

Tight Hardness of the Non-commutative Grothendieck Problem

Jop Briët, Oded Regev
Courant Institute of Mathematical Sciences, New York University
New York City, USA

Rishi Saket
IBM Research
Bangalore, India

Abstract

We prove that it is NP-hard to approximate the non-commutative Grothendieck problem to within any constant factor larger than one-half, which matches the approximation ratio of the algorithm of Naor, Regev, and Vidick (STOC'13). Our proof uses an embedding of finite-dimensional Hilbert spaces into the space of matrices endowed with the trace norm with the property that the image of standard basis vectors is longer than that of unit vectors with no large coordinates.

Keywords

Grothendieck inequality, hardness, semidefinite programming

I. INTRODUCTION

The subject of this paper, the *non-commutative Grothendieck problem*, has its roots in celebrated work of Grothendieck [1], sometimes (jokingly?) referred to as “*Grothendieck’s résumé*.” His paper laid the foundation for the study of the geometry of tensor products of Banach spaces, though its significance only became widely recognized after it was revamped by Lindenstrauss and Pełczyński [2]. The main result of the paper, now known as *Grothendieck’s inequality*, shows a close relationship between the following two quantities. For a complex $d \times d$ matrix M let

$$\text{OPT}(M) = \sup_{\alpha_i, \beta_j} \left| \sum_{i,j=1}^d M_{ij} \alpha_i \overline{\beta_j} \right|, \quad (1)$$

where the supremum goes over scalars on the complex unit circle, and let

$$\text{SDP}(M) = \sup_{a_i, b_j} \left| \sum_{i,j=1}^d M_{ij} \langle a_i, b_j \rangle \right|, \quad (2)$$

where the supremum goes over vectors on a complex Euclidean unit sphere of any dimension. Since the circle is the sphere in dimension one, we clearly have $\text{SDP}(M) \geq \text{OPT}(M)$. Grothendieck’s inequality states that there exists a universal constant $K_G^{\mathbb{C}} < \infty$ such that for each positive integer d and any $d \times d$ matrix M , we also have $\text{SDP}(M) \leq K_G^{\mathbb{C}} \text{OPT}(M)$. This result found an enormous number of applications both within and far beyond its original scope and we give some examples below (see [3], [4] for extensive surveys). Despite this, finding the optimal value of $K_G^{\mathbb{C}}$ is the only one of six problems posed in [1] that remains unsolved today; the current best upper and lower bounds are 1.4049 [5] and 1.338 [6], respectively. The situation is similar for the real variant of the problem, where all objects involved are over the real numbers and the constant is denoted K_G ; see [7] for recent progress on the problem of determining K_G .

The non-commutative Grothendieck problem, which we will refer to simply as the *NCG*, is the optimization problem in which we are asked to maximize a given bilinear form over two unitary matrices. More explicitly, we are given a four-dimensional array of complex numbers $(T_{ijkl})_{i,j,k,l=1}^d$ and are asked to find or approximate the value

$$\text{OPT}(T) = \sup_{A, B} \left| \sum_{i,j,k,l=1}^d T_{ijkl} A_{ij} \overline{B_{kl}} \right|, \quad (3)$$

where the supremum is over pairs of $d \times d$ unitary matrices. (The word “non-commutative” simply refers to the fact that optimization is over matrices.) It is not difficult to see that the (commutative) Grothendieck problem of computing $\text{OPT}(M)$ as in (1) is the special case where T has $T_{iijj} = M_{ij}$ and zeros elsewhere. Seen at first, the problem might seem overly abstract, but in fact, as we will illustrate below, it captures many natural questions as special cases. Grothendieck conjectured that his namesake inequality has an extension that relates (3) and the quantity

$$\text{SDP}(T) = \sup_{\vec{A}, \vec{B}} \left| \sum_{i,j,k,l=1}^d T_{ijkl} \langle \vec{A}_{ij}, \vec{B}_{kl} \rangle \right|, \quad (4)$$

where \vec{A}, \vec{B} range over all $d \times d$ matrices whose entries are complex *vectors* of arbitrary dimension satisfying a certain “unitarity” constraint.¹ Namely, he conjectured that there exists a universal constant $K < \infty$ such that for every positive integer d and array T as above, we have $\text{OPT}(T) \leq \text{SDP}(T) \leq K \text{OPT}(T)$, where the first inequality follows immediately from the definition. Over twenty-five years after being posed, the non-trivial content of Grothendieck’s conjecture, $\text{SDP}(T) \leq K \text{OPT}(T)$, was finally settled in the positive by Pisier [8]. This result is now known as the non-commutative Grothendieck inequality. In contrast with the commutative case, and somewhat surprisingly, the optimal value of K is known: Haagerup [9] lowered Pisier’s original estimate to $K \leq 2$ and this was later shown to be sharp by Haagerup and Itoh [10].

Algorithmic applications: The importance of Grothendieck’s inequality to computer science was pointed out by Alon and Naor [11], who placed it in the context of approximation algorithms for combinatorial optimization problems. They observed that computing $\text{SDP}(M)$ is a semidefinite programming (SDP) problem that can be solved efficiently (to within arbitrary precision), and they translated an upper bound of about 1.78 on K_G due to Krivine [12] to an efficient rounding scheme that turns SDP vectors into a feasible solution for the real Grothendieck problem (1) achieving value at least $\text{SDP}(M)/1.78$. It is known that whatever the value of K_G is, there exists an efficient algorithm achieving value at least $(1/K_G - \varepsilon) \text{SDP}(M)$ for any constant $\varepsilon > 0$. (This was first shown in [13], but can also be derived from the results in [14] using a simple discretization argument combined with a brute-force search.) The Grothendieck problem shows up in a number of different areas such as graph partition problems and computing the Cut-Norm of a matrix [11], in statistical physics where it gives ground state energies in the spin glass model [15], and in quantum physics where it is related to Bell inequalities [16].

In the same spirit, Naor, Regev, and Vidick [17] recently translated the *non-commutative* Grothendieck inequality into an efficient SDP-based approximation algorithm for the NCG problem (3) that achieves value at least $\text{SDP}(T)/2$. They also considered the real variant and a Hermitian variant, for which they gave analogous algorithms achieving value at least $\text{SDP}(T)/2\sqrt{2}$. This in turn implies efficient constant-factor approximation algorithms for a variety of problems, including the Procrustes problem and two robust versions of Principle Component Analysis [17] and quantum XOR games [18]. In a related result, Bandeira et al. [19] considered a special case of the NCG (in fact a special case even of the Little NCG defined below) and showed how to obtain better approximation factors for it. They also show that it is rich enough to capture some applications such as the Procrustes problem and another natural problem called the Global Registration Problem.

Hardness of approximation: For simplicity we momentarily turn to the real setting, but similar results hold over the complex numbers. Since the Grothendieck problem contains MAXCUT as a special case (by taking M to be the positive semidefinite Laplacian matrix of a graph), Håstad’s inapproximability result [20] implies that it is NP-hard to approximate the value (1) to any factor larger than $16/17 \approx .941$. Based on the current best-known lower bound of about 1.676 on K_G , Khot and O’Donnell [21] proved that (1) is Unique-Games-hard to approximate to within a factor larger than $1/1.676 \approx .597$. Moreover, despite the fact that the exact value of K_G is still unknown, Raghavendra and Steurer [13] were able to improve this Unique Games hardness to $1/K_G$. (See for instance [22], [23] for background on the Unique Games conjecture.)

Our result: Whereas the hardness situation for the commutative version of Grothendieck’s problem is reasonably well understood (apart from the yet-unknown exact value of K_G), no tight hardness result was previously known for the non-commutative version. In fact, we are not even aware of any hardness result that is better than what follows from the commutative case. Here we settle this question.

Theorem I.1. *For any constant $\varepsilon > 0$ it is NP-hard to approximate the optimum (3) of the non-commutative Grothendieck problem to within a factor greater than $1/2 + \varepsilon$.*

Little Grothendieck: In fact, we prove a stronger result than Theorem I.1 that concerns a special case of the NCG called the *Little NCG*. Let us start by describing the (real case of the) *commutative* Little Grothendieck problem (a.k.a. the positive-semidefinite Grothendieck problem). A convenient way to phrase it is as asking for the operator norm of a linear map $\mathcal{F} : \mathbb{R}^n \rightarrow \ell_1^d$, defined as $\|\mathcal{F}\| = \sup_a \|\mathcal{F}(a)\|_{\ell_1}$ where the vector a ranges over the n -dimensional Euclidean unit ball. It turns out that that this is a special case of (the real version of) Eq. (1): for any \mathcal{F} there exists a positive semidefinite $d \times d$ matrix M such that $\text{OPT}(M) = \|\mathcal{F}\|^2$; and vice versa, one can also map any such M into a corresponding operator \mathcal{F} (see, e.g., [4] or Section VI). We wish to highlight that for such instances, the constant K_G may be replaced by the smaller value $\pi/2$ [24] and that this value is known to be optimal [1]. Moreover, Nesterov made this algorithmic, namely, he showed an algorithm that approximates $\|\mathcal{F}\|$ as above to within $\sqrt{2/\pi}$ [25]. Finally, Khot and Naor [26], as part of a more general result, showed that this is tight: the Unique-Games-hardness threshold for the Little Grothendieck problem is exactly $\sqrt{2/\pi}$.

