Towards Computing the Grothendieck constant

PRASAD RAGHAVENDRA[*] DAVID STEURER[†]

July 3, 2008

Abstract

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

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

where $B^{(n)}$ is the unit ball in $\mathbb{R}^n$. Despite several efforts [Kri78, Ree93], 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 $K_{N,N}$-QUADRATICPROGRAMMING 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 [Rag08] to the $K_{N,N}$-QUADRATICPROGRAMMING problem. Using some standard but non-trivial modifications, the reduction in [Rag08] yields the following hardness result: Assuming the Unique Games Conjecture, it is NP-hard to approximate the $K_{N,N}$-QUADRATICPROGRAMMING problem to any factor better than the Grothendieck constant K_G.

By adapting a "bootstrapping" argument used in a proof of Grothendieck inequality [BL01], we perform a tighter analysis than [Rag08]. Through this careful analysis, we obtain the following new results:

- An approximation algorithm for $K_{N,N}$-QUADRATICPROGRAMMING 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. To achieve this, we give a direct conversion from dictatorship tests to integrality gap instances bypassing the use of a reduction from UNIQUEGAMES and the Khot-Vishnoi [KV05] integrality gap instances for UNIQUEGAMES.

[*]University of Washington, Seattle, WA.
[†]Princeton University, Princeton, NJ.

1 Introduction

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

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

where K_G is a universal constant. The smallest value of K_G for which the inequality holds, is referred to as the Grothendieck constant. Since the inequality was first discovered [Gro53], the inequality has not only undergone various restatements under different frameworks of analysis (see [LP68]), it has also found numerous applications in functional analysis.

In recent years, the Grothendieck's inequality has found algorithmic applications in efficient construction of Szemeredi partitions of graphs and estimation of cut norms of matrices [AN06], in turn leading to efficient approximation algorithms for dense graph problems [FK99]. The inequality has also proved useful in certain lower bound techniques for communication complexity [LS07]. Among its various applications, here we shall elaborate on the $K_{N,N}$-QuadraticProgramming problem. In this problem, the objective is to maximize the following quadratic program given as input the matrix $A = (a_{ij})$.

$$\text{Maximize} \qquad \sum_{i,j} a_{ij} x_i y_j \qquad \text{Subject to: } x_i, y_j \in \{1, -1\}$$

Alternatively, the problem amounts to computing the norm $\|A\|_{\infty \to 1}$ of the matrix A. The $K_{N,N}$-QuadraticProgramming problem is a formulation of the correlation clustering problem with two clusters . The following natural SDP relaxation to the problem is obtained by relaxing the variables x_i, y_j to unit vectors.

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

The Grothendieck constant K_G is precisely the integrality gap of this SDP relaxation for the $K_{N,N}$-QuadraticProgramming 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 [Kri78], while the best known lower bound is roughly 1.67 [Ree93]. More importantly, very little seems to be known about the matrices A for which the inequality is tight. Approximating the Grothendieck constant 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) [Kho02].

In a recent work [Rag08], 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 $K_{N,N}$-QuadraticProgramming. In this light, the current work applies the techniques of Raghavendra [Rag08] to the $K_{N,N}$-QuadraticProgramming.

1.1 Results

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

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

Although $K_{N,N}$-QuadraticProgramming falls in the "generalized constraint satisfaction problem" framework of Raghavendra [Rag08], the above result does not immediately follow from [Rag08] 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 [Mos08].

On the other hand, the optimal algorithm for CSPs in [Rag08] does not directly translate in to an algorithm for $K_{N,N}$-QuadraticProgramming. The main issue is the constant additive error incurred in all the reductions of [Rag08]. For a CSP, the objective function is guaranteed to be at least a fixed constant fraction (say 0.5), and hence the additive constant error is negligible. In case of $K_{N,N}$-QuadraticProgramming, the value of the optimum solution could be $1/\log n$, in which case an additive constant error destroys the approximation ratio.

To obtain better bound on the error, we use a bootstrapping argument similar to the Gaussian Hilbert space approach to Grothendieck inequality [BL01] (this approach is used for algorithmic purposes in [AN06, AMMN05, KNS08]). Using ideas from the proof of Grothendieck inequality, we perform a tighter analysis of the reduction in [Rag08] for the special case of $K_{N,N}$-QuadraticProgramming. 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* $K_{N,N}$-QuadraticProgramming *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 tighter running time analysis could improve the $O(1/\eta^3)$, 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 [Rag08], our reduction does not use the Khot-Vishnoi [KV05] 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 $Q^{(k)}$ be the set of linear operators A on functions $f : \mathbb{R}^k \to \mathbb{R}$ of the form $A = \sum_{d \in \mathbb{N}} \lambda_d Q_d$, where Q_d is the orthogonal projector on the span of k-multivariate Hermite polynomials of degree d. Then,*

$$K_G = \sup_{\substack{d,k \in \mathbb{N},\\ A \in Q^{(k)}}} \frac{\sup_{f : \mathbb{R}^k \to [-1,1]} \int |Af(\boldsymbol{x})| \, \mathrm{d}\gamma(\boldsymbol{x})}{\sup_{f : \mathbb{R}^k \to B^{(d)}} \int \|Af(\boldsymbol{x})\| \, \mathrm{d}\gamma(\boldsymbol{x})}.$$

Here γ denotes the k-dimensional Gaussian probability measure, and for a function $f : \mathbb{R}^k \to \mathbb{R}^d$, 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.

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. $K_{N,N}$ Quadratic programming is the special case where the graph G is the complete bipartite graph.

The Grothendieck problem on a complete graph admits a $O(\log n)$ approximation [NRT99, Meg01, CW04] and has applications in correlation clustering [CW04]. For the Grothendieck problem on general graphs, [AMMN05] obtain an approximation that depends on the Lovasz θ 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, [KNS08] obtain UGC hardness results and approximation algorithms for the L_p-Grothendieck problem.

On the hardness side, [ABH+05] show a $O(\log^c n)$-NP hardness for the Grothendieck problem on the complete graph for some fixed constant $c < 1$. Integrality gaps for the Grothendieck problem on complete graphs were exhibited in [KO06, AMMN05]. For the $K_{N,N}$-Quadratic Programming problem, a UGC-based hardness of roughly 1.67 was shown in [KO06]. The reduction uses the explicit operator constructed in the proof of 1.67 lower bound [Ree93] for the Grothendieck constant.

2 Preliminaries

Problem 2.1 ($K_{N,N}$-QuadraticProgramming)**.** 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 always attained for numbers with $|x_i| = |y_j| = 1$.

Problem 2.2 ($K_{N,N}$-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 $\|\boldsymbol{u}_i\| = \|\boldsymbol{v}_j\| = 1$.

Definition 2.3. The Grothendieck constant K_G is the supremum of $\mathrm{sdp}(A)/\mathrm{opt}(A)$ over all matrices A.

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. Here, we will consider two kinds of probability spaces. One is the uniform distribution over the Hamming cube $\{1, -1\}^k$, denoted $\Omega = \mathcal{H}^k$. The other one is the Gaussian distribution over $\mathbb{R}^k$, denoted $\Omega = \mathcal{G}^k$. 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 2.4. *Given an operator A on $L_2(\Omega^k)$, and functions $f, g, f', g' \in L_2^{(d)}(\Omega^k)$ satisfying $\|f\|, \|g\|, \|f'\|, \|g'\| \leqslant 1$, then*

$$|\langle f, Ag \rangle - \langle f', Ag' \rangle| \leqslant \|A\| (\|f - f'\| + \|g - g'\|) \,.$$

Lemma 2.5 (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*

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

Definition 2.6 (Noise Operator)**.** For $\Omega = \mathcal{H}^k$ or $\Omega = \mathcal{G}^k$, 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 $\{\boldsymbol{x}_S := \prod_{i\in S} x_i \mid S \subseteq [k], |S| = d\}$.

Hermite Polynomials and Gaussian Noise operator. Let $\mathcal{G}$ be the probability space over $\mathbb{R}$ with Gaussian probability measure. The Hermite polynomials $\{H_d(x) | d \in \mathbb{N}\}$ form an orthonormal basis for $L_2(\mathcal{G})$. An orthogonal basis for $L_2(\mathcal{G}^t)$ is given by the set of functions $\{H_\sigma(\boldsymbol{x}) := \prod_{i=1}^{t} H_{\sigma_i}(x_i) \mid \sigma \in \mathbb{N}_0^t\}$.

Let Q_d denote the projection operator that maps an arbitrary function $F \in L_2(\mathcal{G}^t)$ to the degree d component.

