable algorithm for sparse diagonal CCA. In *Proceedings of the 33rd International Conference on Machine Learning*, volume 48 of *PMLR*, pp. 1148–1157, 2016a.

Megasthenis Asteris, Anastasios Kyrillidis, Dimitris Papailiopoulos, and Alexandros Dimakis. Bipartite correlation clustering: Maximizing agreements. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *PMLR*, pp. 121–129, 2016b.

Boaz Barak, Prasad Raghavendra, and David Steurer. Rounding semidefinite programming hierarchies via global correlation. In *Proceedings of the 2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, FOCS '11, pp. 472–481, USA, 2011. IEEE Computer Society. ISBN 9780769545714. doi: 10.1109/FOCS.2011.95. URL https://doi.org/10.1109/FOCS.2011.95.

Thomas D. Barrett, Christopher W. F. Parsonson, and Alexandre Laterre. Learning to solve combinatorial graph partitioning problems via efficient exploration. *arXiv preprint arXiv:2205.14105*, 2022.

Samuel Burer and Renato D. C. Monteiro. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming*, 95(2):329–357, 2003.

Yang Cao, Haoran Qi, and Hanzhi Wang. ℓ_p-norm maximization over zonotopes is W[1]-hard, 2026.

Yuxin Chen, Chen Cheng, and Jianqing Fan. Asymmetry helps: Eigenvalue and eigenvector analyses of asymmetrically perturbed low-rank matrices. *The Annals of Statistics*, 49(1):435–458, 2021.

Timothy A. Davis and Yifan Hu. The University of Florida sparse matrix collection. *ACM Transactions on Mathematical Software (TOMS)*, 38(1):1–25, 2011.

Jean-Albert Ferrez, Komei Fukuda, and Thomas M. Liebling. Solving the fixed rank convex quadratic maximization in binary variables by a parallel zonotope construction algorithm. *European Journal of Operational Research*, 166(1):35–50, 2005.

Alan Frieze and Mark Jerrum. Improved approximation algorithms for MAX k-CUT and MAX BISECTION. *Algorithmica*, 18(1):67–81, 1997.

Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42 (6):1115–1145, 1995.

Michel X. Goemans and David P. Williamson. Approximation algorithms for MAX-3-CUT and other problems via complex semidefinite programming. *Journal of Computer and System Sciences*, 68(2):442–470, 2004.

Jihong Gui, Zhipeng Jiang, and Suixiang Gao. PCI planning based on binary quadratic programming in LTE/LTE-A networks. *IEEE Access*, 7:203–214, 2019.

Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. Exploring network structure, dynamics, and function using NetworkX. In Gaël Varoquaux, Travis Vaught, and Jarrod Millman (eds.), *Proceedings of the 7th Python in Science Conference*, pp. 11–15, Pasadena, CA USA, 2008.

David S. Johnson, Cecilia R. Aragon, Lyle A. McGeoch, and Catherine Schevon. Optimization by simulated annealing: an experimental evaluation; part II, graph coloring and number partitioning. *Operations Research*, 39(3):378–406, 1991.

George N. Karystinos and Athanasios P. Liavas. Efficient computation of the binary vector that maximizes a rank-deficient quadratic form. *IEEE Transactions on Information Theory*, 56(7): 3581–3593, 2010. doi: 10.1109/TIT.2010.2048450.

Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, 2007.

Anastasios Kyrillidis and George N. Karystinos. Rank-deficient quadratic-form maximization over M-phase alphabet: Polynomial-complexity solvability and algorithmic developments. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3856–3859. IEEE, 2011.

Anastasios Kyrillidis and George N. Karystinos. Fixed-rank Rayleigh quotient maximization by an MPSK sequence. *IEEE Transactions on Communications*, 62(3):961–975, 2014.

Jing Lei and Alessandro Rinaldo. Consistency of spectral clustering in stochastic block models. *The Annals of Statistics*, 43(1):215–237, 2015. doi: 10.1214/14-AOS1274.

Jure Leskovec and Rok Sosič. SNAP: A general-purpose network analysis and graph-mining library. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 8(1):1–20, 2016.

Kin-Kwong Leung, Chi Wan Sung, Majid Khabbazian, and Mohammad Ali Safari. Optimal phase control for equal-gain transmission in MIMO systems with scalar quantization: Complexity and algorithms. *IEEE Transactions on Information Theory*, 56(7):3343–3355, 2010.

Panxing Li and Jing Wang. PCI planning method based on genetic algorithm in LTE network. *Telecommunications Science*, 32(3):147–152, 2016. In Chinese.

Fuda Ma and Jin-Kao Hao. A multiple search operator heuristic for the max-k-cut problem. *Annals of Operations Research*, 248(1–2):365–403, 2017.

