Towards Computing the Grothendieck Constant

Prasad Raghavendra* David Steurer†

Abstract

The *Grothendieck constant* K_G is the smallest constant such that for every $d \in \mathbb{N}$ and every matrix $A = (a_{ij})$,

$$\sup_{\boldsymbol{u}_i,\boldsymbol{v}_j \in B^{(d)}} \sum_{ij} a_{ij} \langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle \leqslant K_G \cdot \sup_{x_i,y_j \in [-1,1]} \sum_{ij} a_{ij} x_i y_j \,,$$

where $B^{(d)}$ is the unit ball in $\mathbb{R}^d$. Despite several efforts [15, 23], the value of the constant K_G remains unknown. The Grothendieck constant K_G is precisely the integrality gap of a natural SDP relaxation for the $K_{M,N}$-Quadratic Programming problem. The input to this problem is a matrix $A = (a_{ij})$ and the objective is to maximize the quadratic form $\sum_{ij} a_{ij} x_i y_j$ over $x_i, y_j \in [-1, 1]$.

In this work, we apply techniques from [22] to the $K_{M,N}$-Quadratic Programming problem. Using some standard but non-trivial modifications, the reduction in [22] yields the following hardness result: Assuming the Unique Games Conjecture [9], it is NP-hard to approximate the $K_{M,N}$-Quadratic Programming problem to any factor better than the Grothendieck constant K_G.

By adapting a "bootstrapping" argument used in a proof of Grothendieck inequality [5], we are able to perform a tighter analysis than [22]. Through this careful analysis, we obtain the following new results:

- An approximation algorithm for $K_{M,N}$-Quadratic Programming that is guaranteed to achieve an approximation ratio arbitrarily close to the Grothendieck constant K_G (optimal approximation ratio assuming the Unique Games Conjecture).
- We show that the Grothendieck constant K_G can be computed within an error η, in time depending only on η. Specifically, for each η, we formulate an explicit finite linear program, whose optimum is η-close to the Grothendieck constant.

We also exhibit a simple family of operators on the Gaussian Hilbert space that is guaranteed to contain tight examples for the Grothendieck inequality.

1 Introduction

The Grothendieck inequality states that for every $m \times n$ matrix $A = (a_{ij})$ and every choice of unit vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, there exists a choice of signs $x_1, \ldots, x_m, y_1, \ldots, y_n \in \{1, -1\}$ such that

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle \leqslant K \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_i y_j \,,$$

where K is a universal constant. The smallest value of K for which the inequality holds, is referred to as the *Grothendieck constant* K_G. Since the inequality was first discovered [8], the inequality has not only undergone various restatements under different frameworks of analysis (see [16]), it has also found numerous applications in functional analysis.

In recent years, the Grothendieck's inequality has found algorithmic applications in efficient construction of Szemerédi partitions of graphs and estimation of cut norms of matrices [2], in turn leading to efficient approximation algorithms for problems in dense and quasi-random graphs [7, 4]. The inequality has also proved useful in certain lower bound techniques for communication complexity [17]. Among its various applications, we shall elaborate here on the $K_{M,N}$-Quadratic Programming problem. In this problem, the objective is to maximize the following quadratic program with the matrix $A = (a_{ij})$ given as input.

$$\text{Maximize } \sum_{i,j} a_{ij} x_i y_j \quad \text{subject to } x_i, y_j \in \{1, -1\}\,.$$

Alternatively, the problem amounts to computing the norm $\|A\|_{\infty\to 1}$ of the matrix A, which is defined as

$$\|A\|_{\infty\to 1} := \max_{\boldsymbol{x} \in \mathbb{R}^n} \frac{\|A\boldsymbol{x}\|_1}{\|\boldsymbol{x}\|_\infty}\,.$$

The $K_{M,N}$-Quadratic Programming problem is a formulation of the correlation clustering problem for two clusters on a bipartite graph [6]. The following natural SDP relaxation to the problem is obtained by relaxing

*Department of Computer Science & Engineering, University of Washington, Seattle, WA 98195. Work done while the author was visiting Princeton University. Supported by NSF grant CCF-0343672.

†Princeton University, Princeton, NJ. Supported by NSF grants MSPA-MCS 0528414, ITR-0205594, and CCF-0832797.

the variables x_i, y_j to unit vectors.

$$\text{Maximize } \sum_{i,j} a_{ij} \langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle$$
$$\text{subject to } \|\boldsymbol{u}_i\| = \|\boldsymbol{v}_j\| = 1\,.$$

The Grothendieck constant K_G is precisely the integrality gap of this SDP relaxation for the $K_{M,N}$-Quadratic Programming problem.

Despite several proofs and reformulations, the value of the Grothendieck constant K_G still remains unknown. In his original work, Grothendieck showed that $\frac{\pi}{2} \leqslant K_G \leqslant 2.3$. The upper bound has been later improved to $\pi/2\log(1+\sqrt{2}) \approx 1.78$ by Krivine [15], while the best known lower bound is roughly 1.67 [23]. More importantly, very little seems to be known about the matrices A for which the inequality is tight [14]. Computing the Grothendieck constant approximatively and characterizing the tight examples for the inequality form the original motivation for this work. Towards this goal, we will harness the emerging connections between semidefinite programming (SDP) and hardness of approximation based on the Unique Games Conjecture (UGC) [9].

In a recent work [22], the first author obtained general results connecting SDP integrality gaps to UGC-based hardness results for arbitrary constraint satisfaction problems (CSP). These connections yielded optimal algorithms and inapproximability for every CSP assuming the Unique Games Conjecture. Further, for the special case of 2-CSPs, it yielded an algorithm to compute the value of the integrality gap of a natural SDP.

Recall that the Grothendieck constant is precisely the integrality gap of the SDP for $K_{M,N}$-Quadratic Programming. In this light, the current work applies the techniques of Raghavendra [22] to the $K_{M,N}$-Quadratic Programming.

1.1 Results

We obtain the following UGC-based hardness result for $K_{M,N}$-Quadratic Programming.

Theorem 1.1. *Assuming the Unique Games Conjecture, it is NP-hard to approximate $K_{M,N}$-Quadratic Programming by any constant factor smaller than the Grothendieck constant K_G.*

Although $K_{M,N}$-Quadratic Programming falls in the "generalized constraint satisfaction problem" framework of Raghavendra [22], the above result does not immediately follow from [22] since the reduction does not preserve bipartiteness. The main technical hurdle in obtaining a bipartiteness-preserving reduction, is to give a stronger analysis of the dictatorship test so as to guarantee a common influential variable. This is achieved using a standard truncation argument as outlined in [19].

Even with the above modification, the optimal algorithm for CSPs in [22] does not directly translate to an algorithm for $K_{M,N}$-Quadratic Programming. The main issue is the additive error of constant magnitude incurred in all the reductions of [22]. For a CSP, the objective function is guaranteed to be at least a fixed constant fraction (say 0.5). Hence, it is sufficient if the additive error term(say η) in the reduction can be bounded by an arbitrarily small constant. In case of $K_{M,N}$-Quadratic Programming, the value of the optimum solution could be as small as $1/\log n$. Here an additive constant error would completely change the approximation ratio.

To obtain better bounds on the error, we use a bootstrapping argument similar to the Gaussian Hilbert space approach to the Grothendieck inequality [5] (this approach is used for algorithmic purposes in [2, 1, 13]). Using ideas from the proof of the Grothendieck inequality, we perform a tighter analysis of the reduction in [22] for the special case of $K_{M,N}$-Quadratic Programming. This tight analysis yields the following new results:

Theorem 1.2. *For every $\eta > 0$, there is an efficient algorithm that achieves an approximation ratio $K_G - \eta$ for $K_{M,N}$-Quadratic Programming running in time $F(\eta) \cdot poly(n)$ where $F(\eta) = \exp(\exp(O(1/\eta^3)))$.*

Theorem 1.3. *For every $\eta > 0$, the Grothendieck constant K_G can be computed within an error η in time proportional to $\exp(\exp(O(1/\eta^3)))$.*

A more careful analysis could lower the degree of the polynomial $O(1/\eta^3)$ in the above bounds, but reducing the number of exponentiations seems to require new ideas.