¹Namely, we require that $\vec{A}^* \vec{A} = \mathbb{1}$ and $\vec{A} \vec{A}^* = \mathbb{1}$ and similarly for \vec{B} , where the multiplication of two vector-entried matrices is a scalar-valued matrix computed just like a normal matrix multiplication except the scalar multiplication is replaced by an inner product, e.g., the (i, j) -coordinate of $\vec{A} \vec{A}^*$ is given by $\sum_k \langle \vec{A}_{ik}, \vec{A}_{jk} \rangle$.

As an aside, we note that other operator norms, in particular of operators in $\mathbb{R}^n \rightarrow \ell_4^d$, have played an important role recently in theoretical computer science (see, e.g., [27]).

The Little *non-commutative* Grothendieck problem is formulated in terms of the (normalized) trace norm, also known as the Schatten-1 norm, which for a $d \times d$ matrix A is given by $\|A\|_{S_1} = d^{-1} \text{Tr} \sqrt{A^*A}$. In other words, $\|A\|_{S_1}$ is the average of the singular values of A . The space of matrices endowed with this norm is denoted by S_1 and by S_1^d if we restrict to $d \times d$ matrices. The problem then asks for the operator norm of a linear map $\mathcal{F} : \mathbb{C}^n \rightarrow S_1^d$. This problem is a special case of the NCG where $\text{OPT}(T) = \|\mathcal{F}\|^2$ (see Section VI). In particular, it follows from [9], [17] that there is an efficient SDP-based $1/\sqrt{2}$ -approximation algorithm for the Little NCG. Our stronger result alluded to above shows tight hardness for the Little NCG, which directly implies Theorem I.1.

Theorem I.2. *For any constant $\varepsilon > 0$ it is NP-hard to approximate the Little non-commutative Grothendieck problem to within a factor greater than $1/\sqrt{2} + \varepsilon$.*

While this result applies to the complex case, an easy transformation shows that it directly implies the same result for the real and Hermitian cases introduced in [17] (see Section V-A). Finally, as we show in the “warm-up” section of this paper (Section IV), we also get a tight NP-hardness result for the *commutative* Little Grothendieck problem, strengthening the unique-games-based result of [26].

Theorem I.3. *For any constant $\varepsilon > 0$ it is NP-hard to approximate the real Little commutative Grothendieck problem to within a factor greater than $\sqrt{2/\pi} + \varepsilon$. Similarly, the complex case is NP-hard to approximate to within a factor greater than $\sqrt{\pi/4} + \varepsilon$.*

Techniques: Nearly all recent work on hardness of approximation, including for commutative Grothendieck problems [13], [26], uses the machinery of Fourier analysis over the hypercube, influences, or the majority is stablest theorem. Our attempts to apply these techniques here failed. Instead, we use a more direct approach similar to that taken in [28] and avoid the use of the hypercube altogether. The role of dictator functions is played in our proof simply by the standard basis vectors of \mathbb{C}^n . The dictatorship test, which is our main technical contribution, comes in the form of a linear operator $\mathcal{F} : \mathbb{C}^n \rightarrow S_1^d$ with the following notable property: it maps the n standard basis vectors to matrices with trace norm 1, and it maps any unit vector with no large coordinate to a matrix with trace norm close to $1/\sqrt{2}$. Roughly speaking, one can think of \mathcal{F} as identifying an interesting subspace of S_1^d in which the unit ball looks somewhat like the intersection of the Euclidean ball with a (scaled) ℓ_∞ ball.

A first attempt to construct an operator \mathcal{F} as above might be to map each standard basis vector to a random unitary matrix. This, however, leads to a very poor map – while standard basis vectors are mapped to matrices of trace norm 1, vectors with no large coordinates are mapped to matrices of trace norm close to $8/(3\pi) \approx 0.848$ by Wigner’s semicircle law. Another natural approach is to look at the construction by Haagerup and Itoh [10] (see also [4, Section 11] for a self-contained description) which shows the factor-2 lower bound in the non-commutative Grothendieck inequality, i.e., the tight integrality gap of the SDP (4). Their construction relies on the so-called CAR algebra (after canonical anticommutation relations) and provides an *isometric* mapping from \mathbb{C}^n to S_1 , i.e., all unit vectors are mapped to matrices of trace norm 1. Directly modifying this construction (akin to how, e.g., Khot et al. [29] obtained tight hardness of MAXCUT by restricting the tight integrality gap instances by Feige and Schechtman [30] from the sphere to the hypercube) does not seem to work. Instead, our construction of \mathcal{F} relies on a different (yet related) algebra known as the Clifford algebra. The Clifford algebra was used before in a celebrated result by Tsirelson [16] (to show that Grothendieck’s inequality can be interpreted as a statement about XOR games with entanglement). His result crucially relies on the fact that the Clifford algebra gives an *isometric* mapping from \mathbb{R}^n to S_1 . Notice that this is again an isometric embedding, but now only over the reals. Our main observation here (Lemma V.2) is that the same mapping, when extended to \mathbb{C}^n , exhibits intriguing cancellations when the phases in the input vectors are not aligned, and this leads to the construction of \mathcal{F} (Lemma V.1). Even though the proof of this fact is simple, we find it surprising; we are not aware of any previous application of those complex Clifford algebras in computer science (or elsewhere for that matter).

Open questions: For the real and Hermitian cases there is a gap of $\sqrt{2}$ between the guarantee of the [17] algorithms and our hardness result. It also would be interesting to explore whether hardness of approximation results can be derived to some of the applications of the NCG, including the Procrustes problem and robust Principle Component Analysis. We believe that our embedding would be useful there too.

Outline: The rest of the paper is organized as follows. In Section II, we set some notational conventions, gather basic preliminary facts about relevant Banach spaces, and give a detailed formulation of the Smooth Label Cover problem. In Section III, we prove hardness of approximation for the problem of computing the norm of a general class of Banach-space-valued functions, closely following [28]. In Section IV, as a “warm up,” we prove Theorem I.3 using the generic result of

Section III and straightforward applications of real and complex versions of the Berry-Esséen Theorem. Section V contains our main technical contribution, which we use there to finish the proof of our main result (Theorem I.2).

II. PRELIMINARIES

Notation and relevant Banach spaces: For a positive integer n we denote $[n] = \{1, \dots, n\}$. For a graph G and vertices $v, w \in V(G)$ we write $v \sim w$ to denote that v and w are adjacent. We write $\Pr_{e \sim v}[\cdot]$ for the expectation with respect to a uniformly distributed random edge with v as an endpoint. For a finite set U we denote by $\mathbf{E}_{u \in U}[\cdot]$ the expectation with respect to the uniform distribution over U . All Banach spaces are assumed to be finite-dimensional. Recall that for Banach spaces X, Y the operator norm of a linear operator $\mathcal{F} : X \rightarrow Y$ is given by

$$\|\mathcal{F}\| = \sup_{x \in X: \|x\|_X \leq 1} \|\mathcal{F}(x)\|_Y.$$

For a real number $p \geq 1$, the p -norm of a vector $a \in \mathbb{C}^n$ is given by

$$\|a\|_{\ell_p} = \left(\sum_{i=1}^n |a_i|^p \right)^{1/p}.$$

As usual we implicitly endow \mathbb{C}^n with the Euclidean norm $\|a\|_{\ell_2}$. For a finite set U endowed with the uniform probability measure we denote by $L_p(U)$ the space of functions $f : U \rightarrow \mathbb{C}$ with the norm

$$\|f\|_{L_p(U)} = \left(\mathbf{E}_{u \in U} [|f(u)|^p] \right)^{1/p}.$$

More generally, for a Banach space X we denote by $L_p(U, X)$ the space of functions $f : U \rightarrow X$ with the norm

$$\|f\|_{L_p(U, X)} = \left(\mathbf{E}_{u \in U} [\|f(u)\|_X^p] \right)^{1/p}.$$

We will write $L_p(X)$ if U is not explicitly given and $\|f\|_{L_p}$ instead of $\|f\|_{L_p(U, X)}$ when there is no danger of ambiguity. Note that $L_2(U, \mathbb{C}^n)$ is a Hilbert space.

Smooth Label Cover: An instance of Smooth Label Cover is given by a quadruple $(G, [n], [k], \Sigma)$ that consists of a regular connected (undirected) graph $G = (V, E)$, a label set $[n]$ for some positive integer n , and a collection $\Sigma = ((\pi_{ev}, \pi_{ew}) : e = (v, w) \in E)$ of pairs of maps both from $[n]$ to $[k]$ associated with the endpoints of the edges in E . Given an assignment $A : V \rightarrow [n]$, we say that an edge $e = (v, w) \in E$ is *satisfied* if $\pi_{ev}(A(v)) = \pi_{ew}(A(w))$.