Definition 2.7 (Gaussian Noise Operator)**.** Let U_ρ denote the linear operator on $L_2(\mathcal{G}^k)$ defined as

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

where Q_d denotes the orthogonal projector on the subspace of $L_2(\mathcal{H}^k)$ spanned by the degree-d Hermite polynomials $\left\{H_\sigma(\boldsymbol{x}) := \prod_{i\in[k]} H_{\sigma_i}(x_i) \mid \sigma \in \mathbb{N}_0^k, \sum \sigma_i = d\right\}$.

Variable Influences. For a function $f \in L_2(\mathcal{H}^k)$, we define $\mathrm{Inf}_i f = \sum_{\sigma \ni i} \hat{f}_\sigma^2$, where $\hat{f}$ is the Fourier-transform of f. Let us denote $\mathrm{MaxInf}\, f := \max_{i\in[k]} \mathrm{Inf}_i f$. For a pair of functions $f, g \in L_2(\mathcal{H}^k)$, 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*.

Fact 2.8. *For $f \in L_2(\mathcal{H}^k)$ and $\gamma \in [0, 1]$, we have $\sum_{i=1}^{k} \mathrm{Inf}_i T_{1-\gamma} f \leqslant \|f\|^2/\gamma$. Similarly, for $f \in L_2(\mathcal{G}^k)$ and $\gamma \in [0, 1]$, $\sum_{i=1}^{k} \mathrm{Inf}_i U_{1-\gamma} f \leqslant \|f\|^2/\gamma$.*

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

Lemma 2.9. *Let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^d$ be two unit vectors, and $f, g \in L_2(\mathcal{H}^k)$. Then,*

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

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. It suffices to show the lemma for the case that both f and g are the same multilinear monomial. Since the variables are independent, one may assume that the monomial has degree 1. For this case, it is trivial. □

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

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

Theorem 2.10 (Invariance Principle, [MOO08])**.** *There is a universal constant C such that, for all $\rho = 1-\gamma \in (0,1)$ the following holds: Let $f \in L_2(\mathcal{H}^k)$ with $\|f\|_\infty \leqslant 1$ and $\mathrm{Inf}_i(T_\rho f) \leqslant \tau$ for all $i \in [k]$. Then,*

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

where $\bar{f} \in L_2(\mathcal{G}^k)$ 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 [Rag08]. We begin by defining dictatorship tests in the current context.

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

$$B = \sum_{d=0}^{k} \lambda_d P_d$$

where P_d is the projection operator on to the degree d part, and $|\lambda_1| \geqslant |\lambda_d|$ for all d. We define two parameter of B:

$$\mathsf{Completeness}(B) := \inf_i \langle \chi_i, B\chi_i \rangle\,, \text{ where } \chi_i(\boldsymbol{x}) = x_i \text{ is the } i^{th} \text{ dictator function.}$$

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

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 $L_2(\mathcal{H}^k)$ 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 $L_2(\mathcal{H}^k)$ as follows:

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

where $\rho = 1 - \eta$.

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

Lemma 3.3. *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 $f, g \in L_2(\mathcal{H}^k)$ and $\rho < 1$. 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 3.4. $\mathsf{Round}_{\eta,f,g}(A)$ 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 the invariance principle [MOO08].

Theorem 3.5. *Let A be a matrix. For functions $f, g \in L_2(\mathcal{H}^k)$ 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 $f', g' \in L_2(\mathcal{H}^k)$ 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 $\min \mathrm{Inf}_i T_\rho f', \mathrm{Inf}_i T_\rho g' \leqslant \tau$ *for all i.*

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

Corollary 3.6. *For every matrix A and $\eta > 0$,*

$$\mathsf{Soundness}_{\eta,\tau}(\mathcal{D}_\eta(A)) \leqslant \left(\sup_{\substack{f,g \in L_2(\mathcal{H}^k),\\ \mathrm{MaxComInf}(T_\rho f, T_\rho g) \leqslant \tau}} \mathsf{Round}_{\eta,f,g}(A) \right) + \left(10\tau^{C\eta/8}/\sqrt{\eta}\right) \cdot \mathrm{sdp}(A)\,,$$

where $\rho = 1-\eta$.

As $\mathsf{Round}_{\eta,f,g}$ is the expected value of a $\{1,-1\}$ solution, it is necessarily smaller than $\mathrm{opt}(A)$. 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 3.7. *For a fixed $\eta > 0$, if $\tau \leqslant 2^{-100 \log \eta / C\eta}$ then , 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 [Rag08], 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 3.8. Given an dictatorship test B on $L_2(\mathcal{H}^k)$ of the form $B = \sum_d \lambda_{d=0}^k P_d$, define the corresponding operator $\mathcal{G}_\eta(B)$ on $L_2(\mathcal{G}^k)$ 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 Appendix A.

Theorem 3.9. *For all $\eta > 0$, there exists k, τ such that following holds: For any dictatorship test B on $L_2(\mathcal{H}^k)$, we have:*

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

$$\mathrm{opt}(\mathcal{G}(B)) \leqslant \mathsf{Soundness}_{\eta,\tau}(B)\,(1 + \eta) + \eta\mathsf{Completeness}(B) \tag{2}$$

In particular, the choices $\tau = O(2^{-100/\eta^3})$ and $k = \Omega(2^{200/\eta^3})$ 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 3.10. *For all $\eta > 0$, there exists k, τ such that: For any dictatorship test B on $L_2(\mathcal{H}^k)$,*

$$\frac{\mathsf{Soundness}_{\eta,\tau}(B)}{\mathsf{Completeness}(B)} \geqslant \frac{1}{K_G} - \eta\,. \tag{3}$$

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

Corollary 3.11. *The equation 2 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 Appendix B.

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 function $f, g \in L_2(\mathcal{H}^k)$ 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 $\sup_{f,g \in L_2(\mathcal{H}^k)} \mathsf{Round}_{\eta,f,g}(A)$. From Corollary 3.6 we have

$$\begin{aligned}\sup_{f,g \in L_2(\mathcal{H}^k), \|f\|,\|g\| \leqslant 1} \mathsf{Round}_{\eta,f,g}(A) &\geqslant \sup_{\substack{f,g \in L_2(\mathcal{H}^k),\\ \mathrm{MaxComInf}(T_\rho f, T_\rho g) \leqslant \tau}} \mathsf{Round}_{\eta,f,g}(A) \\ &\geqslant \mathsf{Soundness}_{\eta,\tau}(\mathcal{D}_\eta(A)) - \left(10\tau^{C\eta/8}/\sqrt{\eta}\right) \cdot \mathrm{sdp}(A) \end{aligned} \tag{4}$$

From Lemma 3.3, we know $\mathsf{Completeness}(\mathcal{D}_\eta(A)) = \mathrm{sdp}(A)(1 - \eta)$. By the choice of k, τ, we can apply Corollary 3.10 on $\mathcal{D}_\eta(A)$ to conclude

$$\mathsf{Soundness}_{\eta,\tau}(\mathcal{D}_\eta(A)) \geqslant \mathsf{Completeness}(\mathcal{D}_\eta(A))\Big(\frac{1}{K_G} - \eta\Big) \geqslant \mathrm{sdp}(A)\Big(\frac{1}{K_G} - \eta\Big)(1 - \eta) \tag{5}$$

From Equations 4 and 5, we conclude that the value returned by the algorithm is at least

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

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 $L_2(\mathcal{H}^k)$ using a ϵ-net in the L_2 norm. As k is a fixed constant depending on η, there is a finite ϵ-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 following lemma (see Appendix C for the proof):

Lemma 3.12. *For $f, g, f', g' \in L_2(\mathcal{H}^k)$ with $\|f\|, \|g\|, \|f'\|, \|g'\| \leqslant 1$,*

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

4 From Integrality gaps to Dictatorship Tests

4.1 Rounding Scheme

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

$\mathsf{Round}_{f,g}$

Input : An $m \times n$ matrix $A = (a_{ij})$ with SDP solution $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\}, \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} \subset B^{(d)}$

- 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 $1/\sqrt{d}$.
- Assign $x_i = \mathrm{trunc}\, T_\rho \bar{f}(\Phi \boldsymbol{u}_i)$ and $y_j = \mathrm{trunc}\, T_\rho \bar{g}(\Phi \boldsymbol{v}_j)$.