With the intent of characterizing the tight cases for the Grothendieck inequality, we perform a non-standard reduction from dictatorship tests to integrality gaps. Unlike the reduction in [22], our reduction does not use the Khot–Vishnoi [12] integrality gap instance for Unique games. This new reduction yields a simple family of operators which are guaranteed to contain the tight cases for the Grothendieck inequality. Specifically, we show the following result:

Theorem 1.4. *Let $\mathcal{Q}^{(k)}$ be the set of linear operators A on functions $f\colon \mathbb{R}^k \to \mathbb{R}$ of the form $A = \sum_{d=0}^{\infty} \lambda_d Q_d$, where Q_d is the orthogonal projector on the span of k-multivariate Hermite polynomials of degree d. There exists operators in $\mathcal{Q}^{(k)}$ for which the Grothendieck inequality is near tight. More precisely, for every η,* *there exists an operator $A \in \mathcal{Q}^{(k)}$ for some k, such that*

$$\sup_{f\colon \mathbb{R}^k \to B^{(d)}} \int \|Af(\boldsymbol{x})\| \,\mathrm{d}\gamma(\boldsymbol{x}) \geqslant (K_G - \eta) \cdot \sup_{f\colon \mathbb{R}^k \to [-1,1]} \int |Af(\boldsymbol{x})| \,\mathrm{d}\gamma(\boldsymbol{x}) \,.$$

Here γ denotes the k-dimensional Gaussian probability measure, and for a function $f\colon \mathbb{R}^k \to \mathbb{R}^d$, we denote by $Af(\boldsymbol{x})$ the vector $(Af_1(\boldsymbol{x}), \ldots, Af_d(\boldsymbol{x}))$ where $f_1, \ldots, f_d$ are the coordinates of f.

We remark that Theorem 1.4 can also be shown in a direct way without using dictatorship tests (details in the full version). We can strengthen the statement of Theorem 1.4 in the following way: For every $\eta > 0$, there exists a linear operator A on functions $f\colon \mathbb{R}^k \to \mathbb{R}$ of the form $A = \sum_{i=0}^{\infty} \lambda_{2i+1} Q_{2i+1}$ such that (1) $k = \mathrm{poly}(1/\eta)$, (2) $\lambda_1 = \max_d |\lambda_d|$, and (3) $K_G = \lambda_1 / \|A\|_{\infty\to 1} \pm \eta$.

In [14], some evidence is given that the operator $A = \sum_{i=0}^{\infty} (-1)^i Q_{2i+1}$ is a tight instance for Grothendieck's inequality when k tends to ∞.

1.2 Prior Work

The general Grothendieck problem on a graph G amounts to maximizing a quadratic polynomial $\sum_{ij} a_{ij} x_i x_j$ over $\{1, -1\}$ values, where a_{ij} is non zero only for edges (i,j) in G. The $K_{M,N}$-Quadratic Programming is the special case where G is a complete bipartite graph $K_{M,N}$.

The Grothendieck problem on a complete graph admits a $O(\log n)$ approximation [21, 18, 6] and has applications in correlation clustering [6]. For the Grothendieck problem on general graphs, [1] obtain an approximation that depends on the Lovász number of the graph.

In an alternate direction, the Grothendieck problem has been generalized to the L_p-Grothendieck problem where the L_p-norm of the assignment is bounded by 1. The traditional Grothendieck corresponds to the case when $p = \infty$. In a recent work, [13] obtain UGC-based hardness results and approximation algorithms for the L_p-Grothendieck problem.

On the hardness side, [3] show a $O(\log^{\gamma} n)$ hardness for the Grothendieck problem on the complete graph for some fixed constant $\gamma > 0$. Tight integrality gaps for the Grothendieck problem on complete graphs were exhibited in [11, 1]. For the $K_{N,N}$-Quadratic Programming problem, a UGC-based hardness of roughly 1.67 was shown in [11]. The reduction uses the explicit operator constructed in the proof of a lower bound [23] for the Grothendieck constant.

1.3 Organization of the Paper

In Section 2, we formally define the Grothendieck constant, and review the notions of Noise Operators, Hermite polynomials, Multilinear extensions and influences. The overall structure of the reductions, along with key definitions and lemmas are described in Section 3. This overview includes reductions from integrality gaps to dictatorships (Subsection 3.1) and vice versa (Subsection 3.2). Using these reductions, we outline the proofs of Theorems 1.1, 1.2, 1.3 and 1.4 in Section 3. Finally, in Sections 4, 5 we present the proof details for the reduction from integrality gaps to dictatorship tests and vice versa.

2 Preliminaries

Problem 1. ($K_{M,N}$-Quadratic Programming) *Given an $m \times n$ matrix $A = (a_{ij})$, compute the optimal value of the following optimization problem,*

$$\mathrm{opt}(A) := \max \sum_{ij} a_{ij} x_i y_j \,,$$

where the maximum is over all $x_1, \ldots, x_m \in [-1,1]$ and $y_1, \ldots, y_n \in [-1,1]$. Note that the optimum value $\mathrm{opt}(A)$ is attained for numbers with $|x_i| = |y_j| = 1$.

Problem 2. ($K_{M,N}$-SemidefiniteProgramming) *Given an $m \times n$ matrix $A = (a_{ij})$, compute the optimal value of the following optimization problem,*

$$\mathrm{sdp}(A) := \max \sum_{ij} a_{ij} \langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle \,,$$

where the maximum is over all vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m \in B^{(d)}$ and all vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in B^{(d)}$. Here $B^{(d)}$ denotes the unit ball in $\mathbb{R}^d$ and we choose $d \geqslant m + n$. Note that the optimum value $\mathrm{sdp}(A)$ is always attained for vectors with $\|\boldsymbol{u}_i\| = \|\boldsymbol{v}_j\| = 1$.

For every matrix A, we have $\mathrm{opt}(A) \leqslant \mathrm{sdp}(A)$. Hence, $K_{M,N}$-SemidefiniteProgramming is a relaxation of $K_{M,N}$-Quadratic Programming. The value $\mathrm{sdp}(A)$ can be computed in polynomial time (up to arbitrarily small numerical error).

Definition 2.1. *The Grothendieck constant K_G is the supremum of $\mathrm{sdp}(A)/\mathrm{opt}(A)$ over all matrices A.*

2.1 Notation.

For a probability space Ω, let $L_2(\Omega)$ denote the Hilbert space of real-valued random variables over Ω with finite second moment,

$$L_2(\Omega) := \{f\colon \Omega \to \mathbb{R} \mid \mathop{\mathbb{E}}_{\omega \leftarrow \Omega} f(\omega)^2 < \infty\}$$

Here, we will consider two kinds of probability spaces. One is the uniform distribution over the Hamming cube $\{1,-1\}^k$, denoted $\Omega = \boldsymbol{\mathcal{H}}^k$. The other one is the Gaussian distribution over $\mathbb{R}^k$, denoted $\Omega = \boldsymbol{\mathcal{G}}^k$. For $f, g \in L_2(\Omega)$, we denote $\langle f, g\rangle := \mathbb{E}\, fg$, $\|f\| := \sqrt{\mathbb{E}\, f^2}$, and $\|f\|_\infty := \sup_{\boldsymbol{x}\in\Omega} f(\boldsymbol{x})$. We have $\|f\| \leqslant \|f\|_\infty$.