The following hardness result for Smooth Label Cover, given in [28],² is a slight variant of the original construction due to Khot [31]. The theorem also describes the various structural properties, including smoothness, that are satisfied by the hard instances.

Theorem II.1. *For any positive real numbers ζ, γ there exist positive integers $n = n(\zeta, \gamma)$, $k = k(\zeta, \gamma)$, and $t = t(\zeta)$, and a family of Smooth Label Cover instances $(G, [n], [k], \Sigma)$ as above such that:*

- *(Hardness): It is NP-hard to distinguish between the following two cases:*
 - *(YES Case): There is an assignment that satisfies all edges.*
 - *(NO Case): Every assignment satisfies less than a ζ -fraction of the edges.*
- *(Structural properties):*
 - *(Smoothness): For every vertex $v \in V$ and distinct $i, j \in [n]$, we have*

$$\Pr_{e \sim v} [\pi_{ev}(i) = \pi_{ev}(j)] \leq \gamma. \tag{5}$$

- *For every vertex $v \in V$, edge $e \in E$ incident on v , and $i \in [k]$, we have $|\pi_{ev}^{-1}(i)| \leq t$; that is, at most t elements in $[n]$ are mapped to the same element in $[k]$.*
- *(Weak Expansion): For any $\delta > 0$ and vertex subset $V' \subseteq V$ such that $|V'| = \delta \cdot |V|$, the number of edges between the vertices in V' is at least $(\delta^2/2)|E|$.*

²For convenience, we make implicit some of the parameters in the statement of the theorem.

III. HARDNESS FOR GENERAL BANACH-SPACE VALUED OPERATORS

The following proposition shows hardness of approximation for the problem of computing the norm of a linear map from \mathbb{C}^n to any Banach space that allows for a “dictatorship test,” namely, a linear function that maps the standard basis vectors to long vectors, and maps “spread” unit vectors to short vectors. As stated, the proposition assumes the underlying field to be \mathbb{C} ; we note that the proposition holds with exactly the same proof also in the case of the real field \mathbb{R} .

Theorem III.1. *Let $(X_n)_{n \in \mathbb{N}}$ be a family of finite-dimensional Banach spaces, and η and τ be positive numbers such that $\eta > \tau$. Suppose that for each positive integer n there exists a linear operator $f : \mathbb{C}^n \rightarrow X_n$ with the following properties:*

- For any vector $a \in \mathbb{C}^n$, we have $\|f(a)\|_{X_n} \leq \|a\|_{\ell_2}$.
- For each standard basis vector e_i , we have $\|f(e_i)\|_{X_n} \geq \eta$.
- For any $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon) > 0$ such that $\|f(a)\|_{X_n} > (\tau + \varepsilon)\|a\|_{\ell_2}$ implies $\|a\|_{\ell_4} > \delta\|a\|_{\ell_2}$.

Then, for any $\varepsilon' > 0$ there exists a positive integer n such that it is NP-hard to approximate the norm of an explicitly given linear operator $\mathcal{F} : L_2 \rightarrow L_1(X_n)$ to within a factor greater than $(\tau/\eta) + \varepsilon'$.

A. The hardness reduction

To set up the reduction, we begin by defining a linear operator $\mathcal{F} = \mathcal{F}_{\zeta, \gamma}$ for any choice of positive real numbers ζ, γ . Afterwards we show that there is a choice of these parameters giving the desired result. For positive real numbers ζ, γ , let n, k , and t be positive integers (depending on ζ, γ) and $(G, [n], [k], \Sigma)$ a Smooth Label Cover instance as in Theorem II.1, where $G = (V, E)$ is a regular graph. Note that ζ controls the “satisfiability” of the instance in the NO case, that γ controls the “smoothness,” and that t depends on ζ only. Endow the vertex set V with the uniform probability measure. To define \mathcal{F} we consider a special linear subspace \mathcal{H} of the Hilbert space $L_2(V, \mathbb{C}^n)$. It will be helpful to view a vector $\mathbf{a} \in L_2(V, \mathbb{C}^n)$ as an assignment $\mathbf{a} = (a_v)_{v \in V}$ of vectors $a_v \in \mathbb{C}^n$ to V . Let $\mathcal{H} \subseteq L_2(V, \mathbb{C}^n)$ be the subspace of vectors $\mathbf{a} = (a_v)_{v \in V}$ that satisfy for every $e = (v, w) \in E$ and $j \in [k]$ the homogeneous linear constraint

$$\sum_{i \in \pi_{ev}^{-1}(j)} a_v(i) = \sum_{i \in \pi_{ew}^{-1}(j)} a_w(i), \quad (6)$$

where $a_v(i)$ denotes the i th coordinate of the vector a_v . Notice that if an assignment $A : V \rightarrow [n]$ satisfies the edge $e = (v, w)$, then the standard basis vectors $a_v = e_{A(v)}$ and $a_w = e_{A(w)}$ satisfy (6); indeed, if $\pi_{ev}(A(v)) = \pi_{ew}(A(w)) = j'$ then both sides of (6) equal 1 if $j = j'$ and equal zero otherwise.

Now let η, τ and f be as in Theorem III.1. We associate with the Smooth Label Cover instance $(G, [n], [k], \Sigma)$ from above the linear operator $\mathcal{F} : \mathcal{H} \rightarrow L_1(V, X_n)$ given by,

$$(\mathcal{F}(\mathbf{a})) (v) = f(a_v). \quad (7)$$

The operator \mathcal{F} thus maps a \mathbb{C}^n -valued assignment $\mathbf{a} = (a_v)_{v \in V}$ satisfying (6) to an X_n -valued assignment given by $f(a_v)$ for each $v \in V$. Theorem III.1 follows from the following two lemmas, which we prove in Sections III-B and III-C, respectively.

Lemma III.2 (Completeness). *Suppose that there exists an assignment $A : V \rightarrow [n]$ that satisfies all the edges in E . Then, $\|\mathcal{F}\| \geq \eta$.*

Lemma III.3 (Soundness). *For any $\varepsilon > 0$ there exists a choice of $\zeta, \gamma > 0$ such that if $\|\mathcal{F}\| > \tau + 4\varepsilon$ then there exists an assignment that satisfies at least a ζ -fraction of the edges of G .*

Proof of Theorem III.1: Let $\varepsilon > 0$ be arbitrary, let ζ, γ be as in Lemma III.3, and let $n = n(\zeta, \gamma)$ and $k = k(\zeta, \gamma)$ be as in Theorem II.1. We use the reduction described above, which maps a Smooth Label Cover instance $(G, [n], [k], \Sigma)$ to the linear operator $\mathcal{F} : L_2 \rightarrow L_1(X_n)$ specified in (7). By Lemma III.2, YES instances are mapped to \mathcal{F} satisfying $\|\mathcal{F}\| \geq \eta$, whereas by Lemma III.3, NO instances are mapped to \mathcal{F} satisfying $\|\mathcal{F}\| < \tau + 4\varepsilon$. We therefore obtain hardness of approximation to within a factor $(\tau + 4\varepsilon)/\eta$. Since ε is arbitrary, we are done. ■

B. Completeness

Here we prove Lemma III.2.

Proof of Lemma III.2: Let $A : V \rightarrow [n]$ be an assignment that satisfies all the edges. Consider the vector $\mathbf{a} \in L_2(V, \mathbb{C}^n)$ where $a_v = e_{A(v)}$ and notice that $\|\mathbf{a}\|_{L_2} = 1$. Since A satisfies all edges, \mathbf{a} satisfies the constraint (6) for every $e \in E$ and $j \in [k]$, and thus \mathbf{a} lies in the domain \mathcal{H} of \mathcal{F} . Moreover, by the second property of f given in Theorem III.1,

$$\|\mathcal{F}(\mathbf{a})\|_{L_1(V, X_n)} = \mathbf{E}_{v \in V} [\|f(a_v)\|_{X_n}] \geq \eta.$$

Hence, $\|\mathcal{F}\| \geq \eta$. ■

C. Soundness

Here we prove Lemma III.3 and show that among the family of operators $\mathcal{F} = \mathcal{F}_{\zeta, \gamma}$ as in (7), for any $\varepsilon > 0$ there is a choice of $\zeta, \gamma > 0$ such that if $\|\mathcal{F}\| > \tau + 4\varepsilon$, then there exists an assignment satisfying a ζ -fraction of the edges in the Smooth Label Cover instance associated with \mathcal{F} . To begin, assume that $\|\mathcal{F}\| > \tau + 4\varepsilon$ for some $\varepsilon > 0$. Let $\mathbf{b} \in \mathcal{H}$ be a vector such that $\|\mathbf{b}\|_{L_2} = 1$ and

$$\mathbf{E}_{v \in V} [\|f(b_v)\|] = \|\mathcal{F}(\mathbf{b})\|_{L_1} \geq \tau + 4\varepsilon. \quad (8)$$

The weak expansion property in Theorem II.1 implies that it suffices to find a ‘‘good’’ assignment for a large subset of the vertices, as any large set of vertices will induce a large set of edges. For $\delta = \delta(\varepsilon)$ as in Theorem III.1, we will consider set of vertices

$$V_0 = \{v \in V \mid \|b_v\|_{\ell_4} > \delta\varepsilon \text{ and } \|b_v\|_{\ell_2} \leq 1/\varepsilon\}. \quad (9)$$

The following lemma shows that V_0 contains a significant fraction of vertices.