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

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

4.2 Relaxed Influence Condition

The following lemma shows that we could replace the condition $\mathrm{MaxComInf}(T_\rho f, T_\rho g) \leqslant \tau$ in Definition 3.1 by the condition $\mathrm{MaxInf}\, T_\rho f, \mathrm{MaxInf}\, T_\rho g \leqslant \sqrt{\tau}$ with a small loss in the soundness. The proof is presented in Appendix C.

Lemma 4.1. *Let A be a dictatorship test on $L_2(\mathcal{H}^k)$, and let f, g be a pair of functions in $L_2(\mathcal{H}^k)$ with $\|f\|_\infty, \|g\|_\infty \leqslant 1$ and $\mathrm{MaxComInf}(T_\rho f, T_\rho g) \leqslant \tau$ for $\rho = 1 - \eta$. Then for every $\tau' > 0$, there are functions $f', g' \in L_2(\mathcal{H}^k)$ with $\|f'\|_\infty, \|g'\|_\infty \leqslant 1$ and $\mathrm{MaxInf}\, T_\rho f', \mathrm{MaxInf}\, T_\rho g' \leqslant \tau'$ 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 3.5

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

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

On the other hand, we have

$$\langle f', \mathcal{D}_\eta(A) g'\rangle = \sum_{ij} \sum_{d=0}^{k} \langle T_\rho f', a_{ij} \langle \boldsymbol{u}_i, \boldsymbol{v}_j\rangle^d P_d(T_\rho g')\rangle = \sum_{ij} a_{ij} \langle T_\rho f', T_{\langle \boldsymbol{u}_i, \boldsymbol{v}_j\rangle}(T_\rho g')\rangle \tag{6}$$

We can assume that all vectors $\boldsymbol{u}_i$ and $\boldsymbol{v}_j$ have unit norm. By Lemma 2.9 , we have

$$\mathop{\mathbb{E}}_{\Phi} \sum_{ij} a_{ij} T_\rho \bar{f}'(\Phi \boldsymbol{u}_i) T_\rho \bar{g}'(\Phi \boldsymbol{v}_j) = \sum_{ij} a_{ij} \langle T_\rho f', T_{\langle \boldsymbol{u}_i, \boldsymbol{v}_j\rangle}(T_\rho g')\rangle \tag{7}$$

From the above equations we have

$$\langle f, \mathcal{D}_\eta(A) g\rangle = \mathop{\mathbb{E}}_{\Phi} \sum_{ij} a_{ij} T_\rho \bar{f}'(\Phi \boldsymbol{u}_i) T_\rho \bar{g}'(\Phi \boldsymbol{v}_j) \tag{8}$$

By the invariance principle (Theorem 2.10), we have

$$\|T_\rho \bar{f}' - \operatorname{trunc} T_\rho \bar{f}'\| \leqslant \tau^{C\eta/2} \qquad \text{and } \|T_\rho \bar{g}' - \operatorname{trunc} T_\rho \bar{g}'\| \leqslant \tau^{C\eta/2}\,. \tag{9}$$

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

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

and

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

Equation (9) 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 2.5), we finish the proof

$$\begin{aligned}
\mathsf{Round}_{\eta, f', g'}(A) = \sum_{ij} a_{ij} \langle \boldsymbol{u}_i'', \boldsymbol{v}_j''\rangle &= \sum_{ij} a_{ij} \langle \boldsymbol{u}_i', \boldsymbol{v}_j'\rangle - \sum_{ij} a_{ij} \langle \boldsymbol{u}_i' - \boldsymbol{u}_i'', \boldsymbol{v}_j'\rangle - \sum_{ij} a_{ij} \langle \boldsymbol{u}_i'', \boldsymbol{v}_j' - \boldsymbol{v}_j''\rangle \\
&\overset{(9)}{\geqslant} \sum_{ij} a_{ij} \langle \boldsymbol{u}_i', \boldsymbol{v}_j'\rangle - 2\tau^{C\gamma} \operatorname{opt}(A) - 2\tau^{C\gamma} \operatorname{opt}(A) \\
&\geqslant \langle f, \mathcal{D}_\eta A g\rangle - 4\tau^{C\eta/2} \operatorname{opt}(A) - 4\tau^{1/4} \operatorname{opt}(A)/\sqrt{\eta}\,.
\end{aligned} \tag{10}$$

□

5 From Dictatorship Tests to Integrality Gaps

In this section, we outline the key ideas in the proof of 3.9. For the convenience of the reader, a full self contained proof is presented in Appendix A.

5.1 $\operatorname{sdp}(\mathcal{G}(B)) \geqslant \mathsf{Completeness}(B)\,(1 - 5\eta)$

To prove this claim, we need to construct an SDP solution to $\operatorname{sdp}(\mathcal{G}(B))$ that achieves nearly the same value as $\mathsf{Completeness}(B)$. Formally, we need to construct functions f, g whose domain is $\mathcal{G}^t$ 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 [KO06, Ree93], we choose the following function $f(\boldsymbol{x}) = \boldsymbol{x}/\|\boldsymbol{x}\|$ which always outputs unit vectors, and very close to the linear function $f(\boldsymbol{x}) = \boldsymbol{x}/\sqrt{t}$ as t increases. From Lemma A.1, for $t > 1/\eta^5$, we have

$$\operatorname{sdp}(\mathcal{G}(B)) \geqslant \mathsf{Completeness}(B)(1 - 5\eta)$$

5.2 $\mathrm{opt}(\mathcal{G}(B)) \leqslant \mathsf{Soundness}_{\eta,\tau}(B)\,(1+\eta) + \eta\mathsf{Completeness}(B)$

For the sake of contradiction, let us suppose $\mathrm{opt}(\mathcal{G}(B)) \geqslant \mathsf{Soundness}_{\eta,\tau}(B)(1+\eta) + \eta\mathsf{Completeness}(B)$. Let the optimum solution be given by two functions $f, g \in L_2(\mathcal{G}^t)$. 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 $L_2(\mathcal{H}^k)$ 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.

$$\begin{aligned}
f'(\boldsymbol{x}) &= f\Big(\frac{1}{\sqrt{R}}\sum_{i=1}^{R} x_i, \frac{1}{\sqrt{R}}\sum_{i=R+1}^{2R} x_i, \ldots, \frac{1}{\sqrt{R}}\sum_{i=(R-1)t+1}^{Rt} x_i\Big) \\
g'(\boldsymbol{x}) &= g\Big(\frac{1}{\sqrt{R}}\sum_{i=1}^{R} x_i, \frac{1}{\sqrt{R}}\sum_{i=R+1}^{2R} x_i, \ldots, \frac{1}{\sqrt{R}}\sum_{i=(R-1)t+1}^{Rt} x_i\Big)
\end{aligned}$$

In Lemma A.2, we show that for $R = \lceil 1/\eta\tau \rceil$, the functions $f', g' \in L_2(\mathcal{G}^{t'})$ for $t' = t \cdot R$ satisfy the requisite properties.

– In the second step, we apply the invariance principle to construct functions on $L_2(\mathcal{H}^k)$ with the same properties as f', g'. However, the invariance principle of [MOO08] 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. The technical details are described in Lemma A.4. In conclusion, by applying Lemma A.4, we obtain functions f'' and g'' in $L_2(\mathcal{H}^{t'D})$ that have the following properties :

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

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

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

Recall that $\|B\| = \lambda_1 = \mathsf{Completeness}(B)$. By the choice of $k > t'D$, the functions $f'', g'' \in L_2(\mathcal{H}^{t'D}) \subset L_2(\mathcal{H}^k)$. 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.

References

[ABH+05] Sanjeev Arora, Eli Berger, Elad Hazan, Guy Kindler, and Muli Safra, *On non-approximability for quadratic programs*, FOCS, 2005, pp. 206–215.

[AMMN05] Noga Alon, Konstantin Makarychev, Yury Makarychev, and Assaf Naor, *Quadratic forms on graphs*, STOC, 2005, pp. 486–493.

[AN06] Noga Alon and Assaf Naor, *Approximating the cut-norm via Grothendieck's inequality*, SIAM J. Comput **35** (2006), no. 4, 787–803.

[BL01] B.W.Johnson and Johnson Lindenstrauss, *Basic concepts in the geometry of Banach spaces*, Handbook of the geometry of Banach spaces, North Holland, Amsterdam **1** (2001), 1–84.

[CW04] Moses Charikar and Anthony Wirth, *Maximizing quadratic programs: Extending Grothendieck's inequality*, FOCS, 2004, pp. 54–60.

[FK99] Alan Frieze and Ravi Kannan, *Quick approximation to matrices and applications*, Combinatorica **19** (1999), no. 2, 175–220.