Lemma 2.1. (Bootstrapping Lemma) *Given an $m \times n$ matrix $A = (a_{ij})$, and vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, then*

$$\sum_{ij} a_{ij} \langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle \leqslant \big(\max_i \|\boldsymbol{u}_i\|\big)\big(\max_j \|\boldsymbol{v}_j\|\big) \cdot \mathrm{sdp}(A)$$
$$\leqslant 2\big(\max_i \|\boldsymbol{u}_i\|\big)\big(\max_j \|\boldsymbol{v}_j\|\big) \cdot \mathrm{opt}(A)$$

Definition 2.2. (Noise Operator) *For $\Omega = \boldsymbol{\mathcal{H}}^k$ or $\Omega = \boldsymbol{\mathcal{G}}^k$, let T_ρ denote the linear operator on $L_2(\Omega)$ defined as*

$$T_\rho := \sum_{d=0}^{k} \rho^d P_d \,,$$

where P_d denotes the orthogonal projector on the subspace of $L_2(\Omega)$ spanned by the (multilinear) degree-d monomials $\big\{\chi_S(\boldsymbol{x}) := \prod_{i \in S} x_i \mid S \subseteq [k], |S| = d\big\}$.

Fact 2.1. *For every function $f \in L_2(\boldsymbol{\mathcal{H}}^k)$.*

$$(T_\rho f)(\boldsymbol{x}) = \mathop{\mathbb{E}}_{\boldsymbol{y} \sim_\rho \boldsymbol{x}} f(\boldsymbol{y}) \,.$$

Here, $\boldsymbol{y} \sim_\rho \boldsymbol{x}$ means that $\boldsymbol{y}$ is obtained from $\boldsymbol{x}$ by replacing each coordinate i independently with probability $1-\rho$ by a random sign. In particular, $\mathbb{E}_{\boldsymbol{y}\sim_\rho \boldsymbol{x}} \langle \boldsymbol{x}, \boldsymbol{y}\rangle = \rho$.

2.2 Hermite Polynomials and Gaussian Noise operator. Let $\boldsymbol{\mathcal{G}}$ be the probability space over $\mathbb{R}$ with Gaussian probability measure. The set of (univariate) *Hermite polynomials* $\{H_d \mid d \in \mathbb{N}\}$ forms an orthonormal basis for $L_2(\boldsymbol{\mathcal{G}})$. The degree of $H_d \in \mathbb{R}[x]$ is equal to d. The first Hermite polynomials are 1, x, x^2-1, and x^3-3x. An orthogonal basis for $L_2(\boldsymbol{\mathcal{G}}^k)$ is given by the set of functions $\{H_\sigma(\boldsymbol{x}) := \prod_{i=1}^k H_{\sigma_i}(x_i) \mid \sigma \in \mathbb{N}_0^k\}$.

Definition 2.3. (Gaussian Noise Operator) *Let U_ρ denote the linear operator on $L_2(\boldsymbol{\mathcal{G}}^k)$ defined as*

$$U_\rho := \sum_{d=0}^{k} \rho^d Q_d \,,$$

where Q_d denotes the orthogonal projector on the subspace of $L_2(\boldsymbol{\mathcal{H}}^k)$ spanned by the set of k-variate degree-d Hermite polynomials $\big\{H_\sigma(\boldsymbol{x}) \mid \sigma \in \mathbb{N}_0^k, \sum \sigma_i = d\big\}$.

Fact 2.2. *For every function $f \in L_2(\boldsymbol{\mathcal{G}}^k)$.*

$$(U_\rho f)(\boldsymbol{x}) = \mathop{\mathbb{E}}_{\boldsymbol{y} \sim_\rho \boldsymbol{x}} f(\boldsymbol{y}) \,.$$

Here, $\boldsymbol{y} \sim_\rho \boldsymbol{x}$ means that $\boldsymbol{y}$ can be written as $\boldsymbol{y} = \rho \boldsymbol{x} + \sqrt{1-\rho^2}\boldsymbol{z}$ for a random Gaussian vector $\boldsymbol{z}$. In particular, $\mathbb{E}_{\boldsymbol{y}\sim_\rho \boldsymbol{x}} \langle \boldsymbol{x}, \boldsymbol{y}\rangle = \rho$.

2.3 Variable Influences. For a function $f \in L_2(\boldsymbol{\mathcal{H}}^k)$, we define $\mathrm{Inf}_i f = \sum_{S \ni i} \hat{f}_S^2$, where $\hat{f}_S$ are the Fourier coefficients of f,

$$f = \sum_{S \subseteq [k]} \hat{f}_S \chi_S \,.$$

Let us denote $\mathrm{MaxInf}\, f := \max_{i \in [k]} \mathrm{Inf}_i f$. For a pair of functions $f, g \in L_2(\boldsymbol{\mathcal{H}}^k)$, we define $\mathrm{MaxComInf}(f,g) := \max_{i\in[k]} \min\{\mathrm{Inf}_i f, \mathrm{Inf}_i g\}$ to be the *maximum common influence.*

Similarly, for $f \in L_2(\boldsymbol{\mathcal{G}}^k)$, we denote by $\mathrm{Inf}_i f = \sum_{\sigma; \sigma_i > 0} \hat{f}_\sigma^2$ the *influence* of coordinate i. Here, $\hat{f}_\sigma$ are the Hermite coefficients of f,

$$f = \sum_{\sigma \in \mathbb{N}_0^k} \hat{f}_\sigma H_\sigma \,.$$

Fact 2.3. *For $f \in L_2(\boldsymbol{\mathcal{H}}^k)$ and $\gamma \in [0,1]$, we have $\sum_{i=1}^k \mathrm{Inf}_i T_{1-\gamma} f \leqslant \|f\|^2/\gamma$. Similarly, for $f \in L_2(\boldsymbol{\mathcal{G}}^k)$ and $\gamma \in [0,1]$, $\sum_{i=1}^k \mathrm{Inf}_i U_{1-\gamma} f \leqslant \|f\|^2/\gamma$.*

2.4 Multilinear Extensions. For $f \in L_2(\boldsymbol{\mathcal{H}}^k)$, let $\bar{f} \in L_2(\boldsymbol{\mathcal{G}}^k)$ denotes the (unique) multilinear extension of f to $\mathbb{R}^k$.

Lemma 2.2. *Let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^d$ be two unit vectors, and $f, g \in L_2(\boldsymbol{\mathcal{H}}^k)$. Then,*

$$\mathop{\mathbb{E}}_{\Phi} \bar{f}(\Phi \boldsymbol{u}) \bar{g}(\Phi \boldsymbol{v}) = \langle f, T_{\langle \boldsymbol{u}, \boldsymbol{v}\rangle} g \rangle \qquad (2.1)$$

where Φ is a $k \times d$ Gaussian matrix, that is, the entries of Φ are mutually independent normal variables with standard deviation $\frac{1}{\sqrt{d}}$.

Proof. Note that $\Phi \boldsymbol{u} \sim_\rho \Phi \boldsymbol{v}$ for $\rho = \langle \boldsymbol{u}, \boldsymbol{v}\rangle$. Hence, the left-hand side of equation (2.1) is equal to $\langle \bar{f}, U_\rho \bar{g}\rangle$. Since $\bar{g}$ is multilinear, we have $Q_d \bar{g} = P_d \bar{g}$. Therefore, $\langle \bar{f}, U_\rho \bar{g}\rangle = \langle f, T_\rho g\rangle$, as desired.