Lemma III.4. *For $V_0 \subseteq V$ defined as in (9), we have $|V_0| \geq \varepsilon^2|V|$.*

Proof:

Define the sets

$$\begin{aligned} V_1 &= \{v \in V \mid \|b_v\|_{\ell_4} \leq \delta\varepsilon \text{ and } \|b_v\|_{\ell_2} < \varepsilon\}, \\ V_2 &= \{v \in V \mid \|b_v\|_{\ell_4} \leq \delta\varepsilon \text{ and } \|b_v\|_{\ell_2} \geq \varepsilon\}, \\ V_3 &= \{v \in V \mid \|b_v\|_{\ell_2} > 1/\varepsilon\}. \end{aligned}$$

From (8), we have

$$\sum_{v \in V_0} \|f(b_v)\|_{X_n} + \sum_{v \in V_1} \|f(b_v)\|_{X_n} + \sum_{v \in V_2} \|f(b_v)\|_{X_n} + \sum_{v \in V_3} \|f(b_v)\|_{X_n} \geq (\tau + 4\varepsilon)|V|. \quad (10)$$

We bound the four sums on the left-hand side of (10) individually. Since (by the first item in Theorem III.1) we have $\|f(b_v)\|_{X_n} \leq \|b_v\|_{\ell_2}$, and since $\|b_v\|_{\ell_2} \leq 1/\varepsilon$ for every $v \in V_0$, the first sum in (10) can be bounded by

$$\sum_{v \in V_0} \|f(b_v)\|_{X_n} \leq |V_0|/\varepsilon. \quad (11)$$

Similarly, using the definition of V_1 the second sum in (10) is at most $\varepsilon|V|$. Next, from the third property of f in Theorem III.1, for each $v \in V_2$, we have $\|f(b_v)\|_{X_n} \leq (\tau + \varepsilon)\|b_v\|_{\ell_2}$. Therefore, the third sum in (10) is bounded as

$$\begin{aligned} \sum_{v \in V_2} \|f(b_v)\|_{X_n} &\leq (\tau + \varepsilon) \sum_{v \in V_2} \|b_v\|_{\ell_2} \\ &\leq (\tau + \varepsilon)|V_2|^{\frac{1}{2}} \left(\sum_{v \in V_2} \|b_v\|_{\ell_2}^2 \right)^{\frac{1}{2}} \quad (\text{By Cauchy-Schwartz}) \\ &\leq (\tau + \varepsilon)|V|^{\frac{1}{2}} \left(\sum_{v \in V} \|b_v\|_{\ell_2}^2 \right)^{\frac{1}{2}} \\ &= (\tau + \varepsilon)|V|, \end{aligned} \quad (12)$$

where the last inequality uses $\|\mathbf{b}\|_{L_2} = 1$. Finally, the fourth sum in (10) is bounded by

$$\begin{aligned} \sum_{v \in V_3} \|f(b_v)\|_{X_n} &\leq \sum_{v \in V_3} \|b_v\|_{\ell_2} \quad (\text{By the property of } f) \\ &< \sum_{v \in V_3} \varepsilon \|b_v\|_{\ell_2}^2 \\ &\leq \varepsilon \sum_{v \in V} \|b_v\|_{\ell_2}^2 \\ &= \varepsilon|V| \|\mathbf{b}\|_{L_2}^2 = \varepsilon|V|. \end{aligned} \quad (13)$$

Combining the above with Equation (10) yields, $|V_0|/\varepsilon \geq \varepsilon|V|$, which proves Lemma III.4. \blacksquare

Lemma III.4 and the weak expansion property implies that the set $E(V_0)$ of edges induced by V_0 has cardinality

$$|E(V_0)| \geq (\varepsilon^4/2)|E|. \quad (14)$$

We set out to show that there exists an assignment to the vertices in V_0 that satisfies a significant fraction of edges in $E(V_0)$. Roughly speaking, we do this by randomly assigning each $v \in V_0$ one of the coordinates of the vector b_v at which it has large magnitude. (Assigning the largest coordinate may not work.) The following simple proposition shows that those vectors indeed have large coordinates.

Proposition III.5. *Let $\beta = \beta(\varepsilon)$ be given by $\beta = \delta^2 \varepsilon^3$. Then, for each $v \in V_0$, we have $\|b_v\|_{\ell_\infty} \geq \beta$.*

Proof: For every $v \in V_0$, we have

$$\delta^4 \varepsilon^4 \leq \|b_v\|_{\ell_4}^4 \leq \|b_v\|_{\ell_\infty}^2 \|b_v\|_{\ell_2}^2 \leq \|b_v\|_{\ell_\infty}^2 / \varepsilon^2, \quad (15)$$

giving the claim. ■

For the to-be-determined value of ζ let $t = t(\zeta)$ be as in Theorem II.1 and for each $v \in V_0$ let

$$A_1^v = \left\{ i \in [n] \mid |b_v(i)| \geq \frac{\beta}{4} \right\} \quad \text{and} \quad A_2^v = \left\{ i \in [n] \mid |b_v(i)| \geq \frac{\beta}{4t} \right\}.$$

By Proposition III.5 these sets are nonempty and clearly $A_1^v \subseteq A_2^v$. Moreover, since $\|b_v\|_{\ell_2} \leq 1/\varepsilon$, we have,

$$|A_1^v| \leq \frac{16}{\varepsilon^2 \beta^2} \quad \text{and} \quad |A_2^v| \leq \frac{16t^2}{\varepsilon^2 \beta^2}. \quad (16)$$

Now consider a random assignment $A : V_0 \rightarrow [n]$ that independently assigns each vertex $v \in V_0$ a uniformly random label from A_1^v and assigns the remaining vertices in V some fixed arbitrary label. The following lemma shows that on average, this assignment satisfies a significant fraction of edges.

Lemma III.6. *There exists a $\gamma > 0$ depending only on ε and ζ such that for some absolute constant $c > 0$ the expected fraction of edges in E satisfied by the random assignment A given above is at least $c\varepsilon^8 \beta^4$.*

Setting γ appropriately as in the above lemma and $\zeta = c\varepsilon^8 \delta^4$ then gives Lemma III.3; indeed, notice that then ζ , and therefore also γ , depend on ε alone.

The remainder of this section is devoted to the proof of Lemma III.6. Let $E' \subseteq E(V_0)$ be the subset of edges $e = (v, w)$ whose projections π_{ev} and π_{ew} are injective on the subsets A_2^v and A_2^w respectively. Formally,

$$E' = \{e = (v, w) \in E(V_0) \mid |\pi_{ev}(A_2^v)| = |A_2^v|, \text{ and } |\pi_{ew}(A_2^w)| = |A_2^w|\}. \quad (17)$$

We set the parameter γ according to the following proposition which shows a lower bound on $|E'|$ using the smoothness property. Recall that t is a function of ζ only.

Proposition III.7. *There exists an absolute constant $c' > 0$ such that for any $\gamma \leq c'\varepsilon^8 \beta^4 / t^4$, the set E' has cardinality $|E'| \geq (\varepsilon^4/4)|E|$.*

Proof: Consider any vertex $v \in V_0$. By the smoothness property of Theorem II.1 and a union bound over all distinct pairs $i, j \in A_2^v$, the fraction of edges $e \in E$ incident on v that do not satisfy

$$|\pi_{ev}(A_2^v)| = |A_2^v|, \quad (18)$$

is at most

$$\frac{\gamma |A_2^v|^2}{2} \leq \frac{1}{2} \left(\frac{\varepsilon^8 \beta^4}{2^{10} \cdot t^4} \right) \left(\frac{16^2 \cdot t^4}{\varepsilon^4 \beta^4} \right) = \frac{\varepsilon^4}{8},$$

via an appropriate setting of c' . Therefore, the number of edges in E that are incident on some $v \in V_0$ and do not satisfy (18) is at most

$$\sum_{v \in V_0} \frac{\varepsilon^4}{8} \deg(v) \leq \frac{\varepsilon^4}{8} \sum_{v \in V} \deg(v) \leq \frac{\varepsilon^4}{4} |E|.$$

Thus,

$$|E'| \geq |E(V_0)| - (\varepsilon^4/4)|E| \geq (\varepsilon^4/4)|E|,$$

by Equation (14). ■

The following proposition shows that for an edge $e = (v, w) \in E'$, the label sets A_1^v and A_1^w intersect under projections given by e .