[Gro53] Alexander Grothendieck, *Résumé de la théorie métrique des produits tensoriels topologiques*, Bol. Soc. Mat **8** (1953), 1–79.

[Kho02] Subhash Khot, *On the power of unique 2-prover 1-round games*, STOC, 2002, pp. 767–775.

[KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell, *Optimal inapproximability results for MAX-CUT and other 2-variable CSPs?*, SIAM J. Comput **37** (2007), no. 1, 319–357.

[KNS08] Guy Kindler, Assaf Naor, and Gideon Schechtman, *The UGC hardness threshold of the ℓ_p Grothendieck problem*, SODA, 2008, pp. 64–73.

[KO06] Subhash Khot and Ryan O'Donnell, *SDP gaps and UGC-hardness for MAXCUTGAIN*, FOCS, 2006, pp. 217–226.

[Kri78] J. L. Krivine, *Constantes de Grothendieck et fonctions de type positif sur les sphères*, Adv. Math. **31** (1977-1978), 16–30.

[KV05] Subhash Khot and Nisheeth K. Vishnoi, *The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1*, FOCS, 2005, pp. 53–62.

[LP68] Johnson Lindenstrauss and A. Pelczyński, *Absolutely summing operators in L_p spaces and their applications*, Studia Math **29** (1968), 275–326.

[LS07] Nati Linial and Adi Shraibman, *Lower bounds in communication complexity based on factorization norms*, STOC, 2007, pp. 699–708.

[Meg01] A. Megretski, *Relaxation of quadratic programming in operator theory and systems analysis*, Systems, approximation, singular integral operators and related topics (2001).

[MOO08] E. Mossel, R. O'Donnell, and K. Oleszkiewicz, *Noise stability of functions with low influences: invariance and optimality*, To appear in Ann. Math., 2008.

[Mos08] E. Mossel, *Gaussian bounds for noise correlation of functions*, Submitted. Extended Abstract to Appear in FOCS, 2008.

[NRT99] A. Nemirovski, C. Roos, and T. Terlaky, *On maximization of quadratic form over intersection of ellipsoids with common center*, Mathematical Programming **86(3)** (1999), 463–473.

[Rag08] Prasad Raghavendra, *Optimal algorithms and inapproximability results for every CSP?*, STOC, 2008, pp. 245–254.

[Ree93] J. Reeds, *A new lower bound on the real Grothendieck constant*, Unpublished Manuscript, available at http://www.dtc.umn.edu/ reedsj (1993).

A From Dictatorship Tests to Integrality Gaps

In this section, we give a self contained proof of Theorem 3.9 for the convenience of the reader.

Proof of Theorem 3.9 :

Proof. Define $D = \lceil 2\log_{1-\eta}\eta/16\rceil$, $\tau \leqslant O(2^{-35D^2\log D})$, $t > 1/\eta^5$ and $k > tD\lceil 1/\eta\tau\rceil$.

Firstly, we need to construct an SDP solution to $\mathrm{sdp}(\mathcal{G}(B))$ that achieves nearly the same value as $\mathsf{Completeness}(B)$. Formally, we need to construct functions f, g whose domain is $\mathcal{G}^t$ 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 [KO06, Ree93], we choose the following function $f(\boldsymbol{x}) = \boldsymbol{x}/\|\boldsymbol{x}\|$ which always outputs unit vectors, and very close to the linear function $f(\boldsymbol{x}) = \boldsymbol{x}/\sqrt{t}$ as t increases. From Lemma A.1, for $t > 1/\eta^5$, we have

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

For the sake of contradiction, let us suppose $\mathrm{opt}(\mathcal{G}(B)) \geqslant \mathsf{Soundness}_{\eta,\tau}(B)(1+\eta)+\eta\mathsf{Completeness}(B)$. Let the optimum solution be given by two functions $f, g \in L_2(\mathcal{G}^t)$. 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 $L_2(\mathcal{H}^k)$ 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.

$$\begin{aligned} f'(\boldsymbol{x}) &= f\Big(\frac{1}{\sqrt{R}}\sum_{i=1}^{R} x_i, \frac{1}{\sqrt{R}}\sum_{i=R+1}^{2R} x_i, \ldots, \frac{1}{\sqrt{R}}\sum_{i=(R-1)t+1}^{Rt} x_i\Big) \\ g'(\boldsymbol{x}) &= g\Big(\frac{1}{\sqrt{R}}\sum_{i=1}^{R} x_i, \frac{1}{\sqrt{R}}\sum_{i=R+1}^{2R} x_i, \ldots, \frac{1}{\sqrt{R}}\sum_{i=(R-1)t+1}^{Rt} x_i\Big) \end{aligned}$$

  In Lemma A.2, we show that for $R = \lceil 1/\eta\tau\rceil$, the functions $f', g' \in L_2(\mathcal{G}^{t'})$ for $t' = t\cdot R$ satisfy the requisite properties.

– In the second step, we apply the invariance principle to construct functions on $L_2(\mathcal{H}^k)$ with the same properties as f', g'. However, the invariance principle of [MOO08] 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. The technical details are described in Lemma A.4. In conclusion, by applying Lemma A.4, we obtain functions f'' and g'' in $L_2(\mathcal{H}^{t'D})$ that have the following properties :

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

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

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

Recall that $\|B\| = \lambda_1 = \mathsf{Completeness}(B)$. By the choice of $k > t'D$, the functions $f'', g'' \in L_2(\mathcal{H}^{t'D}) \subset L_2(\mathcal{H}^k)$. 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.

It is easy to check that our choices of k, τ satisfy $\tau < O(2^{-100/\eta^3})$ and $k > \Omega(2^{200/\eta^3})$. □

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

Proof. For the sake of brevity, let us denote $C = \mathcal{G}(B)$. Consider the functions $F, G : \mathbb{R}^t \to B^{(t)}$ given by $F(\boldsymbol{x}) = \frac{\boldsymbol{x}}{\|x\|}$. The functions F, G form a feasible SDP solution, since they associate a unit vector $F(\boldsymbol{x})$ with each point $\boldsymbol{x}$ is $\mathbb{R}^t$. Define a related function $F'(\boldsymbol{x}) = \frac{\boldsymbol{x}}{\sqrt{t}}$ also satisfying $\|F\| = 1$. As $\|\boldsymbol{x}\|$ is concentrated around $\sqrt{t}$, the functions F, F' should be close to each other. Formally,

$$\begin{aligned}\|F - F'\|^2 &= \mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\left\|\frac{\boldsymbol{x}}{\sqrt{t}} - \frac{\boldsymbol{x}}{\|x\|}\right\|^2\right] = \frac{1}{t}\mathop{\mathbb{E}}_{\boldsymbol{x}}\left[(\|\boldsymbol{x}\| - \sqrt{t})^2\right] \\ &= 2 - \frac{2\,\mathbb{E}_{\boldsymbol{x}}[\|\boldsymbol{x}\|]}{\sqrt{t}} \qquad (\text{ using } \mathop{\mathbb{E}}_{\boldsymbol{x}}[\|\boldsymbol{x}\|^2] = t)\end{aligned}$$

Using the well known fact $\mathbb{E}_{\boldsymbol{x}}[\|\boldsymbol{x}\|] > \sqrt{t} - O(\sqrt{\log t})$, we get $\|F - F'\| \leqslant (\frac{\log t}{t})^{\frac{1}{4}}$. Observe that,

$$\langle F', CG'\rangle = \frac{1}{t}\langle \boldsymbol{x}, (\sum_{d=0}^{k}\rho^{4d}\lambda_d Q_d)\boldsymbol{x}\rangle = \rho^4\lambda_1$$

Now using Lemma 2.4

$$\langle F, CG\rangle \geqslant \lambda_1\rho^4 - \|C\|(\|F' - F\| + \|G - G'\|) \geqslant \lambda_1\left(\rho^4 - 2\left(\frac{\log t}{t}\right)^{\frac{1}{4}}\right)$$

This completes the proof since $\lambda_1 = \mathsf{Completeness}(B)$. □

Lemma A.2. *Given two functions $f, g \in L_2(\mathcal{G}^t)$ with $\|f\|_\infty, \|g\|_\infty \leqslant 1$, there exists $f', g' \in L_2(\mathcal{G}^{t\cdot\lceil 1/(1-\rho)\tau\rceil})$ with $\|f\|'_\infty, \|g\|'_\infty \leqslant 1$ and $\max_i \mathrm{Inf}_i(U_\rho f'), \max_j \mathrm{Inf}_j(U_\rho g') \leqslant \tau$ and*