2.5 Truncation of Low-influence Functions. For $f \colon \mathbb{R}^k \to \mathbb{R}$, let $\mathrm{trunc}\, f \colon \mathbb{R}^k \to [-1,1]$ denote the function

$$\mathrm{trunc}\, f(\boldsymbol{x}) := \begin{cases} 1 & \text{if } f(\boldsymbol{x}) > 1\,, \\ f(\boldsymbol{x}) & \text{if } -1 < f(\boldsymbol{x}) < 1\,, \\ -1 & \text{if } f(\boldsymbol{x}) < -1\,. \end{cases}$$

In our context, the invariance principle [20] roughly says that if f is a bounded function on $\boldsymbol{\mathcal{H}}^k$ with no influential coordinate, then the multilinear extension of f as function on $\boldsymbol{\mathcal{G}}^k$ is close to a bounded function (its truncation).

THEOREM 2.1. (INVARIANCE PRINCIPLE, [20]) *There is a universal constant C such that, for all $\rho = 1-\gamma \in (0,1)$ the following holds: Let $f \in L_2(\boldsymbol{\mathcal{H}}^k)$ with $\|f\|_\infty \leqslant 1$ and $\mathrm{Inf}_i(T_\rho f) \leqslant \tau$ for all $i \in [k]$. Then,*

$$\left\|T_\rho \bar{f} - \operatorname{trunc} T_\rho \bar{f}\right\| \leqslant \tau^{C \cdot \gamma}$$

where $\bar{f} \in L_2(\boldsymbol{\mathcal{G}}^k)$ denotes the (unique) multilinear extension of f to $\mathbb{R}^k$.

3 Proof Overview

In this section, we will outline the overall structure of the reductions, state the key definitions and lemmas, and show how they connect with each other. In the subsequent sections, we will present the proofs of the lemmas used. The overall structure of the reduction is along the lines of [22]. We begin by defining dictatorship tests in the current context.

DEFINITION 3.1. *A* dictatorship test *B is an operator on $L_2(\boldsymbol{\mathcal{H}}^k)$ of the following form:*

$$B = \sum_{d=0}^{k} \lambda_d P_d \qquad (\lambda_1 \geqslant |\lambda_d| \text{ for all } d)$$

where P_d is the projection operator on to the degree-d part. We define two parameters of B:

$$\mathsf{Completeness}(B) := \inf_i \langle \chi_i, B\chi_i \rangle = \lambda_1 \,,$$

where $\chi_i(\boldsymbol{x}) = x_i$ is the i^{th} dictator function, and

$$\mathsf{Soundness}_{\eta,\tau}(B) := \sup_{\substack{f,g \in L_2(\boldsymbol{\mathcal{H}}^k),\\ \|f\|_\infty, \|g\|_\infty \leqslant 1,\\ \mathrm{MaxComInf}(T_\rho f, T_\rho g) \leqslant \tau}} \langle T_\rho f, B T_\rho g \rangle \,,$$

where $\rho = 1-\eta$.

3.1 From Integrality Gaps to Dictatorship Tests:

In the first step, we describe a reduction from a matrix A of arbitrary size, to a dictatorship test $\mathcal{D}(A)$ on $L_2(\boldsymbol{\mathcal{H}}^k)$ for a constant k independent of the size of A.

Towards this, let us set up some notation. Let $A = (a_{ij})$ be an $m \times n$ matrix with SDP value $\mathrm{sdp}(A)$. Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m \in B^{(d)}$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in B^{(d)}$ be an SDP solution such that

$$\sum_{ij} a_{ij} \langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle = \mathrm{sdp}(A) \,.$$

In general, an optimal SDP solution $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ might not be unique. In the following, we will however assume that for every instance A we can uniquely associate an optimal SDP solution, e.g., the one computed by a given implementation of the ellipsoid method.

With this notation, we are ready to define the dictatorship test $\mathcal{D}(A)$.

DEFINITION 3.2. *For $d \in \mathbb{N}$, let us define coefficients $\lambda_d \in \mathbb{R}$,*

$$\lambda_d := \sum_{ij} a_{ij} \langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle^d \,.$$

Define linear operators $\mathcal{D}(A), \mathcal{D}_\eta(A)$ on $L_2(\boldsymbol{\mathcal{H}}^k)$,

$$\mathcal{D}(A) := \sum_{d=0}^{k} \lambda_d P_d = \sum_{ij} a_{ij} T_{\langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle}$$

$$\mathcal{D}_\eta(A) := T_\rho \mathcal{D}(A) T_\rho = \sum_{d=0}^{k} \rho^{2d} \lambda_d P_d \,,$$

where $\rho = 1-\eta$.

By the definition of $\mathsf{Completeness}(\mathcal{D}_\eta(A))$, we have:

LEMMA 3.1. *For all matrices A,*

$$\mathsf{Completeness}(\mathcal{D}_\eta(A)) = \lambda_1 \rho^2 \geqslant \mathrm{sdp}(A)(1-2\eta) \,.$$

Towards bounding $\mathsf{Soundness}_{\eta,\tau}(\mathcal{D}_\eta(A))$, we define a rounding scheme $\mathsf{Round}_{\eta,f,g}$ for every pair of functions $f, g \in L_2(\boldsymbol{\mathcal{H}}^k)$ and $\eta > 0$. The rounding scheme $\mathsf{Round}_{\eta,f,g}$ is an efficient randomized procedure that takes as input the optimal SDP solution for A, and outputs a solution $x_1, \ldots, x_m, y_1, \ldots, y_n \in [-1, 1]$. The details of the randomized rounding procedure are described in Section 4.

DEFINITION 3.3. *$\mathsf{Round}_{\eta,f,g}(A)$ is the expected value of the solution returned by the randomized rounding procedure $\mathsf{Round}_{\eta,f,g}$ on the input A.*

The following relationship between performance of rounding schemes and soundness of the dictatorship test is proven using Theorem 2.1 (invariance principle [20]).

THEOREM 3.1. *Let A be a matrix. For functions $f, g \in L_2(\boldsymbol{\mathcal{H}}^k)$ satisfying $\|f\|_\infty, \|g\|_\infty \leqslant 1$ and $\mathrm{MaxComInf}(T_\rho f, T_\rho g) \leqslant \tau$ for $\rho = 1-\eta$, there exists functions $f', g' \in L_2(\boldsymbol{\mathcal{H}}^k)$ such that*

$$\langle f, \mathcal{D}_\eta(A) g \rangle \leqslant \mathsf{Round}_{\eta,f',g'}(A) + \left({}^{10\tau^{C\eta/8}}\!/\!\sqrt{\eta}\right) \cdot \mathrm{sdp}(A) \,.$$

Further, the functions f', g' satisfy $\|f\|_\infty, \|g\|_\infty \leqslant 1$.

By taking the supremum on both sides of the above inequality over all low influence functions, one obtains the following corollary.