Proposition III.8. *For every edge $e = (v, w) \in E'$, we have $\pi_{ev}(A_1^v) \cap \pi_{ew}(A_1^w) \neq \emptyset$.*

Proof: From Proposition III.5, let $i^* \in [n]$ be such that $|b_v(i^*)| \geq \beta$. Note that $i^* \in A_1^v$. Let $j^* = \pi_{ev}(i^*)$. Clearly it suffices to show that there exists an $i' \in A_1^w$ such that $\pi_{ew}(i') = j^*$, as this implies that $j^* \in \pi_{ev}(A_1^v) \cap \pi_{ew}(A_1^w)$.

Recall that since $\mathbf{b} \in \mathcal{H}$, the vector \mathbf{b} satisfies the constraint (6), in particular,

$$\left| \sum_{i \in \pi_{ev}^{-1}(j^*)} b_v(i) \right| = \left| \sum_{i \in \pi_{ew}^{-1}(j^*)} b_w(i) \right|. \quad (19)$$

We show that because $i^* \in A_1^v$, the left-hand side must be large. Therefore the right hand side is also large, from which we conclude that there must exist a coordinate $i' \in \pi_{ew}^{-1}(j^*)$ such that $|b_w(i')|$ is large, and so $i' \in A_1^w$.

Recall from the second structural property in Theorem II.1 that $|\pi_{ev}^{-1}(j^*)| \leq t$. Moreover, since π_{ev} acts injectively on the set A_2^v and since $i^* \in A_2^v$, no index $i \neq i^*$ such that $\pi_{ev}(i) = \pi_{ev}(i^*)$ can belong to A_2^v . Hence, by the triangle inequality, the left-hand side of (19) is at least

$$|b_v(i^*)| - \sum_{\substack{i \in \pi_{ev}^{-1}(j^*) \\ i \neq i^*}} |b_v(i)| \geq \beta - t \cdot \left(\frac{\beta}{4t} \right) = \frac{3\beta}{4}. \quad (20)$$

Combining (19), (20), and the triangle inequality lets us bound the right-hand side of (19) by

$$\begin{aligned} \frac{3\beta}{4} &\leq \left| \sum_{i \in \pi_{ew}^{-1}(j^*)} b_w(i) \right| \\ &\leq \sum_{i \in \pi_{ew}^{-1}(j^*) \cap A_2^w} |b_w(i)| + \sum_{i \in \pi_{ew}^{-1}(j^*) \setminus A_2^w} |b_w(i)| \\ &\leq \sum_{i \in \pi_{ew}^{-1}(j^*) \cap A_2^w} |b_w(i)| + t \frac{\beta}{4t}, \end{aligned} \quad (21)$$

where the last inequality uses the same facts as above. Since π_{ew} acts injectively on A_2^w , there is at most one index $i \in \pi_{ew}^{-1}(j^*)$ that also belongs to A_2^w , meaning that the sum in (21) consists of at most one term. We see that sum must be at least $\beta/2$ and in particular, there is an $i' \in \pi_{ew}^{-1}(j^*)$ such that $|b_w(i')| \geq \beta/2$. We conclude that $i' \in A_1^w$ and $\pi_{ew}(i') = j^* = \pi_{ev}(i^*)$, proving the claim. \blacksquare

Proof of Lemma III.6: By Proposition III.8 and Equation (16) any edge $e = (v, w) \in E'$ is satisfied by the assignment A with probability at least $1/(|A_1^v||A_1^w|) \geq (\varepsilon^4 \beta^4)/256$. Since by Proposition III.7, we have $|E'| \geq (\varepsilon^4/4)|E|$, the expected fraction of satisfied edges is at least $\varepsilon^8 \beta^4/1024$. \blacksquare

IV. THE COMMUTATIVE CASE

Recall that the *commutative* Little Grothendieck problem asks for the norm of a linear operator $\mathcal{F} : L_2 \rightarrow L_1$. In this section we use Theorem III.1 to prove Theorem I.3, the tight hardness result for this problem. We first consider the real case of Theorem I.3, and then the complex case in Section IV-B.

A. The real case

The real case of Theorem I.3 follows easily by combining Theorem III.1 with the following simple lemma.

Lemma IV.1. *For every positive integer n there exists a map $f : \mathbb{R}^n \rightarrow L_1$ with the following properties:*

- For any vector $a \in \mathbb{R}^n$, we have $\|f(a)\|_{L_1} \leq \|a\|_{\ell_2}$.
- For each standard basis vector e_i , we have $\|f(e_i)\|_{L_1} = 1$. If $\|f(a)\|_{L_1} > (\sqrt{2/\pi} + \varepsilon)\|a\|_{\ell_2}$ then $\|a\|_{\ell_4} > (\varepsilon/K)\|a\|_{\ell_2}$, where $K < \infty$ is a universal constant.

This shows that there is an L_1 -valued function f that satisfies the conditions of the real variant of Theorem III.1 for $\tau = \sqrt{2/\pi}$, $\eta = 1$ and $\delta(\varepsilon) = (\varepsilon/K)$. Hence, it is NP-hard to approximate the norm of a linear operator $\mathcal{F} : L_2 \rightarrow L_1(L_1)$ over \mathbb{R} to a factor $\sqrt{2/\pi} + \varepsilon$ for any $\varepsilon > 0$. The real case of Theorem I.3 then follows from the fact that $L_1(L_1)$ is isometrically isomorphic to L_1 .

The proof of Lemma IV.1 uses the following version of the Berry–Esséen Theorem (see for example [32, Chapter 5.2, Theorem 5.16]).

Theorem IV.2 (Berry–Esséen Theorem). *There exists a universal constant $K < \infty$ such that the following holds. Let n be a positive integer and let Z_1, \dots, Z_n be independent centered $\{-1, 1\}$ -valued random variables. Then, for any vector $a \in \mathbb{R}^n$ such that $\|a\|_{\ell_\infty} \leq \varepsilon \|a\|_{\ell_2}$, we have*

$$\mathbf{E} \left[\left| \sum_{i=1}^n a_i Z_i \right| \right] \leq \left(\sqrt{\frac{2}{\pi}} + K\varepsilon \right) \|a\|_{\ell_2}.$$

Proof of Lemma IV.1: Endow $\{-1, 1\}^n$ with the uniform probability measure and define the function $f : \mathbb{R}^n \rightarrow L_1(\{-1, 1\}^n)$ by

$$(f(a))(Z_1, \dots, Z_n) = \sum_{i=1}^n a_i Z_i.$$

The first property follows since

$$\|f(a)\|_{L_1} \leq \|f(a)\|_{L_2} = \left(\mathbf{E} \left[\left| \sum_{i=1}^n a_i Z_i \right|^2 \right] \right)^{1/2} = \|a\|_{\ell_2}.$$

The second property is trivial. The third property follows from Theorem IV.2. Indeed, the theorem implies that if for some $\varepsilon > 0$, we have

$$\|f(a)\|_{L_1} = \mathbf{E} \left[\left| \sum_{i=1}^n a_i Z_i \right| \right] > \left(\sqrt{\frac{2}{\pi}} + \varepsilon \right) \|a\|_{\ell_2},$$

then $\|a\|_{\ell_\infty} > (\varepsilon/K) \|a\|_{\ell_2}$. Since $\|a\|_{\ell_4} \geq \|a\|_{\ell_\infty}$ the last property follows. \blacksquare

B. The complex case

A similar argument to the one above shows the complex case of Theorem I.3. This follows from the following complex analogue of Lemma IV.1.

Lemma IV.3. *For every positive integer n there exists a map $f : \mathbb{C}^n \rightarrow L_1$ with the following properties*

- For any vector $a \in \mathbb{C}^n$, we have $\|f(a)\|_{L_1} \leq \|a\|_{\ell_2}$.
- For each standard basis vector e_i , we have $\|f(e_i)\|_{L_1} = 1$. If $\|f(a)\|_{L_1} > (\sqrt{\pi/4} + \varepsilon) \|a\|_{\ell_2}$ then $\|a\|_{\ell_4} > (\varepsilon^2/K) \|a\|_{\ell_2}$, where $K < \infty$ is a universal constant.

This shows that there is an L_1 -valued function f that satisfies the conditions of Theorem III.1 for $\tau = \sqrt{\pi/4}$, $\eta = 1$ and $\delta(\varepsilon) = (\varepsilon^2/K)$. Hence, it is NP-hard to approximate the norm of a linear operator $\mathcal{F} : L_2 \rightarrow L_1(L_1)$ over \mathbb{C} to a factor $\sqrt{\pi/4} + \varepsilon$ for any $\varepsilon > 0$.