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

Proof. Again for the sake of brevity, let us denote $C = \mathcal{G}(B)$ and $R = \lceil 1/(1-\rho)\tau \rceil$. Define f', g' as follows :

$$\begin{aligned} f'(\boldsymbol{x}) &= f\Big(\frac{1}{\sqrt{R}}\sum_{i=1}^{R} x_i, \frac{1}{\sqrt{R}}\sum_{i=R+1}^{2R} x_i, \ldots, \frac{1}{\sqrt{R}}\sum_{i=(R-1)t+1}^{Rt} x_i\Big) \\ g'(\boldsymbol{x}) &= g\Big(\frac{1}{\sqrt{R}}\sum_{i=1}^{R} x_i, \frac{1}{\sqrt{R}}\sum_{i=R+1}^{2R} x_i, \ldots, \frac{1}{\sqrt{R}}\sum_{i=(R-1)t+1}^{Rt} x_i\Big) \end{aligned}$$

For succinctness, we will use $M : \mathbb{R}^{tR} \to \mathbb{R}^t$ to denote the linear transformation that maps $(x_1, \ldots, x_{tR}) \to \big(1/\sqrt{R}\sum_{i=1}^{R} x_i, 1/\sqrt{R}\sum_{i=R+1}^{2R} x_i, \ldots, 1/\sqrt{R}\sum_{i=(R-1)t+1}^{Rt} x_i\big)$. In this notation, $f'(\boldsymbol{x}) = f(M\boldsymbol{x})$ and $g'(\boldsymbol{x}) = g(M\boldsymbol{x})$.

Firstly, we will show that $\langle f', \mathcal{G}(B)g'\rangle = \langle f, \mathcal{G}(B)g\rangle$. As $\mathcal{G}(B)$ is a linear combination of the projection operators Q_d, it is enough to show that for all projections less than d, we have $\langle f', Q_d g'\rangle = \langle f, Q_d g\rangle$. Let us denote $f^{=d} = Q_d f, g^{=d} = Q_d g, f'^{=d} = Q_d f'$ and $g'^{=d} = Q_d g'$. We need to show that for each d

$$\langle f^{=d}, g^{=d}\rangle = \langle f'^{=d}, g'^{=d}\rangle$$

Towards this, we make the following claim:

Claim A.3. *For any β, $(U_\beta f)(M\boldsymbol{x}) = (U_\beta f')(\boldsymbol{x})$*

Proof.

$$(U_\beta f')(\boldsymbol{x}) = \mathop{\mathbb{E}}_{\boldsymbol{z}}[f'(\beta\boldsymbol{x} + \sqrt{(1-\beta^2)}\boldsymbol{z})] = \mathop{\mathbb{E}}_{\boldsymbol{z}}[f(\beta(M\boldsymbol{x}) + \sqrt{(1-\beta^2)}(M\boldsymbol{z}))]$$

For random Gaussian vector $\boldsymbol{z} \in \mathcal{G}^{tR}$, the distribution of $M\boldsymbol{z}$ is identical to a random Gaussian vector $\boldsymbol{z}' \in \mathcal{G}^t$. Thus we have

$$(U_\beta f')(\boldsymbol{x}) = \mathop{\mathbb{E}}_{\boldsymbol{z}'}[f(\beta(M\boldsymbol{x}) + \sqrt{(1-\beta^2)}\boldsymbol{z}')] = (U_\beta f)(M\boldsymbol{x})$$

□

From the above claim, we have:

$$\sum_d \beta^d f'^{=d}(\boldsymbol{x}) = (U_\beta f')(\boldsymbol{x}) = (U_\beta f)(M\boldsymbol{x}) = \sum_d \beta^d f^{=d}(M\boldsymbol{x})$$

for all $\beta \in [0,1]$ and all x. Any two power series that agree on an interval must have matching coeffcients. From this, we get $f'^{=d}(\boldsymbol{x}) = f^{=d}(M\boldsymbol{x})$ for all x. Now we have,

$$\langle f'^{=d}, g'^{=d}\rangle = \mathop{\mathbb{E}}_{\boldsymbol{x}}\left[f'^{=d}(\boldsymbol{x})g'^{=d}(\boldsymbol{x})\right] = \mathop{\mathbb{E}}_{\boldsymbol{x}}\left[f^{=d}(M\boldsymbol{x})g^{=d}(M\boldsymbol{x})\right] = \langle f^{=d}, g^{=d}\rangle$$

The last equality uses the fact that $M\boldsymbol{x}$ has the same distribution as a random vector $\boldsymbol{z}$ generated from the Gaussian measure $\mathcal{G}^t$.

To finish the proof of the lemma, we need to show that $\max_i(\mathrm{Inf}_i(U_\rho f'), \max_i(U_\rho g') \leqslant \tau$. From Fact 2.8, we know $\sum_i \mathrm{Inf}_i(U_\rho f') \leqslant \|f'\|^2/1-\rho$. Further, due to the symmetries among the input variables to f', we have $\mathrm{Inf}_{(i-1)R+1}(U_\rho f') = \mathrm{Inf}_{(i-1)R+2}(U_\rho f') \ldots = \ldots \mathrm{Inf}_{iR}(U_\rho f')$. Hence we immediately conclude $\max_i \mathrm{Inf}_i U_\rho f' \leqslant 1/(1-\rho)R \leqslant \tau$. The same argument applied to g' finishes the proof. □

Lemma A.4. *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 $L_2(\mathcal{H}^{tD})$: Given two functions $f, g : L_2(\mathcal{G}^t)$ with $\|f\|_\infty, \|g\|_\infty \leqslant 1$ and $\max_i \mathrm{Inf}_i(U_\rho f), \mathrm{Inf}_j(U_\rho g) \leqslant \tau$, there exists functions $f', g' \in L_2(\mathcal{H}^{tD})$ satisfying $\|f'\|_\infty, \|g'\|_\infty \leqslant 1$, $\max_i \mathrm{Inf}_i(T_\rho f'), \max_j \mathrm{Inf}_j(T_\rho g') \leqslant \tau$*

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

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

Proof. . For the sake of brevity, let us denote $C = \mathcal{G}(B)$. Notice that by definition of $\mathcal{G}(B)$, we have $C = U_{\rho^2} N U_{\rho^2}$. Define $\tilde{f} = U_\rho f$ and $\tilde{g} = U_\rho g$. By definition, the functions $\tilde{f}, \tilde{g}$ satisfy $\|\tilde{f}\|_\infty, \|\tilde{g}\|_\infty \leqslant 1$. Further,

$$\langle f, Cg \rangle = \langle f, U_\rho^2 B U_\rho^2 g \rangle = \langle U_\rho f, U_\rho B U_\rho g \rangle = \langle \tilde{f}, U_\rho B U_\rho \tilde{g} \rangle \tag{11}$$

Let $f^{\leqslant D}, g^{\leqslant D}$ denote the functions obtained by truncating $\tilde{f}, \tilde{g}$ to degree D. Formally, we have $f^{\leqslant D} = \sum_{d=0}^{D} \rho^d Q_d f$ and $g^{\leqslant D} = \sum_{d=0}^{D} \rho^d Q_d g$. Observe that

$$\|\tilde{f} - f^{\leqslant D}\|^2 = \sum_{|\sigma| > D} \tilde{f}_\sigma'^2 = \sum_{|\sigma| > D} \rho^{|\sigma|} f_\sigma^2 \leqslant \eta^2/256 \tag{12}$$

From Lemma 2.4 we have:

$$\begin{aligned} \langle f^{\leqslant D}, U_\rho B U_\rho g^{\leqslant D} \rangle &\geqslant \langle \tilde{f}, U_\rho B U_\rho \tilde{g} \rangle - \|U_\rho B U_\rho\| (\|\tilde{f} - f^{\leqslant D}\| + \|\tilde{g} - g^{\leqslant D}\|) \\ &\geqslant \langle f, Cg \rangle - \eta \|B\|/16 \end{aligned} \tag{13}$$

Although the range of functions $\tilde{f}, \tilde{g}$ is $[-1, 1]$, their low degree truncations $f^{\leqslant D}, g^{\leqslant D}$ are not necessarily bounded in $[-1, 1]$. However, they are nearly always bounded in the following sense:

$$\begin{aligned} \|f^{\leqslant D} - \mathrm{trunc}\, f^{\leqslant D}\| \leqslant& \quad \|f^{\leqslant D} - \tilde{f}\| + \|\tilde{f} - \mathrm{trunc}\, \tilde{f}\| + \|\mathrm{trunc}\, \tilde{f} - \mathrm{trunc}\, f^{\leqslant D}\| \\ \leqslant& \quad \eta/16 + 0 + \eta/16 = \eta/8 \end{aligned} \tag{14}$$

The final inequality is obtained using Equation (12), $\|\tilde{f} - \mathrm{trunc}\, \tilde{f}\|_\infty = 0$ and $\|\mathrm{trunc}\, \tilde{f} - \mathrm{trunc}\, f^{\leqslant D}\| \leqslant \|\tilde{f} - f^{\leqslant D}\|$. Similarly, for $g^{\leqslant D}$ we have $\|g^{\leqslant D} - \mathrm{trunc}\, g^{\leqslant D}\| \leqslant \eta/8$.