Corollary 3.1. *For every matrix A and $\eta > 0$,*

$$\mathsf{Soundness}_{\eta,\tau}(\mathcal{D}_\eta(A)) \leqslant \sup_{\substack{f,g \in L_2(\boldsymbol{\mathcal{H}}^k),\\ \|f\|_\infty, \|g\|_\infty \leqslant 1}} \mathsf{Round}_{\eta,f,g}(A) + \frac{10\tau^{C\eta/8}\mathrm{sdp}(A)}{\sqrt{\eta}},$$

As $\mathsf{Round}_{\eta,f,g}$ is the expected value of a $[-1,1]$ solution, it is necessarily at most $\mathrm{opt}(A)$. Further by Grothendieck's inequality, $\mathrm{sdp}(A)$ and $\mathrm{opt}(A)$ are within constant factor of each other. Together, these facts immediately imply the following corollary:

Corollary 3.2. *For $\eta > 0$, if $\tau \leqslant 2^{-100\log(1/\eta)/C\eta}$, then for all matrices A,*

$$\mathsf{Soundness}_{\eta,\tau}(\mathcal{D}_\eta(A)) \leqslant \mathrm{opt}(A)(1+\eta).$$

3.2 From Dictatorship Tests to Integrality Gaps The next key step is the conversion from arbitrary dictatorship tests back to integrality gaps. Unlike many previous works [22], we obtain a simple direct conversion without using the unique games hardness reduction or the Khot–Vishnoi integrality gap instance. In fact, the integrality gap instances produced have the following simple description:

Definition 3.4. *Given an dictatorship test B on $L_2(\boldsymbol{\mathcal{H}}^k)$ of the form $B = \sum_{d=0}^{k} \lambda_d P_d$, define the corresponding operator $\mathcal{G}_\eta(B)$ on $L_2(\boldsymbol{\mathcal{G}}^k)$ as*

$$\mathcal{G}_\eta(B) = \sum_d \lambda_d Q_d \rho^{2d},$$

where $\rho = 1 - \eta$.

We present the proof of the following theorem in Section 5.

Theorem 3.2. *For all $\eta > 0$, there exists k, τ such that following holds: For any dictatorship test B on $L_2(\boldsymbol{\mathcal{H}}^k)$, we have:*

$$\mathrm{sdp}(\mathcal{G}(B)) \geqslant \mathsf{Completeness}(B)\,(1-5\eta)\,, \tag{3.2}$$

$$\mathrm{opt}(\mathcal{G}(B)) \leqslant \mathsf{Soundness}_{\eta,\tau}(B)\,(1+\eta) + \eta\mathsf{Completeness}(B)\,. \tag{3.3}$$

In particular, the choices $\tau = 2^{-100/\eta^3}$ and $k = 2^{200/\eta^3}$ suffice.

By Grothendieck's theorem, the ratio of $\mathrm{sdp}(\mathcal{G}(B))$ and $\mathrm{opt}(\mathcal{G}(B))$ is at most K_G. Hence as a simple corollary, one obtains the following limit to dictatorship testing:

Corollary 3.3. *For all $\eta > 0$, there exists k, τ such that for any dictatorship test B on $L_2(\boldsymbol{\mathcal{H}}^k)$,*

$$\frac{\mathsf{Soundness}_{\eta,\tau}(B)}{\mathsf{Completeness}(B)} \geqslant \frac{1}{K_G} - \eta\,. \tag{3.4}$$

From the above corollary, we know that $\mathsf{Soundness}_{\eta,\tau}(B)$ and $\mathsf{Completeness}(B)$ are within constant factors of each other. Consequently, we have

Corollary 3.4. *The equation (3.3) can be replaced by*

$$\mathrm{opt}(\mathcal{G}(B)) \leqslant \mathsf{Soundness}_{\eta,\tau}(B)\,(1+5\eta)\,.$$

We present the proof of the Theorems 1.2 to illustrate how the two conversions outlined in this section come together. The proofs of the remaining theorems are deferred to the full version.

3.3 Proof of Theorem 1.2 Consider the following idealized algorithm for the $K_{N,N}$-Quadratic Programming problem

- Find the optimal SDP solution $\boldsymbol{u}_i, \boldsymbol{v}_j$
- Fix $k = 2^{200/\eta^3}$ and $\tau = 2^{-100/\eta^3}$. For every pair of functions $f, g \in L_2(\boldsymbol{\mathcal{H}}^k)$ with $\|f\|, \|g\| \leqslant 1$, run the rounding scheme $\mathsf{Round}_{\eta,f,g}(A)$ to obtain a $[-1,1]$ solution. Output the solution with the largest value.

The value of the solution obtained is given by $\sup_{f,g\in L_2(\boldsymbol{\mathcal{H}}^k)} \mathsf{Round}_{\eta,f,g}(A)$. From Corollary 3.1, we have

$$\begin{aligned}
&\sup_{f,g\in L_2(\boldsymbol{\mathcal{H}}^k), \|f\|,\|g\|\leqslant 1} \mathsf{Round}_{\eta,f,g}(A) \\
&\geqslant \sup_{f,g\in L_2(\boldsymbol{\mathcal{H}}^k), \|f\|_\infty,\|g\|_\infty\leqslant 1} \mathsf{Round}_{\eta,f,g}(A) \\
&\geqslant \mathsf{Soundness}_{\eta,\tau}(\mathcal{D}_\eta(A)) \\
&\quad - \left(10\tau^{C\eta/8}/\sqrt{\eta}\right)\cdot \mathrm{sdp}(A)\,.
\end{aligned} \tag{3.5}$$

From Lemma 3.1, we know $\mathsf{Completeness}(\mathcal{D}_\eta(A)) \geqslant \mathrm{sdp}(A)(1-2\eta)$. By the choice of k, τ, we can apply Corollary 3.3 on $\mathcal{D}_\eta(A)$ to conclude

$$\mathsf{Soundness}_{\eta,\tau}(\mathcal{D}_\eta(A)) \geqslant \mathsf{Completeness}(\mathcal{D}_\eta(A))\,(1/K_G - \eta) \tag{3.6}$$

$$\geqslant \mathrm{sdp}(A)\,(1/K_G - \eta)\,(1-2\eta) \tag{3.7}$$

From equations (3.5) and (3.6), we conclude that the value returned by the idealized algorithm is at least

$$\mathrm{sdp}(A)\left((1/K_G - \eta)\,(1-\eta) - 10\tau^{C\eta/8}/\sqrt{\eta}\right),$$

which by the choice of τ is at least $\mathrm{sdp}(A)(1/K_G - 4\eta)$.

In order to implement the idealized algorithm, we discretize the unit ball in space $L_2(\boldsymbol{\mathcal{H}}^k)$ using a η-net in the L_2-norm. As k is a fixed constant depending on η, there is a finite η-net that would serve the purpose. To finish the argument, one needs to show that the value of the solution returned is not affected by the discretization. This follows from the next lemma whose proof is deferred to the full version:

LEMMA 3.2. *For functions $f, g, f', g' \in L_2(\boldsymbol{\mathcal{H}}^k)$ with $\|f\|, \|g\|, \|f'\|, \|g'\| \leqslant 1$,*

$$|\mathsf{Round}_{\eta,f,g}(A) - \mathsf{Round}_{\eta,f',g'}(A)| \leqslant \mathrm{sdp}(A)(\|f - f'\| + \|g - g'\|)\,.$$

3.4 Proof of Theorem 1.1 As a rule of thumb, every dictatorship test yields a UG hardness result using by now standard techniques [10, 11, 22]. Specifically, we can show the following :

LEMMA 3.3. *Given a dictatorship test A and a unique games instance G, it is possible to efficiently construct an operator $G \otimes_\eta A$ that satisfies the following to two conditions:*

1. *if* $\mathrm{val}(G) \geqslant 1 - \epsilon$, *then*
$$\mathrm{opt}(G \otimes_\eta A) \geqslant \mathsf{Completeness}(A)(1 - o_{\epsilon,\eta\to 0}(1)),$$
2. *if* $\mathrm{val}(G) < \epsilon$, *then*
$$\mathrm{opt}(G \otimes_\eta A) < \mathsf{Soundness}_{\eta,\tau}(A)(1 + o_{\epsilon,\eta,\tau\to 0}(1))\,.$$

Due to space constraints, we omit the proof of the above lemma here.