The proof of Lemma IV.3 is based on the following complex analogue the Berry–Esséen Theorem. Since we could not find this precise formulation in the literature we include a proof below for completeness.

Lemma IV.4 (Complex Berry–Esséen Theorem). *There exists a universal constant $K < \infty$ such that the following holds. Let Z_1, \dots, Z_n be independent uniformly distributed random variables over $\{1, i, -1, -i\}$. Then, for any vector $a \in \mathbb{C}^n$ such that $\|a\|_{\ell_\infty} \leq \varepsilon \|a\|_{\ell_2}$, we have*

$$\mathbf{E} \left[\left| \sum_{j=1}^n Z_j a_j \right| \right] \leq \left(\sqrt{\frac{\pi}{4}} + K\sqrt{\varepsilon} \right) \|a\|_{\ell_2}.$$

The proof is based on the following multi-dimensional version of the Berry–Esséen theorem due to Bentkus [33, Theorem 1.1].

Theorem IV.5 (Bentkus). *Let X_1, \dots, X_n be independent \mathbb{R}^d -valued random variables such that $\mathbf{E}[X_j] = 0$ for each $j \in [n]$. Let $S = X_1 + \dots + X_n$ and assume that the covariance matrix of S equals $\mathbb{1}_d$. Let $g \sim \mathcal{N}(0, \mathbb{1}_d)$ be standard Gaussian vector in \mathbb{R}^d with the same covariance matrix as S . Then, for any measurable convex set $A \subseteq \mathbb{R}^d$, we have*

$$|\Pr[S \in A] - \Pr[g \in A]| \leq c(d) \sum_{j=1}^n \mathbf{E}[\|X_j\|_{\ell_2}^3],$$

where $c(d) = O(d^{1/4})$.

We also use the following standard tail bound.

Lemma IV.6 (Hoeffding's inequality [34]). *Let X_1, \dots, X_n be independent real-valued random variables such that for each $i \in [n]$, $X_i \in [a_i, b_i]$ for some $a_i < b_i$. Let $S = X_1 + \dots + X_n$. Then, for any $t > 0$,*

$$\Pr[|S - \mathbf{E}[S]| > t] \leq 2e^{-2t^2 / \sum_{i=1}^n (b_i - a_i)^2}.$$

Proof of Lemma IV.4: Let $a \in \mathbb{C}^n$ be some vector. By homogeneity we may assume that $\|a\|_{\ell_2} = 1$. Set $\varepsilon = \|a\|_{\ell_\infty}$. For each $j \in [n]$ define the random vector $X_j \in \mathbb{R}^2$ by $X_j = \sqrt{2}[\Re(Z_j a_j), \Im(Z_j a_j)]^\top$. Let $S = X_1 + \dots + X_n$, and let $T \geq 1$ be some number to be set later. We have

$$\mathbf{E}[\|S\|_{\ell_2}] = \int_0^\infty \Pr[\|S\|_{\ell_2} > t] dt = \int_0^T \Pr[\|S\|_{\ell_2} > t] dt + \int_T^\infty \Pr[\|S\|_{\ell_2} > t] dt. \quad (22)$$

We now bound each integral separately. Notice that $\mathbf{E}[X_j] = 0$, and $\mathbf{E}[X_j X_j^\top] = |a_j|^2 \mathbf{1}_2$. It follows that the covariance matrix of S equals $\mathbf{1}_2$. If we let $g \sim \mathcal{N}(0, \mathbf{1}_2)$ be a standard Gaussian vector in \mathbb{R}^2 , then it follows from Theorem IV.5 (for $d = 2$) that for any $t > 0$, we have

$$\begin{aligned} |\Pr[\|S\|_{\ell_2} \leq t] - \Pr[\|g\|_{\ell_2} \leq t]| &\leq c \sum_{j=1}^n \mathbf{E}[\|X_j\|_{\ell_2}^3] \\ &\leq c \max_{j \in [n]} \mathbf{E}[\|X_j\|_{\ell_2}] \sum_{k=1}^n \mathbf{E}[\|X_k\|_{\ell_2}^2] \\ &\leq 2\sqrt{2}c\varepsilon. \end{aligned} \quad (23)$$

By (23), the first integral in (22) is at most

$$2\sqrt{2}c\varepsilon T + \mathbf{E}[\|g\|_{\ell_2}] = 2\sqrt{2}c\varepsilon T + \sqrt{\pi/2}, \quad (24)$$

where we used that $\|g\|_{\ell_2}$ is distributed according to a χ_2 distribution.

We now bound the second integral in (22). The first coordinate S_1 is a sum of independent random variables, $\sqrt{2}\Re(Z_j a_j)$, which are centered and have magnitude at most $\sqrt{2}|a_j|$. Similarly, the same holds for S_2 . Lemma IV.6 therefore gives,

$$\begin{aligned} \int_T^\infty \Pr[\|S\|_{\ell_2} > t] dt &\leq \int_T^\infty \Pr[|S_1| > t/\sqrt{2}] dt + \int_T^\infty \Pr[|S_2| > t/\sqrt{2}] dt \\ &\leq 4 \int_T^\infty e^{-t^2/4} dt \\ &\leq 4 \int_T^\infty t e^{-t^2/4} dt \\ &= 8e^{-T^2/4}, \end{aligned} \quad (25)$$

where in the last inequality we used the assumption $T \geq 1$.

Now set $T = \sqrt{4/\varepsilon}$. Combining (22), (24) and (25), we get

$$\mathbf{E}\left[\left\|\sum_{j=1}^n Z_j a_j\right\|\right] = \frac{1}{\sqrt{2}} \mathbf{E}[\|S\|_{\ell_2}] \leq \frac{1}{\sqrt{2}} \left(\sqrt{\frac{\pi}{2}} + 2\sqrt{2}c\varepsilon T + 8e^{-T^2/4} \right) \leq \sqrt{\frac{\pi}{4}} + K\sqrt{\varepsilon}.$$

The proof of Lemma IV.3 is nearly identical to that of Lemma IV.1, now based on Lemma IV.4 and the function $f: \mathbb{C}^n \rightarrow L_1(\{1, i, -1, -i\}^n)$ given by $(f(a))(Z_1, \dots, Z_n) = a_1 Z_1 + \dots + a_n Z_n$, where $\{1, i, -1, -i\}^n$ is endowed with the uniform probability measure. ■

V. THE NON-COMMUTATIVE CASE

In this section we complete the proof of our main theorem (Theorem I.2). The following lemma gives the linear matrix-valued map f mentioned in the introduction that will allow us to conclude.

Lemma V.1. *Let n be a positive integer and let $d = 2^{2n+\lceil n/2 \rceil}$. Then, there exists a linear operator $f : \mathbb{C}^n \rightarrow \mathbb{C}^{d \times d}$ such that for any vector $a \in \mathbb{C}^n$, we have*

$$\|f(a)\|_{S_1} \leq \sqrt{\frac{\|a\|_{\ell_2}^2 + \|a\|_{\ell_4}^2}{2}}.$$

In particular, $\|f(a)\|_{S_1} \leq (\|a\|_{\ell_2} + \|a\|_{\ell_4})/\sqrt{2}$. Moreover, for each basis vector e_i we have $\|f(e_i)\|_{S_1} = 1$.

Theorem I.2 now follows easily by combining the above lemma with Theorem III.1. Indeed, Lemma V.1 shows that the conditions of Theorem III.1 hold for $\tau = 2^{-1/2}$, $\eta = 1$ and $\delta(\varepsilon) = \sqrt{2}\varepsilon$. It is therefore NP-hard to approximate the norm of a linear operator $\mathcal{F} : L_2 \rightarrow L_1(S_1)$ to within a factor $1/\sqrt{2} + \varepsilon$ for any $\varepsilon > 0$. This implies the theorem because $L_1(S_1)$ embeds isometrically into S_1 . To see the last fact, we use the map that takes a matrix-valued function g on a finite measure space U to a block diagonal matrix with blocks proportional to $g(u)$ for $u \in U$ and use the fact that the trace norm of a block diagonal matrix is the average trace norm of the blocks.

The rest of this section is devoted to the proof of Lemma V.1. For a complex vector $a \in \mathbb{C}^n$ let $\Re(a), \Im(a) \in \mathbb{R}^n$ denote its real and imaginary parts, respectively, and define

$$\Lambda(a) = \sqrt{\|\Re(a)\|_{\ell_2}^2 \|\Im(a)\|_{\ell_2}^2 - \langle \Re(a), \Im(a) \rangle^2}. \quad (26)$$

Note that this value is the area of the parallelogram in \mathbb{R}^n generated by the vectors $\Re(a)$ and $\Im(a)$.