Now we shall apply the invariance principle to obtain functions f', g' on $\{1, -1\}^k$ from $f^{\leqslant D}, g^{\leqslant D}$. Towards this, we define certain ensembles of random variables. Firstly, let $\mathcal{X}$ denote the ensemble consisting of the first $D + 1$ Hermite polynomials. Specifically,

$$\mathcal{X} = \{H_0(x) = 1, H_1(x), \ldots, H_D(x) | x \in \mathcal{G}\}$$

Correspondingly, define the ensemble $\mathcal{Y}$ over the space $\{1, -1\}^D$ as follows:

$$\mathcal{Y} = \left\{ \chi_i(\boldsymbol{y}) = \prod_{j=1}^{i} y_i \middle| 0 \leqslant i \leqslant D, \boldsymbol{y} \text{ is uniform in } \{1, -1\}^D \right\}$$

We wish to point out that $\mathcal{Y}$ does not form a basis for $L_2(\{1, -1\}^D)$ since it consists of only $D + 1$ of the 2^D characters. Observe that $\mathcal{X}$ and $\mathcal{Y}$ are orthonormal ensemble of random variables. Let $\mathcal{X}^t = (\mathcal{X}^{(1)}, \ldots, \mathcal{X}^{(t)})$ and $\mathcal{Y}^t = (\mathcal{Y}^{(1)}, \ldots, \mathcal{Y}^{(t)})$ denote the product space consisting of t independent copies of $\mathcal{X}$ and $\mathcal{Y}$ respectively.

Recall that $f^{\leqslant D}$ and $g^{\leqslant D}$ are polynomials of total degree at most D in t variables. In other words, $f^{\leqslant D}$ can be written as :

$$f^{\leqslant D}(x^{(1)}, x^{(2)}, \ldots, x^{(t)}) = \sum_{|\sigma| \leqslant D} \hat{f}_\sigma \prod_{i=1}^{t} H_{\sigma_i}(x^{(i)})$$

Hence $f^{\leqslant D}, g^{\leqslant D}$ are multilinear polynomials in the ensemble $\mathcal{X}^t$. Let $\bar{f}$ and $\bar{g}$ denote the multilinear polynomials obtained by substituting $\mathcal{Y}^t$ in place of $\mathcal{X}^t$. So $\bar{f}$ and $\bar{g}$ are functions over the domain $(\{1,-1\}^D)^t = \{1,-1\}^{tD}$. More precisely,

$$\bar{f}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(t)}) = \sum_{|\sigma| \leqslant D} \hat{f}_\sigma \prod_{i=1}^{t} \chi_{\sigma_i}(\boldsymbol{y}^{(i)}) \qquad\qquad \text{where } \boldsymbol{y}^{(i)} \in \{1,-1\}^D$$

The ensemble $\mathcal{Y}$ is chosen so that term $\prod_{i=1}^t H_{\sigma_i}(x^i)$ is mapped to a unique term $\prod_{i=1}^t \chi_{\sigma_i}(\boldsymbol{y}^{(i)})$ of the same total degree. Hence we have the following identity :

$$\langle f^{\leqslant D}, U_\rho B U_\rho g^{\leqslant D} \rangle = \sum_{d=0}^{D} \rho^{2d} \lambda_d \sum_{|\sigma|=d} \hat{f}_\sigma \hat{g}_\sigma = \langle T_\rho \bar{f}, B T_\rho \bar{g} \rangle \tag{15}$$

Finally, define the functions f' and g' as

$$f' = \operatorname{trunc} \bar{f} \qquad\qquad g' = \operatorname{trunc} \bar{g}$$

By definition, the functions f', g' satisfy $\|f\|_\infty, \|g\|_\infty \leqslant 1$. Further, for any coordinate i, we have

$$\operatorname{Inf}_i(T_\rho f') \leqslant \operatorname{Inf}_i(T_\rho \bar{f}) \leqslant \operatorname{Inf}_i(\bar{f}) = \operatorname{Inf}_i(f^{\leqslant D}) \leqslant \operatorname{Inf}_i(\tilde{f}) = \operatorname{Inf}_i(U_\rho f) \leqslant \tau$$

Thus the functions f', g' satisfy the maximum influence condition too.

Firstly, we appeal to the invariance principle to bound the truncation errors : $\|f' - \bar{f}\|$ and $\|g' - \bar{g}\|$. Towards this, note that there exists some finite constant $\alpha(D)$ depending only on D such that ensembles $\mathcal{X}$ and $\mathcal{Y}$ are $(2, 3, \beta(D))$-hypercontractive(see [MOO08] for definitions). In particular, we will show the following in Appendix D

Lemma A.5. *For all D, the ensemble $\mathcal{X} = \{H_0(x), \ldots, H_D(x)\}$ is $(2, 3, \alpha(D) = 2^{-5D \log D})$ hypercontractive. The ensemble $\mathcal{Y} = \{\chi_i(y)\}$ is $(2, 3, 2^{-D/6-1})$ hypercontractive.*

Now we shall apply Theorem 3.19 in [MOO08] under the hypothesis **H1** to the degree D polynomial $(f^{\leqslant D} + 1)/2$ to conclude

$$\left| \mathbb{E}\left[\zeta\left(\frac{f^{\leqslant D}(\mathcal{X}) + 1}{2}\right)\right] - \mathbb{E}\left[\zeta\left(\frac{\bar{f}(\mathcal{Y}) + 1}{2}\right)\right] \right| \leqslant O(D^{2/3} \alpha(D)^{-2D} \tau^{1/3})$$

where $\zeta(x)$ is defined as

$$\zeta(x) = \begin{cases} (x-1)^2 & \text{if } x > 1 \\ 0 & \text{if } 0 \leqslant x \leqslant 1 \\ x^2 & \text{if } x < 0 \end{cases}$$

By the choice of τ, $O(D^{2/3} \alpha(D)^{-2D} \tau^{1/3}) = O(2^{11 D^2 \log D} \tau^{1/3}) \leqslant \eta^2/256$. Clearly we have $\zeta(\frac{f^{\leqslant D}(x)+1}{2}) = (f^{\leqslant D}(x) - \operatorname{trunc} f^{\leqslant D}(x))^2/4$. Thus we get

$$\|\bar{f} - \operatorname{trunc} \bar{f}\|^2 \leqslant \|f^{\leqslant D} - \operatorname{trunc} f^{\leqslant D}\|^2 + \eta^2/256$$

Along with Equation (14) this yields, $\|\bar{f} - f'\| = \|\bar{f} - \operatorname{trunc} \bar{f}\| \leqslant \eta/7$. By a similar argument, we get $\|\bar{g} - g'\| \leqslant \eta/7$. Using Lemma 2.4, we finish the proof:

$$\begin{aligned}
\langle T_\rho f', BT_\rho g'\rangle &\geqslant \langle T_\rho \bar{f}, BT_\rho \bar{g}\rangle - \|B\|(\|T_\rho f - T_\rho \bar{f}\| + \|T_\rho g - T_\rho \bar{g}\|) \\
&\geqslant \langle T_\rho \bar{f}, BT_\rho \bar{g}\rangle - 2\|B\|\eta/7 \\
&\geqslant \langle f^{\leqslant D}, U_\rho B U_\rho g^{\leqslant D}\rangle - 2\|B\|\eta/7 \qquad \text{Equation (15)} \\
&\geqslant \langle f, Cg\rangle - \|B\|\eta/2 \qquad \text{Equation (13)}
\end{aligned}$$

□

B Putting it together : proofs of the Theorems 1.1,1.4, 1.3

B.1 Proof of Theorem 1.1

As a rule of thumb, every dictatorship test yields a UG hardness result using by now standard techniques [KKMO07, KO06, Rag08]. For the sake of completeness, we include the proof of the following lemma in the Appendix C.1.