To finish the proof of Theorem 1.1, let A be a matrix for which the ratio of $\mathrm{sdp}(A)/\mathrm{opt}(A) \geqslant K_G - \eta$. Consider the dictatorship test $\mathcal{D}_\eta(A)$ obtained from the matrix A. By Corollary 3.1, the completeness of $\mathcal{D}_\eta(A)$ is $\mathrm{sdp}(A)(1-\eta)$. Further by Corollary 3.2, the soundness is at most $\mathrm{opt}(A)(1+\eta)$ for sufficiently small choice of τ. Plugging this dictatorship test $\mathcal{D}_\eta(A)$ in to the above lemma, we obtain a UG hardness of $(K_G - \eta)(1-\eta)/(1+\eta) \geqslant K_G - 5\eta$. Since η can be made arbitrarily small, the proof is complete.

3.5 Proof of Theorem 1.4 Let A be an arbitrary finite matrix for which $\mathrm{sdp}(A)/\mathrm{opt}(A) \geqslant K_G - \eta$. Consider the dictatorship test/operator $\mathcal{D}_\eta(A)$ on $L_2(\boldsymbol{\mathcal{H}}^k)$. From Lemma 3.1 and Corollary 3.2, the ratio of $\mathsf{Completeness}(A)$ to $\mathsf{Soundness}_{\eta,\tau}(A)$ is at least $\mathrm{sdp}(A)/\mathrm{opt}(A) - 2\eta$ for sufficiently small choice of τ. Further it is easy to see that the operator $\mathcal{D}_\eta(A)$ is translation invariant by construction. Now using Theorem 3.2, for large enough choice of k, the operator $\mathcal{G}(\mathcal{D}_\eta(A))$ is an operator with $\mathrm{sdp}(A)/\mathrm{opt}(A) \geqslant K_G - 10\eta$. Since the operator $\mathcal{G}(\mathcal{D}_\eta(A))$ belongs to the set $\mathcal{Q}^{(k)}$, this completes the proof of Theorem 1.4.

3.6 Proof of Theorem 1.3 A naive approach to compute the Grothendieck constant, is to iterate over all matrices A and compute the largest possible value of $\mathrm{sdp}(A)/\mathrm{opt}(A)$. However, the set of all matrices is an infinite set, and there is no guarantee on when to terminate.

As there is a conversion from integrality gaps to dictatorship tests and vice versa, instead of searching for the matrix with the worst integrality gap, we shall find the dictatorship test with the worst possible ratio between completeness and soundness. Recall that a dictatorship test is an operator on $L_2(\boldsymbol{\mathcal{H}}^k)$ for a finite k depending only on η the error incurred in the reductions. In principle, this already shows that the Grothendieck constant is computable up to an error η in time depending only on η.

Define K as follows

$$\frac{1}{K} = \inf_{\substack{\lambda_1=1,\\ \lambda_d\in[-1,1]}} \sup_{\substack{f,g\in L_2(\boldsymbol{\mathcal{H}}^k),\\ \mathrm{MaxComInf}(T_\rho f, T_\rho g)\leqslant\tau \\ \|f\|_\infty,\|g\|_\infty\leqslant 1}} \langle f, \sum_{d=0}^{k} \rho^{2d}\lambda_d Q_d g\rangle\,, \qquad ,$$

where $\rho = 1 - \eta$.

Let $\mathcal{P}$ denote the space of all pairs of functions $f, g \in L_2(\boldsymbol{\mathcal{H}}^k)$ with $\mathrm{MaxComInf}(T_\rho f, T_\rho g) \leqslant \tau$ and $\|f\|_\infty, \|g\|_\infty \leqslant 1$. Since $\mathcal{P}$ is a compact set, there exists an η-net of pairs of functions $\mathcal{F} = \{(f_1, g_1), \ldots, (f_N, g_N)\}$ such that: For every point $(f, g) \in \mathcal{P}$, there exists $f_i, g_i \in \mathcal{F}$ satisfying $\|f - f'\| + \|g - g'\| \leqslant \eta$. The size of the η-net is a constant depending only on k and η (note: k depends only on η).

The constant K can be expressed up to an error of $O(\eta)$ using the following finite linear program:

$$\begin{aligned}
&\text{Minimize } \frac{1}{K} = \mu \\
&\text{Subject to } \mu \geqslant \sum_{d=0}^{k} \lambda_d \cdot \langle f, \sum_{d=0}^{k} \rho^{2d} Q_d g\rangle \\
&\qquad\qquad \text{for all functions } f, g \in \mathcal{F}\,, \\
&\qquad \lambda_i \in [-1, 1] \qquad \text{for all } 0 \leqslant i \leqslant k\,, \\
&\qquad \lambda_1 = 1\,.
\end{aligned}$$

4 From Integrality gaps to Dictatorship Tests

4.1 Rounding Scheme For functions $f, g \in L_2(\boldsymbol{\mathcal{H}}^k)$, define the rounding procedure $\mathsf{Round}_{\eta,f,g}$ as follows:

> $\mathsf{Round}_{\eta,f,g}$
> Input: An $m \times n$ matrix $A = (a_{ij})$ with SDP solution $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\}, \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} \subseteq B^{(d)}$
>
> - Compute $\bar{f}, \bar{g}$ the multilinear extensions of f, g.
> - Generate $k \times d$ matrix Φ all of whose entries are mutually independent normal variables of standard deviation $1/\sqrt{d}$.
> - Output the assignment
>
> $$x_i = \operatorname{trunc} T_\rho \bar{f}(\Phi \boldsymbol{u}_i)\,, \qquad y_j = \operatorname{trunc} T_\rho \bar{g}(\Phi \boldsymbol{v}_j)\,.$$

The expected value of the solution returned $\mathsf{Round}_{\eta,f,g}(A)$ is given by:

$$\mathsf{Round}_{\eta,f,g}(A) = \mathop{\mathbb{E}}_{\Phi} \textstyle\sum_{ij} a_{ij} \operatorname{trunc} T_\rho \bar{f}(\Phi \boldsymbol{u}_i) \operatorname{trunc} T_\rho \bar{g}(\Phi \boldsymbol{v}_j).$$

4.2 Relaxed Influence Condition The following lemma shows that we could replace the condition $\operatorname{MaxComInf}(T_\rho f, T_\rho g) \leqslant \tau$ in Definition 3.1 by the condition $\operatorname{MaxInf} T_\rho f, \operatorname{MaxInf} T_\rho g \leqslant \sqrt{\tau}$ with a small loss in the soundness. The proof is omitted here due to space constraints.

LEMMA 4.1. *Let A be a dictatorship test on $L_2(\boldsymbol{\mathcal{H}}^k)$, and let f, g be a pair of functions in $L_2(\boldsymbol{\mathcal{H}}^k)$ with $\|f\|_\infty, \|g\|_\infty \leqslant 1$ and $\operatorname{MaxComInf}(T_\rho f, T_\rho g) \leqslant \tau$ for $\rho = 1 - \eta$. Then for every $\tau' > 0$, there are functions $f', g' \in L_2(\boldsymbol{\mathcal{H}}^k)$ with $\|f'\|_\infty, \|g'\|_\infty \leqslant 1$ and $\operatorname{MaxInf} T_\rho f', \operatorname{MaxInf} T_\rho g' \leqslant \tau'$ such that*

$$\langle T_\rho f', A T_\rho g' \rangle \geqslant \langle T_\rho f, A T_\rho g \rangle - 2\|A\| \sqrt{\tau/\tau' \eta}\,.$$

With this background, we now present the soundness analysis.