Lemma V.2. *Let n be a positive integer and let $d' = 2^{\lceil n/2 \rceil}$. Then, there exists a operator $C : \mathbb{C}^n \rightarrow \mathbb{C}^{d' \times d'}$ such that for any vector $a \in \mathbb{C}^n$, we have*

$$\|C(a)\|_{S_1} = \frac{1}{2} \sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \frac{1}{2} \sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}. \quad (27)$$

Though we will not use it here, let us point out that the map C becomes an isometric embedding if we restrict it to \mathbb{R}^n , since $\Lambda(a) = 0$ for real vectors.

Proof: We begin by defining a set of pairwise anti-commuting matrices as follows. The Pauli matrices are the four Hermitian matrices

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Using these we define $2^{\lceil n/2 \rceil}$ matrices in $\mathbb{C}^{d' \times d'}$ by

$$C_{2j-1} = \underbrace{Z \otimes \cdots \otimes Z}_{j-1 \text{ times}} \otimes X \otimes \underbrace{I \otimes \cdots \otimes I}_{\lceil n/2 \rceil - j \text{ times}}$$

$$C_{2j} = \underbrace{Z \otimes \cdots \otimes Z}_{j-1 \text{ times}} \otimes Y \otimes \underbrace{I \otimes \cdots \otimes I}_{\lceil n/2 \rceil - j \text{ times}},$$

for each $j \in [\lceil n/2 \rceil]$. It is easy to verify that these matrices have trace zero, that they are Hermitian, unitary, and that they pairwise anti-commute. In particular, they satisfy $C_j^2 = I$. For a vector $a \in \mathbb{C}^n$ we define the map C by $C(a) = a_1 C_1 + \cdots + a_n C_n$. Note that for a real vector $x \in \mathbb{R}^n$, the matrix $C(x)$ is Hermitian and that it satisfies $C(x)^2 = \|x\|_{\ell_2}^2 I$. If a real vector $z \in \mathbb{R}^n$ is orthogonal to x then by expanding the definitions of the matrices $C(x)$ and $C(z)$ and using the above properties we find that they anti-commute:

$$\begin{aligned} C(x)C(z) &= \langle x, z \rangle I + \sum_{j \neq k} x_j z_k C_j C_k \\ &= 0 - \sum_{j \neq k} x_j z_k C_k C_j \\ &= -C(z)C(x). \end{aligned}$$

This shows that the matrix $C(x)C(z)$ is skew-Hermitian, which implies that it has purely imaginary eigenvalues. Since this matrix has trace zero and satisfies $C(x)C(z)(C(x)C(z))^* = \|x\|_{\ell_2}^2 \|z\|_{\ell_2}^2 I$, half the eigenvalues equal $i\|x\|_{\ell_2}\|z\|_{\ell_2}$ and the other half equal $-i\|x\|_{\ell_2}\|z\|_{\ell_2}$.

We show that C satisfies (27). Let $x = \Re(a)$ and $y = \Im(a)$ so that $C(a) = C(x) + iC(y)$. Write $y = y^\parallel + y^\perp$ where y^\parallel is parallel to x and y^\perp is orthogonal to x . Then,

$$\begin{aligned} C(a)C(a)^* &= (C(x) + iC(y))(C(x) - iC(y)) \\ &= \|a\|_{\ell_2}^2 I - i(C(x)C(y) - C(y)C(x)) \\ &= \|a\|_{\ell_2}^2 I - 2iC(x)C(y^\perp), \end{aligned}$$

where in the last line we used the fact that $C(y^\parallel)$ commutes with $C(x)$ while $C(y^\perp)$ anti-commutes with $C(x)$. Using what we deduced above for the matrix $C(x)C(y^\perp)$ we see that half of the eigenvalues of $C(a)C(a)^*$ equal $\|a\|_{\ell_2}^2 + 2\|x\|_{\ell_2}\|y^\perp\|_{\ell_2}$ and the other half equal $\|a\|_{\ell_2}^2 - 2\|x\|_{\ell_2}\|y^\perp\|_{\ell_2}$. Hence,

$$\|C(a)\|_{S_1} = \frac{1}{2} \sqrt{\|a\|_{\ell_2}^2 + 2\|x\|_{\ell_2}\|y^\perp\|_{\ell_2}} + \frac{1}{2} \sqrt{\|a\|_{\ell_2}^2 - 2\|x\|_{\ell_2}\|y^\perp\|_{\ell_2}}$$

The claim now follows because $\|x\|_{\ell_2}\|y^\perp\|_{\ell_2}$ is precisely the area of the parallelogram generated by the vectors x and y . ■

We denote the entry-wise product of two vectors $a, b \in \mathbb{C}^n$ by $a \circ b = (a_1 b_1, \dots, a_n b_n)$.

Proposition V.3. *Let ω be a vector chosen uniformly from $\{1, i, -1, -i\}^n$. Then, for any $a \in \mathbb{C}^n$, we have*

$$4\mathbf{E}_\omega[\Lambda(a \circ \omega)^2] = \|a\|_{\ell_2}^4 - \|a\|_{\ell_4}^4.$$

Proof: Fix a vector $a \in \mathbb{C}^n$. Define the random vectors $x_\omega = \Re(a \circ \omega)$ and $y_\omega = \Im(a \circ \omega)$. Then $\Lambda(a \circ \omega)^2 = \|x_\omega\|_{\ell_2}^2 \|y_\omega\|_{\ell_2}^2 - \langle x_\omega, y_\omega \rangle^2$. For each $j \in [n]$ we factor $a_j = \alpha_j e^{i\phi_j}$ and $\omega_j = e^{i\psi_j}$, where $\alpha_j \in \mathbb{R}_+$ and $\phi_j, \psi_j \in [0, 2\pi]$. Note that ψ_1, \dots, ψ_n are independent uniformly distributed random phases in $\{0, \pi/2, \pi, 3\pi/2\}$. Then,

$$\begin{aligned} \|x_\omega\|_{\ell_2}^2 &= \sum_{j=1}^n \alpha_j^2 \cos^2(\phi_j + \psi_j) \\ \|y_\omega\|_{\ell_2}^2 &= \sum_{j=1}^n \alpha_j^2 \sin^2(\phi_j + \psi_j) \\ \langle x_\omega, y_\omega \rangle &= \sum_{j=1}^n \alpha_j^2 \cos(\phi_j + \psi_j) \sin(\phi_j + \psi_j). \end{aligned}$$

With this it is easy to verify that

$$\begin{aligned} \|x_\omega\|_{\ell_2}^2 \|y_\omega\|_{\ell_2}^2 - \langle x_\omega, y_\omega \rangle^2 &= \sum_{j \neq k} \alpha_j^2 \alpha_k^2 \cos^2(\phi_j + \psi_j) \sin^2(\phi_k + \psi_k) - \\ &\quad \sum_{j \neq k} \alpha_j^2 \alpha_k^2 \cos(\phi_j + \psi_j) \sin(\phi_j + \psi_j) \cos(\phi_k + \psi_k) \sin(\phi_k + \psi_k). \end{aligned} \quad (28)$$

By independence of ψ_j and ψ_k when $j \neq k$ and the elementary identities $\mathbf{E}[\cos^2(\phi_j + \psi_k)] = 1/2$, $\mathbf{E}[\sin^2(\phi_j + \psi_k)] = 1/2$ and $\mathbf{E}[\cos(\phi_j + \psi_j) \sin(\phi_j + \psi_j)] = 0$, the expectation of (28) equals

$$\mathbf{E}_\omega[\Lambda(a \circ \omega)^2] = \frac{1}{4} \sum_{j \neq k} a_j^2 a_k^2 = \frac{1}{4} (\|a\|_{\ell_2}^4 - \|a\|_{\ell_4}^4).$$

We remark that in the above proof, it suffices if $\omega \in \{1, i, -1, -i\}^n$ is chosen from a pairwise independent family. Using this in the proof below, allows one to prove Lemma V.1 with a smaller parameter d . ■

Proof of Lemma V.1: Let C be the map given by Lemma V.2. Define the map

$$f(a) = \bigoplus_{\omega} C(a \circ \omega)$$

where ω ranges over $\{1, i, -1, -i\}^n$. By convexity of the square function, Jensen's inequality, and the fact that $\|a \circ \omega\|_{\ell_2} = \|a\|_{\ell_2}$, we have

$$\begin{aligned}
\|f(a)\|_{S_1}^2 &= \left(\mathbf{E}_\omega [\|C(a \circ \omega)\|_{S_1}] \right)^2 \\
&\stackrel{\text{Lemma V.2}}{=} \left(\mathbf{E}_\omega \left[\frac{1}{2} \sqrt{\|a \circ \omega\|_{\ell_2}^2 + 2\Lambda(a \circ \omega)} + \frac{1}{2} \sqrt{\|a \circ \omega\|_{\ell_2}^2 - 2\Lambda(a \circ \omega)} \right] \right)^2 \\
&\leq \mathbf{E}_\omega \left[\left(\frac{1}{2} \sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a \circ \omega)} + \frac{1}{2} \sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a \circ \omega)} \right)^2 \right] \\
&= \frac{\|a\|_{\ell_2}^2}{2} + \frac{1}{2} \mathbf{E}_\omega \left[\sqrt{\|a\|_{\ell_2}^4 - 4\Lambda(a \circ \omega)^2} \right]. \tag{29}
\end{aligned}$$