Lemma B.1. *Given a dictatorship test A with completeness c and soundness s, and a unique games instance G, it is possible to efficiently construct an operator $G \otimes_\rho A$ satisfies the following to two conditions:*

1. *if* $\operatorname{val}(G) \geqslant 1 - \epsilon$, *then* $\operatorname{opt}(G \otimes_\rho A) \geqslant c(1 - o_{\epsilon,\gamma,\tau\to 0}(1))$,
2. *if* $\operatorname{val}(G) < \epsilon$, *then* $\operatorname{opt}(G \otimes_\rho A) < s(1 + o_{\epsilon,\gamma,\tau\to 0}(1))$,

To finish the proof of Theorem 1.1, let A be a matrix for which the ratio of $\operatorname{sdp}(A)/\operatorname{opt}(A) \geqslant K_G - \eta$. Consider the dictatorship test $\mathcal{D}_\eta(A)$ obtained from the matrix A. By Corollary 3.3, the completeness of $\mathcal{D}_\eta(A)$ is $\operatorname{sdp}(A)(1 - \eta)$. Further by Corollary 3.7, the soundness is at most $\operatorname{opt}(A)(1 + \eta)$ 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.

B.2 Proof of Theorem 1.4

Let A be an arbitrary finite matrix for which $\operatorname{sdp}(A)/\operatorname{opt}(A) \geqslant K_G - \eta$. Consider the dictatorship test/operator $\mathcal{D}_\eta(A)$ on $L_2(\mathcal{H}^k)$. From Lemma 3.3 and Corollary 3.7, the ratio of $\mathsf{Completeness}(A)$ to $\mathsf{Soundness}_{\eta,\tau}(A)$ is at least $\operatorname{sdp}(A)/\operatorname{opt}(A) - 2\eta$ 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 3.9, for large enough choice of k, the operator $\mathcal{G}(\mathcal{D}_\eta(A))$ is an operator with $\operatorname{sdp}(A)/\operatorname{opt}(A) \geqslant K_G - 10\eta$. It is easy to see that the operator $\mathcal{G}(\mathcal{D}_\eta(A))$ has all the properties specified in Theorem 3.9.

B.3 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 $\operatorname{sdp}(A)/\operatorname{opt}(A)$. 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 $L_2(\mathcal{H}^k)$ 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

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

Let $\mathcal{P}$ denote the space of all pairs of functions $f, g \in L_2(\mathcal{H}^k)$ with $\mathrm{MaxComInf}(T_\rho f, T_\rho g) \leqslant \tau$ and $\|f\|, \|g\| \leqslant 1$.. 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 using the following finite linear program:

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

C Further Proofs

The following lemma is restatement of Lemma 3.12.

Lemma C.1. *For $f, g, f', g' \in L_2(\mathcal{H}^k)$ with $\|f\|, \|g\|, \|f'\|, \|g'\| \leqslant 1$,*

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

Proof. Define $\boldsymbol{u}'_i = \mathrm{trunc}\, T_\rho \bar{f}(\Phi \boldsymbol{u}_i), \boldsymbol{v}'_j = \mathrm{trunc}\, T_\rho \bar{g}(\Phi \boldsymbol{v}_j)$ and $\boldsymbol{u}''_i = \mathrm{trunc}\, T_\rho \bar{f}'(\Phi \boldsymbol{u}_i), \boldsymbol{v}''_j = \mathrm{trunc}\, \bar{g}'(\Phi \boldsymbol{v}_j)$. Substituting we get,

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

As trunc and T_ρ are contractive operators, $\|\boldsymbol{u}'_i\|, \|\boldsymbol{u}''_i\|, \|\boldsymbol{v}'_j\|, \|\boldsymbol{v}''_j\| \leqslant 1$. Further, observe that $\|\boldsymbol{u}'_i - \boldsymbol{u}''_i\| \leqslant \|f - f'\|$ and $\|\boldsymbol{v}'_j - \boldsymbol{v}''_j\| \leqslant \|g - g'\|$, since for all x, $|\mathrm{trunc}\, f(x) - \mathrm{trunc}\, f'(x)| \leqslant |f(x) - f'(x)|$. Substituting in the above equation, we get the required result. □

The following lemma is a restatement of Lemma 4.1.

Lemma C.2. *Let A be a dictatorship test on $L_2(\mathcal{H}^k)$. Let f, g be a pair of functions in $L_2(\mathcal{H}^k)$ with $\|f\|_\infty, \|g\|_\infty \leqslant 1$ and $\mathrm{MaxComInf}(T_\rho f, T_\rho g) \leqslant \tau$ for $\rho = 1 - \eta$.*

Then for every $\tau' > 0$, there are functions $f', g' \in L_2(\{0,1\}^k$ with $\|f'\|_\infty, \|g'\|_\infty \leqslant 1$ and $\mathrm{MaxInf}\, T_\rho f, \mathrm{MaxInf}\, T_\rho g \leqslant \tau'$ such that

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

Proof. Let J denote the set of variables i with $\mathrm{Inf}_i T_\rho f > \tau'$. Since the total influence of $T_\rho f$ is bounded by $1/\eta$ (see Fact 2.8), the set J has cardinality at most $1/\eta\tau'$. Let M_J be the orthogonal projector on the space of functions that do not depend on any variable in J. We define $f' = M_J f$ and $g' = M_J g$. We still have $\|f'\|_\infty, \|g'\| \leqslant 1$. Note that $\mathrm{Inf}_i T_\rho g \leqslant \tau$ for every $i \in J$. Hence, $\|T_\rho g - T_\rho g'\|^2 \leqslant |J|\tau$. Now,

$$\begin{aligned}\langle T_\rho f, A\, T_\rho g\rangle - \langle T_\rho f, A T_\rho g'\rangle &= \langle T_\rho f, A\, T_\rho (g - g')\rangle \\ &\leqslant \|A\| \|T_\rho f\| \cdot \|T_\rho (g-g')\| \\ &\leqslant \sqrt{|J|\tau}\,.\end{aligned}$$

On the other hand, $\langle T_\rho f, A T_\rho g'\rangle = \langle T_\rho f', A T_\rho g'\rangle$, because

$$\langle T_\rho (f - f'), A T_\rho g'\rangle = \langle T_\rho f, (I - M_J) A M_J T_\rho g\rangle = \langle T_\rho f, A (I - M_J) M_J T_\rho g\rangle = 0\,,$$

where we used the fact that the operators A and M_J commute (as both are diagonalized by the Fourier transform), and that $(I - M_J)M_J = 0$.

We repeat the same argument with the set K of variables i with $\mathrm{Inf}_i T_\rho g > \tau'$. Again, projecting on M_K changes the value of $\langle T_\rho f', A\, T_\rho g'\rangle$ by at most $\sqrt{|J|\tau}$. □

C.1 Reduction from UniqueGames to $K_{N,N}$-QuadraticProgramming

The UGC asserts that for every $\epsilon > 0$, there is a k such that for a unique game G on alphabet $[k]$ it is hard to distinguish between $\mathrm{val}(G) \geqslant 1 - \epsilon$ and $\mathrm{val}(G) < \epsilon$.

Formally, we will represent a unique game G on alphabet $[k]$ as a distribution over triples (u, v, π), where $u \in W_1$ and $v \in W_2$ are vertices, and π is a permutation of $[k]$. Here we can and will assume that the game is bipartite, i.e., W_1 and W_2 are disjoint.

Let $A = \sum_{d \in [k]} \lambda_d P_d$ be a dictatorship test on $L_2(\mathcal{H}^k)$. For $\rho = 1 - \eta$, we define a linear operator $G \otimes_\eta A$ on $\left(L_2(\mathcal{H}^k)\right)^{|W_1|+|W_2|}$ as follows:

$$\langle \boldsymbol{f}, (G \otimes_\eta A)\boldsymbol{g}\rangle := \mathop{\mathbb{E}}_{(u,v,\pi)\sim G} \langle T_\rho(\pi.f_u), A T_\rho g_v\rangle\,,$$

where $\boldsymbol{f} = (f_u)_{u \in W_1}$, $\boldsymbol{g} = (g_v)_{v \in W_2}$, and $\pi.f_u$ denotes the function $f_u(x_{\pi(1)}, \ldots, x_{\pi(k)})$.

We claim the following properties of the reduction $G \mapsto G \otimes_\eta A$. This claim implies Lemma B.1.