4.3 Proof of Theorem 3.1

Proof. By Lemma 4.1, there exists function $f', g' \in L_2(\boldsymbol{\mathcal{H}}^k)$ with $\|f'\|_\infty, \|g'\|_\infty \leqslant 1$ and $\operatorname{MaxInf} T_\rho f', \operatorname{MaxInf} T_\rho g' \leqslant \sqrt{\tau}$ such that

$$\begin{aligned}\langle f', \mathcal{D}_\eta(A) g' \rangle &\geqslant \langle f, \mathcal{D}_\eta(A) g \rangle - 2\|\mathcal{D}(A)\| \cdot \tau^{1/4}/\sqrt{\eta} \\ &\geqslant \langle f, \mathcal{D}_\eta(A) g \rangle - 4\operatorname{opt}(A) \cdot \tau^{1/4}/\sqrt{\eta}\,.\end{aligned}$$

On the other hand, we have

$$\begin{aligned}\langle f', \mathcal{D}_\eta(A) g' \rangle &= \sum_{ij} \sum_{d=0}^{k} \langle T_\rho f', a_{ij} \langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle^d P_d(T_\rho g') \rangle \\ &= \sum_{ij} a_{ij} \langle T_\rho f', T_{\langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle}(T_\rho g') \rangle\end{aligned}$$

We can assume that all vectors $\boldsymbol{u}_i$ and $\boldsymbol{v}_j$ have unit norm. By Lemma 2.2 , we have

$$\mathop{\mathbb{E}}_{\Phi} \sum_{ij} a_{ij} T_\rho \bar{f}'(\Phi \boldsymbol{u}_i) T_\rho \bar{g}'(\Phi \boldsymbol{v}_j) \tag{4.8}$$

$$= \sum_{ij} a_{ij} \langle T_\rho f', T_{\langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle}(T_\rho g') \rangle \tag{4.9}$$

From the above equations we have

$$\langle f', \mathcal{D}_\eta(A) g' \rangle = \mathop{\mathbb{E}}_{\Phi} \sum_{ij} a_{ij} T_\rho \bar{f}'(\Phi \boldsymbol{u}_i) T_\rho \bar{g}'(\Phi \boldsymbol{v}_j) \tag{4.10}$$

By the invariance principle (Theorem 2.1), we have

$$\|T_\rho \bar{f}' - \operatorname{trunc} T_\rho \bar{f}'\| \leqslant \tau^{C\eta/2} \tag{4.11}$$

and

$$\|T_\rho \bar{g}' - \operatorname{trunc} T_\rho \bar{g}'\| \leqslant \tau^{C\eta/2}\,. \tag{4.12}$$

Now we shall apply the simple yet powerful bootstrapping trick. Let us define new vectors in $L_2(\boldsymbol{\mathcal{G}}^{k \times d})$,

$$(\boldsymbol{u}_i')_\Phi = T_\rho \bar{f}'(\Phi \boldsymbol{u}_i) \qquad (\boldsymbol{v}_j')_\Phi = T_\rho \bar{g}'(\Phi \boldsymbol{v}_j)$$

and

$$(\boldsymbol{u}_i'')_\Phi = \operatorname{trunc} T_\rho \bar{f}'(\Phi \boldsymbol{u}_i) \qquad (\boldsymbol{v}_j'')_\Phi = \operatorname{trunc} T_\rho \bar{g}'(\Phi \boldsymbol{v}_j)$$

Equation (4.11) implies that $\|\boldsymbol{u}_i' - \boldsymbol{u}_i''\| \leqslant \tau^{C\eta/2}$ and $\|\boldsymbol{v}_j' - \boldsymbol{v}_j''\| \leqslant \tau^{C\eta/2}$. Using the bootstrapping argument (Lemma 2.1), we finish the proof

$$\mathsf{Round}_{\eta,f',g'}(A) = \sum_{ij} a_{ij} \langle \boldsymbol{u}_i'', \boldsymbol{v}_j'' \rangle \tag{4.13}$$

$$\begin{aligned}&= \sum_{ij} a_{ij} \langle \boldsymbol{u}_i', \boldsymbol{v}_j' \rangle - \sum_{ij} a_{ij} \langle \boldsymbol{u}_i' - \boldsymbol{u}_i'', \boldsymbol{v}_j' \rangle - \sum_{ij} a_{ij} \langle \boldsymbol{u}_i'', \boldsymbol{v}_j' - \boldsymbol{v}_j'' \rangle \\ &\overset{(4.11)}{\geqslant} \sum_{ij} a_{ij} \langle \boldsymbol{u}_i', \boldsymbol{v}_j' \rangle - 2\tau^{C\eta/2}\operatorname{opt}(A) - 2\tau^{C\eta/2}\operatorname{opt}(A) \\ &\geqslant \langle f, \mathcal{D}_\eta A g \rangle - 4\tau^{C\eta/2}\operatorname{opt}(A) - 4\tau^{1/4}\operatorname{opt}(A)/\sqrt{\eta}\,.\end{aligned}$$

5 From Dictatorship Tests to Integrality Gaps

In this section, we outline the key ideas in the proof of Theorem 3.2. Due to space constraints, the details are deferred to the full version.

5.1 $\text{sdp}(\mathcal{G}(B)) \geqslant \mathsf{Completeness}(B)\,(1-5\eta)$**.** To prove this claim, we need to construct an SDP solution to $\text{sdp}(\mathcal{G}(B))$ that achieves nearly the same value as $\mathsf{Completeness}(B)$. Formally, we need to construct functions f, g whose domain is $\boldsymbol{\mathcal{G}}^t$ and outputs are unit vectors. Since we want to achieve a value close to $\mathsf{Completeness}(B) = \lambda_1$, the functions f, g should be linear or near-linear. Along the lines of [11, 23], we choose the following function $f(\boldsymbol{x}) = g(\boldsymbol{x}) = {}^{\boldsymbol{x}}/_{\|\boldsymbol{x}\|}$ which always outputs unit vectors, and very close to the linear function $\boldsymbol{x} \mapsto {}^{\boldsymbol{x}}/_{\sqrt{t}}$ as t increases. Formally, we show the following lemma:

Lemma 5.1.

$$\text{sdp}(\mathcal{G}(B)) \geqslant \mathsf{Completeness}(B)\left(\rho^4 - 2\left(\frac{\log t}{t}\right)^{\frac{1}{4}}\right)$$

For $t > 1/\eta^5$, the value of the SDP solution is at least $\mathsf{Completeness}(B)(1-5\eta)$.

5.2 $\text{opt}(\mathcal{G}(B)) \leqslant \mathsf{Soundness}_{\eta,\tau}(B)\,(1+\eta) + \eta\mathsf{Completeness}(B)$**.** For the sake of contradiction, let us suppose $\text{opt}(\mathcal{G}(B)) \geqslant \mathsf{Soundness}_{\eta,\tau}(B)(1+\eta) + \eta\mathsf{Completeness}(B)$. Let the optimum solution be given by two functions $f, g \in L_2(\boldsymbol{\mathcal{G}}^t)$. By assumption, we have $\|f\|_\infty, \|g\|_\infty \leqslant 1$ and

$$\langle f, \mathcal{G}(B) g\rangle \geqslant \mathsf{Soundness}_{\eta,\tau}(B)(1+\eta) + \eta\mathsf{Completeness}(B)\,.$$

To get a contradiction, we will construct low influence functions in $L_2(\boldsymbol{\mathcal{H}}^k)$ that have a objective value greater than $\mathsf{Soundness}_{\eta,\tau}(B)$ on the dictatorship test B. This construction is obtained in two steps:

In the first step, we obtain functions f', g' over a larger dimensional space with the same objective value but are also guaranteed to have no influential coordinates. This is achieved by defining f', g' as follows for large enough R,

$$f'(\boldsymbol{x}) = f\Big(\tfrac{1}{\sqrt{R}}\sum_{i=1}^{R} x_i, \tfrac{1}{\sqrt{R}}\sum_{i=R+1}^{2R} x_i, \ldots, \tfrac{1}{\sqrt{R}}\sum_{i=(R-1)t+1}^{Rt} x_i\Big)$$
$$g'(\boldsymbol{x}) = g\Big(\tfrac{1}{\sqrt{R}}\sum_{i=1}^{R} x_i, \tfrac{1}{\sqrt{R}}\sum_{i=R+1}^{2R} x_i, \ldots, \tfrac{1}{\sqrt{R}}\sum_{i=(R-1)t+1}^{Rt} x_i\Big).$$

For a sufficiently large choice of R (say $R = \lceil 1/\eta\tau \rceil$), the functions $f', g' \in L_2(\boldsymbol{\mathcal{G}}^{t'})$ have no influential coordinates. Formally, we show the following lemma:

Lemma 5.2. *Given two functions $f, g \in L_2(\boldsymbol{\mathcal{G}}^t)$ with $\|f\|_\infty, \|g\|_\infty \leqslant 1$, there exists $f', g' \in L_2(\boldsymbol{\mathcal{G}}^{t\cdot\lceil 1/(1-\rho)\tau\rceil})$ with $\|f\|'_\infty, \|g\|'_\infty \leqslant 1$ and $\max_i \text{Inf}_i(U_\rho f'), \max_j \text{Inf}_j(U_\rho g') \leqslant \tau$ and*

$$\langle f', \mathcal{G}(B) g'\rangle = \langle f, \mathcal{G}(B) g\rangle.$$

In the second step, we apply the invariance principle to construct functions on $L_2(\boldsymbol{\mathcal{H}}^k)$ with the same properties as f', g'. More precisely, we show

Lemma 5.3. *For any $\eta > 0$, there exists $D, \tau > 0$ such that the following holds for every operator $B = \sum_{d=0}^{tD} \lambda_d P_d$ on $L_2(\boldsymbol{\mathcal{H}}^{tD})$: Given two functions $f, g \in L_2(\mathcal{G}^t)$ with $\|f\|_\infty, \|g\|_\infty \leqslant 1$ and $\max_i \text{Inf}_i(U_\rho f), \text{Inf}_j(U_\rho g) \leqslant \tau$, there exists functions $f', g' \in L_2(\boldsymbol{\mathcal{H}}^{tD})$ satisfying $\|f'\|_\infty, \|g'\|_\infty \leqslant 1$, $\max_i \text{Inf}_i(T_\rho f'), \max_j \text{Inf}_j(T_\rho g') \leqslant \tau$, and*

$$\langle T_\rho f', B T_\rho g'\rangle \geqslant \langle f, \mathcal{G}(B) g\rangle - \eta\|B\|\,.$$

In particular, the choices $D \geqslant 2\log_{1-\eta}\eta/16$ and $\tau \leqslant O(2^{-35D^2\log D})$ suffice.

The invariance principle of [20] only applies to multilinear polynomials, while the functions f', g' need not be multilinear. To overcome this hurdle, we treat a multivariate Hermite expansion as a multilinear polynomial over the ensemble consisting of Hermite polynomials. Unfortunately, this step of the proof is complicated with careful truncation arguments and choice of ensembles to apply invariance principle.

In conclusion, by applying Lemma 5.3, we obtain functions f'' and g'' in $L_2(\boldsymbol{\mathcal{H}}^{t'D})$ that have the following properties:

1. $$\|f''\|_\infty, \|g''\|_\infty \leqslant 1$$
2. $$\max_i \text{Inf}_i(T_\rho f''), \max_j \text{Inf}_j(T_\rho g'') \leqslant \tau.$$

Further the functions f'', g'' satisfy,

$$\langle T_\rho f'', B T_\rho g''\rangle \geqslant \langle f', \mathcal{G}(B) g'\rangle - \eta\|B\| = \langle f, \mathcal{G}(B) g\rangle - \eta\|B\|$$
$$= \mathsf{Soundness}_{\eta,\tau}(B)(1+\eta) + \eta\mathsf{Completeness}(B) - \eta\|B\|\,.$$

Recall that $\|B\| = \lambda_1 = \mathsf{Completeness}(B)$. By the choice of $k > t'D$, the functions $f'', g'' \in L_2(\boldsymbol{\mathcal{H}}^{t'D}) \subset L_2(\boldsymbol{\mathcal{H}}^k)$. Thus we have two functions f'', g'' with no influential variables, but yielding a value higher than the $\mathsf{Soundness}_{\eta,\tau}(B)$. A contradiction.

Acknowledgments We thank Ryan O'Donnell for bringing [14] to our attention.

References

[1] N. Alon, K. Makarychev, Y. Makarychev, and A. Naor. Quadratic forms on graphs. In *STOC*, pages 486–493, 2005.

[2] N. Alon and A. Naor. Approximating the cut-norm via Grothendieck's inequality. *SIAM J. Comput*, 35(4):787–803, 2006.

[3] S. Arora, E. Berger, E. Hazan, G. Kindler, and M. Safra. On non-approximability for quadratic programs. In *FOCS*, pages 206–215, 2005.

[4] A. Coja-Oghlan, C. Cooper, A. Frieze. An efficiently regular concept for sparse graphs and matrices. To appear in *SODA*, 2009.

[5] B.W. Johnson and J. Lindenstrauss. Basic concepts in the geometry of Banach spaces. *Handbook of the geometry of Banach spaces, North Holland, Amsterdam*, 1:1–84, 2001.

[6] M. Charikar and A. Wirth. Maximizing quadratic programs: Extending Grothendieck's inequality. In *FOCS*, pages 54–60, 2004.

[7] A. Frieze and R. Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19(2):175–220, 1999.

[8] A. Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques. *Bol. Soc. Mat*, 8:1–79, 1953.

[9] S. Khot. On the power of unique 2-prover 1-round games. In *STOC*, pages 767–775, 2002.

[10] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comput*, 37(1):319–357, 2007.

[11] S. Khot and R. O'Donnell. SDP gaps and UGC-hardness for MaxCutGain. In *FOCS*, pages 217–226, 2006.

[12] S. Khot and N. K. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1. In *FOCS*, pages 53–62, 2005.

[13] G. Kindler, A. Naor, and G. Schechtman. The UGC hardness threshold of the ℓ_p Grothendieck problem. In *SODA*, pages 64–73, 2008.

[14] H. König. Spherical functions and the Grothendieck inequality. *Unpublished manuscript,* 2005. Available at http://shrimp.bayou.uni-linz.ac.at/Vienna2005/pdf/koenig.pdf.

[15] J. L. Krivine. Constantes de Grothendieck et fonctions de type positif sur les sphères. *Adv. Math.*, 31:16–30, 1977-1978.

[16] J. Lindenstrauss and A. Pelczyński. Absolutely summing operators in L_p spaces and their applications. *Studia Math*, 29:275–326, 1968.

[17] N. Linial and A. Shraibman. Lower bounds in communication complexity based on factorization norms. In *STOC*, pages 699–708, 2007.

[18] A. Megretski. Relaxation of quadratic programming in operator theory and systems analysis. *Systems, approximation, singular integral operators and related topics*, 2001.

[19] E. Mossel. Gaussian bounds for noise correlation of functions. Submitted. Extended abstract appeared in *FOCS*, 2008.

[20] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. To appear in Ann. Math., 2008.

[21] A. Nemirovski, C. Roos, and T. Terlaky. On maximization of quadratic form over intersection of ellipsoids with common center. *Mathematical Programming*, 86(3):463–473, 1999.

[22] P. Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *STOC*, pages 245–254, 2008.

[23] J. Reeds. A new lower bound on the real Grothendieck constant. *Unpublished Manuscript*, 1993. Available at http://www.dtc.umn.edu/~reedsj.