K. M. Mackenthun, Jr. A fast algorithm for multiple-symbol differential detection of MPSK. *IEEE Transactions on Communications*, 42(2/3/4):1471–1474, 1994. doi: 10.1109/TCOMM.1994.582823.

Panos P. Markopoulos, George N. Karystinos, and Dimitris A. Pados. Optimal algorithms for L_1-subspace signal processing. *IEEE Transactions on Signal Processing*, 62(19):5046–5058, 2014. doi: 10.1109/TSP.2014.2338077.

Robby G. McKilliam, Andre Pollok, William Cowley, I. Vaughan L. Clarkson, and Barry G. Quinn. Carrier phase and amplitude estimation for phase shift keying using pilots and data. *IEEE Transactions on Signal Processing*, 62(15):3976–3989, 2014. doi: 10.1109/TSP.2014.2332976.

Idin Motedayen-Aval, Arvind Krishnamoorthy, and Achilleas Anastasopoulos. Optimal joint detection/estimation in fading channels with polynomial complexity. *IEEE Transactions on Information Theory*, 53(1):209–223, 2007. doi: 10.1109/TIT.2006.887504.

Alantha Newman. Complex semidefinite programming and Max-k-Cut. In *1st Symposium on Simplicity in Algorithms (SOSA 2018)*, volume 61 of *OpenAccess Series in Informatics (OASIcs)*, pp. 13:1–13:11. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2018.

Shayan Oveis Gharan and Luca Trevisan. A new regularity lemma and faster approximation algorithms for low threshold rank graphs. In *International Workshop on Approximation Algorithms for Combinatorial Optimization*, pp. 303–316. Springer, 2013.

Dimitris Papailiopoulos, Alexandros Dimakis, and Stavros Korokythakis. Sparse PCA through low-rank approximations. In *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *PMLR*, pp. 747–755, 2013.

Dimitris Papailiopoulos, Ioannis Mitliagkas, Alexandros Dimakis, and Constantine Caramanis. Finding dense subgraphs via low-rank bilinear optimization. In *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *PMLR*, pp. 1890–1898, 2014.

Dimitris S. Papailiopoulos and George N. Karystinos. Maximum-likelihood noncoherent OSTBC detection with polynomial complexity. *IEEE Transactions on Wireless Communications*, 9(6): 1935–1945, 2010.

Dohyung Park, Anastasios Kyrillidis, Constantine Caramanis, and Sujay Sanghavi. Non-square matrix sensing without spurious local minima via the Burer-Monteiro approach. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pp. 65–74, 2017.

Yeqing Qiu, Ye Xue, Akang Wang, Yiheng Wang, Qingjiang Shi, and Zhi-Quan Luo. ROS: A GNN-based relax-optimize-and-sample framework for Max-k-Cut problems. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 50721–50743, 2025.

Shuyi Ren, Kaiming Shen, Xin Li, Xin Chen, and Zhi-Quan Luo. A linear time algorithm for the optimal discrete IRS beamforming. *IEEE Wireless Communications Letters*, 12(3):496–500, 2023. doi: 10.1109/LWC.2022.3232146.

Martin J. A. Schuetz, J. Kyle Brubaker, and Helmut G. Katzgraber. Combinatorial optimization with physics-inspired graph neural networks. *Nature Machine Intelligence*, 4(4):367–377, 2022.

Amit Singer. Angular synchronization by eigenvectors and semidefinite programming. *Applied and Computational Harmonic Analysis*, 30(1):20–36, 2011.

Anthony Man-Cho So, Jiawei Zhang, and Yinyu Ye. On approximating complex quadratic optimization problems via semidefinite programming relaxations. *Mathematical Programming*, 110 (1):93–110, 2007.

Mojtaba Soltanalian and Ahmad Mousavi. The rank-collapse principle for quadratic optimization, 2026.

Wim Sweldens. Fast block noncoherent decoding. *IEEE Communications Letters*, 5(4):132–134, 2001. doi: 10.1109/4234.917091.

Jan Tönshoff, Berke Kisin, Jakob Lindner, and Martin Grohe. One model, any CSP: Graph neural networks as fast global search heuristics for constraint satisfaction. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 4280–4288, 2023.

Per-Åke Wedin. Perturbation bounds in connection with singular value decomposition. *BIT*, 12(1): 99–111, 1972.

Yinyu Ye. The Gset dataset, 2003. URL https://web.stanford.edu/~yyye/yyye/Gset/.

Shuzhong Zhang and Yongwei Huang. Complex quadratic optimization and semidefinite programming. *SIAM Journal on Optimization*, 16(3):871–890, 2006.

Xiayu Zheng, Yao Xie, Jian Li, and Petre Stoica. MIMO transmit beamforming under uniform elemental power constraint. *IEEE Transactions on Signal Processing*, 55(11):5395–5406, 2007.