Claim C.3. *For $\tau, \eta \in [0, 1]$, $\rho = 1 - \eta$, and every unique game G, we have*

1. *If* $\mathrm{val}(G) \geqslant 1 - \epsilon$ *then* $\mathrm{opt}(A \otimes_\eta G) \geqslant \mathsf{Completeness}(A)(1 - O(\epsilon + \eta))$.
2. *If* $\mathrm{val}(G) < (\tau\eta)^3$ *then* $\mathrm{opt}(A \otimes_\eta G) \leqslant \mathsf{Soundness}_{\eta,\tau}(T_\rho A T_\rho) + O(\tau\eta)\mathsf{Completeness}(A)$.

Proof. By scaling[1], we may assume $\lambda_1 = 1$ and $\lambda_d \in [-1, 1]$ for all $d \in [k]$, where $A = \sum_d \lambda_d P_d$. Note that $\mathsf{Completeness}(A) = \lambda_1 = 1$.

Suppose that $\mathrm{val}(G) \geqslant 1 - \epsilon$. Then there exists a labeling $\ell\colon W_1 \cup W_2 \to [k]$ such that

$$\mathop{\mathbb{P}}_{(u,v,\pi)\sim G} \left\{\pi(\ell(u)) = \ell(v)\right\} \geqslant 1 - \epsilon\,.$$

We choose $\boldsymbol{f}$ and $\boldsymbol{g}$ such that $f_u(\boldsymbol{x}) = x_{\ell(u)}$ and $g_v(\boldsymbol{x}) = x_{\ell(v)}$ are dictator functions. If $\pi(\ell(u)) = \ell(v)$, then $\pi.f_u = g_v$. Hence, $\langle T_\rho \pi.f_u, A T_\rho g_v\rangle = \rho^2\lambda_1 = \rho^2$. On the other hand, if $\pi(\ell(u)) \neq \ell(v)$, then clearly $|\langle T_\rho \pi.f_u, A T_\rho g_v\rangle| \leqslant 1$. Thus,

$$\mathop{\mathbb{E}}_{(u,v,\pi)\sim G} \langle T_\rho(\pi.f_u), A T_\rho g_v\rangle \geqslant (1 - \epsilon)\cdot \rho^2 - \epsilon \geqslant 1 - 2\epsilon - 2\eta\,.$$

[1] Note that scaling A by a factor α, scales $\mathrm{opt}(G \otimes_\eta A)$, $\mathsf{Completeness}(A)$, and $\mathsf{Soundness}_{\eta,\tau}(A)$ by the same factor α

It follows that $\mathrm{opt}(G \otimes_\eta A) \geqslant c - o_{\epsilon\to 0}(1)$ for any game G with $\mathrm{val}(G) \geqslant 1-\epsilon$.

Now suppose that $\mathrm{opt}(G \otimes_\eta A) \geqslant \mathsf{Soundness}_{\eta,\tau}(A) + \delta$, where $\delta = \tau\eta$. In this case, we want to show that $\mathrm{val}(G) \geqslant \epsilon$ for $\epsilon = \tau\eta^3$. Let $\boldsymbol{f} = (f_u)$ and $\boldsymbol{g} = (g_v)$ be vectors with $\|f_u\|_\infty, \|g_v\|_\infty \leqslant 1$ that achieve

$$\mathop{\mathbb{E}}_{(u,v,\pi)\sim G} \langle T_\rho(\pi.f_u), AT_\rho g_v\rangle \geqslant \mathsf{Soundness}_{\eta,\tau}(A) + \delta\,. \tag{16}$$

In hindsight, let us define a set of candidate labels for vertices $u \in W_1$ and $v \in W_2$,

$$J_u = \left\{ i \mid \mathrm{Inf}_i T_\rho(f_u) > \tau \right\} \quad \text{and} \quad J_v = \left\{ i \mid \mathrm{Inf}_i T_\rho(g_v) > \tau \right\}\,.$$

Since $\rho = 1-\eta$, we have $|J_u|, |J_v| \leqslant 1/\eta\tau$ (see Fact 2.8). Since A is contracting, we get from equation (16) that

$$\mathop{\mathbb{P}}_{(u,v,\pi)\sim G} \left\{ \langle T_\rho \pi.f_u, AT_\rho g_v\rangle > \mathsf{Soundness}_{\eta,\tau}(A) \right\} \geqslant \delta\,.$$

The situation $\langle T_\rho\pi.f_u, AT_\rho g_v\rangle > \mathsf{Soundness}_{\eta,\tau}(A)$ implies that $\mathrm{Inf}_i T_\rho(\pi.f_u) > \tau$ and $\mathrm{Inf}_i T_\rho g_v > \tau$ for some $i \in [k]$. Of course, $\mathrm{Inf}_i T_\rho(\pi.f_u) > \tau$ just means that variable $\pi^{-1}(i)$ has influence $\mathrm{Inf}_{\pi^{-1}(i)} T_\rho(f_u) > \tau$. It follows that $i \in J_v$ and $\pi(i) \in J_u$. Thus,

$$\mathop{\mathbb{P}}_{(u,v,\pi)\sim G} \left\{ \exists i \in [k].\ i \in J_u \text{ and } \pi(i) \in J_v \right\} \geqslant \delta\,.$$

Hence, if we choose a random element of J_u as the label $\ell(u)$ and a random element of J_v as the label $\ell(v)$, we have

$$\mathop{\mathbb{P}}_{(u,v,\pi)\sim G} \left\{ \exists i \in [k].\ \ell(u) = i \text{ and } \ell(v) = \pi(i) \right\} \geqslant \delta \cdot (\tau\eta)^2\,,$$

where we use the fact that $|J_u||J_v| \leqslant 1/\tau^2\eta^2$. We can conclude that $\mathrm{val}(G) \geqslant \delta\tau^2\eta^2$ for every unique game G with $\mathrm{opt}(G \otimes_\eta A) \geqslant \mathsf{Soundness}_{\eta,\tau}(A) + \delta$. By our choice of τ and δ, we have $\delta\tau^2\eta^2 = \epsilon$. Hence, we get $\mathrm{val}(G) \geqslant (\tau\eta)^3$ for every G with $\mathrm{opt}(G \otimes_\eta A) \geqslant \mathsf{Soundness}_{\eta,\tau}(A) + \delta = \mathsf{Soundness}_{\eta,\tau}(A) + \tau\eta$.

□

D Hypercontractivity, Hermite Polynomials and Invariance Principle

Lemma D.1. *For all D, the ensemble $\mathcal{X} = \{H_0(x), \ldots, H_D(x)\}$ is $(2, 3, \alpha(D) = 2^{-5D\log D})$ hypercontractive. The ensemble $\mathcal{Y} = \{\chi_i(y)\}$ is $(2, 3, 2^{-D/6-1})$ hypercontractive.*

Proof. The ensemble $\mathcal{Y}$ consists of random variables over the finite probability space consisting of uniform measure of $\{1,-1\}^D$. The probability of the smallest atom is 2^{-D}. Hence by Proposition 3.15 in [MOO08], the ensemble $\mathcal{Y}$ is $(2, 3, 2^{-D/6-1})$ hypercontractive.

To show the hypercontractivity of the ensemble $\mathcal{X}$, it is sufficient to show that every linear combination of non-constant random variables from the ensemble is hypercontractive. Specifically, consider a random variable $X = \sum_{d=1}^{D} c_d H_d(x)$ where x is a normal random variable. By Proposition 3.16 in [MOO08], the variable X is $(2, 3, \alpha)$ hypercontractive for $\alpha = \|X\|_2 / 2\sqrt{2}\|X\|_3$. Thus to show hypercontractivity of the ensemble $\mathcal{X}$ it is sufficient to upper bound $\|X\|_3$ for all $X = \sum_{d=1}^{D} c_d H_d(x)$ with $\sum_{d=1}^{D} c_d^2 = 1$.

We will derive some rough upper bound on the $\|X\|_3$. using the following well known facts:

– All the coefficients of the polynomial $H_d(x)$ are strictly bounded in absolute value by $d!$.

– For a normal random variable x, $\mathbb{E}_{x\in\mathcal{G}}[x^{2t}] = \frac{2t!}{2^t t!}$.

In particular, this gives :

$$||X||_3 < \sum_{d=1}^{D} ||H_d(x)||_3 < \sum_{d=1}^{D} d! \cdot (\sum_{i=0}^{d} ||x^i||_3) < D! \cdot D \cdot (3D!)^{\frac{1}{3}} < 2^{5D \log D} \tag{17}$$

Hence the ensemble $\mathcal{X}$ is $(2, 3, \alpha(D) = 2^{-5D \log D})$ hypercontractive. $\square$