Concavity of the square-root function, Jensen's inequality and Proposition V.3 gives that the expectation in (29) is at most

$$\left(\|a\|_{\ell_2}^4 - 4\mathbf{E}[\Lambda(a \circ \omega)^2] \right)^{1/2} = \left(\|a\|_{\ell_2}^4 - \|a\|_{\ell_2}^4 + \|a\|_{\ell_4}^4 \right)^{1/2} = \|a\|_{\ell_4}^2.$$

Hence,

$$\|f(a)\|_{S_1} \leq \sqrt{\frac{\|a\|_{\ell_2}^2 + \|a\|_{\ell_4}^2}{2}}.$$

For the second claim observe that for any standard basis vector e_j and $\omega \in \{1, i, -1, -i\}^n$, the vector $e_j \circ \omega$ is either purely real or purely imaginary. This implies $\Lambda(e_j \circ \omega) = 0$. Hence, by Lemma V.2,

$$\|f(e_i)\|_{S_1} = \frac{1}{2} \mathbf{E}_\omega \left[\sqrt{\|e_j \circ \omega\|_{\ell_2}^2 + 2\Lambda(e_j \circ \omega)} + \sqrt{\|e_j \circ \omega\|_{\ell_2}^2 - 2\Lambda(e_j \circ \omega)} \right] = 1.$$

■

A. The real and Hermitian variants

We end this section by showing that our hardness result of Theorem I.2 also holds for two variants of the Little NCG, the real variant and the Hermitian variant. Both variants were introduced (in the context of the ‘‘big’’ NCG) in [17], partly for the purpose of using them in applications. The real variant asks for the operator norm of a linear map \mathcal{F} from \mathbb{R}^n to a space $\mathbb{R}^{d \times d}$ endowed with the Schatten-1 norm; in the Hermitian variant, the linear map is from \mathbb{R}^n to the space $H_d \subseteq \mathbb{C}^{d \times d}$ of Hermitian matrices, again endowed with the Schatten-1 norm. In both cases the operator norm is given by $\|\mathcal{F}\| = \sup_a \|F(a)\|_{S_1}$ with the supremum over real unit vectors a . Both the real and Hermitian variants follow directly by combining the lemma shown below and the real version of Theorem III.1. Let us denote by $\mathbb{S}^{d \times d} \subseteq \mathbb{R}^{d \times d}$ the space of real symmetric matrices.

Lemma V.4. *Let n be a positive integer and let d be as in Lemma V.1. Then, there exists a linear operator $f : \mathbb{R}^n \rightarrow \mathbb{S}^{4d \times 4d}$ satisfying the conditions stated in Lemma V.1 (with $a \in \mathbb{R}^n$).*

The lemma follows by applying the map ρ of the elementary claim below to the restriction of the operator f of Lemma V.1 to \mathbb{R}^n .

Claim V.5. *For every positive integer d there exists a map $\rho : \mathbb{C}^{d \times d} \rightarrow \mathbb{S}^{4d \times 4d}$ such that for any matrix $A \in \mathbb{C}^{d \times d}$, we have $\|\rho(A)\|_{S_1} = \|A\|_{S_1}$. Moreover, ρ is linear over the real numbers, that is, for any $\alpha \in \mathbb{R}$ and $A, B \in \mathbb{C}^{d \times d}$, we have $\rho(\alpha A) = \alpha \rho(A)$ and $\rho(A + B) = \rho(A) + \rho(B)$.*

Proof: The proof follows by combining two standard transformations taking complex matrices to Hermitian matrices and real matrices, respectively. Let $A \in \mathbb{C}^{d \times d}$ be a matrix with singular values $\sigma_1 \geq \dots \geq \sigma_d$. The first transformation is given by $A \mapsto \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}$. By [35, Theorem 7.3.3], the last matrix has eigenvalues $\sigma_1 \geq \dots \geq \sigma_d \geq -\sigma_d \geq \dots \geq -\sigma_1$. Notice that this transformation is linear over the reals since the adjoint is such. Let $B \in \mathbb{C}^{d \times d}$ be a Hermitian matrix with eigenvalues $\lambda_1 \geq \dots \geq \lambda_d$. The second transformation is given by $B \mapsto \begin{bmatrix} \Re(B) & \Im(B) \\ -\Im(B) & \Re(B) \end{bmatrix}$. Then the last matrix is symmetric and by [35, 1.30.P20 (g), p. 71], that matrix has the same eigenvalues as B but with doubled multiplicities, that is, the matrix has eigenvalues $\lambda_1 \geq \lambda_1 \geq \dots \geq \lambda_d \geq \lambda_d$. Notice that this transformation is also linear over the reals. Let ρ be the composition of these maps. Then the matrix $\rho(A)$ has the same singular values as A but with quadrupled multiplicities, which implies that $\|\rho(A)\|_{S_1} = \|A\|_{S_1}$, and ρ is linear over the reals. ■

VI. LITTLE VERSUS BIG GROTHENDIECK THEOREM

For completeness, we include here the well-known relation between the little and big Grothendieck problems. We focus on the non-commutative case; the commutative case is similar and can be found in, e.g., [4, Section 5]. This discussion clarifies how to derive Theorem I.1 from Theorem I.2.

Consider a linear map $\mathcal{F} : \mathbb{C}^n \rightarrow S_1^d$. A standard and easy-to-prove fact is that for two finite-dimensional Banach spaces X, Y , the operator norm of a linear map $\mathcal{G} : X \rightarrow Y$ equals the norm of its adjoint $\mathcal{G}^* : Y^* \rightarrow X^*$. As a result, $\|\mathcal{F}\| = \|\mathcal{F}^*\|$. Notice that since Hilbert space is self-dual and the dual of S_1 is the space S_∞ of matrices endowed with the Schatten- ∞ norm (i.e., the maximum singular value), we have that $\mathcal{F}^* : S_\infty^d \rightarrow \mathbb{C}^n$. In particular,

$$\|\mathcal{F}^*\| = \sup \| \mathcal{F}^*(A) \|_2 ,$$

where the supremum is taken over all A of Schatten- ∞ norm at most 1. Equivalently, since any matrix with Schatten- ∞ norm at most 1 lies in the convex hull of the set of unitary matrices, we could take the supremum over all unitary matrices A . Next, recall that in the NCG problem we are given a bilinear form $T : \mathbb{C}^{d \times d} \times \mathbb{C}^{d \times d} \rightarrow \mathbb{C}$, and asked to compute $\text{OPT}(T) = \sup_{A, B} |T(A, B)|$, where the supremum ranges over unitary matrices. Define the bilinear form $T(A, B) = \langle \mathcal{F}^*(A), \mathcal{F}^*(B) \rangle$. By Cauchy-Schwarz,

$$\text{OPT}(T) = \sup_A \| \mathcal{F}^*(A) \|_2^2 = \| \mathcal{F}^* \|^2 = \| \mathcal{F} \|^2 ,$$

where the supremum is over all unitary A , showing that the Little NCG is a special case of the “big” NCG.

ACKNOWLEDGMENTS

We thank Steve Heilman and Thomas Vidick for early discussions and we thank Gilles Pisier and anonymous referees for helpful comments on an earlier version of this manuscript. J.B. was supported by a Rubicon grant from the Netherlands Organisation for Scientific Research (NWO). O.R. was supported by the Simons Collaboration on Algorithms and Geometry and by the National Science Foundation (NSF) under Grant No. CCF-1320188. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES

- [1] A. Grothendieck, “Résumé de la théorie métrique des produits tensoriels topologiques (French),” *Bol. Soc. Mat. São Paulo*, vol. 8, pp. 1–79, 1953.
- [2] J. Lindenstrauss and A. Pełczyński, “Absolutely summing operators in L_p -spaces and their applications,” *Studia Math.*, vol. 29, pp. 275–326, 1968.
- [3] S. Khot and A. Naor, “Grothendieck-type inequalities in combinatorial optimization,” *Communications on Pure and Applied Mathematics*, vol. 65, no. 7, pp. 992–1035, 2012.
- [4] G. Pisier, “Grothendieck’s theorem, past and present,” *Bull. Amer. Math. Soc.*, vol. 49, no. 2, pp. 237–323, 2012, also available at arXiv:1101.4195.
- [5] U. Haagerup, “A new upper bound for the complex Grothendieck constant,” *Israel J. Math.*, vol. 60, no. 2, pp. 199–224, 1987. [Online]. Available: <http://dx.doi.org/10.1007/BF02790792>
- [6] A. Davie, “Lower bound for K_G ,” 1984, unpublished.
- [7] M. Braverman, K. Makarychev, Y. Makarychev, and A. Naor, “The Grothendieck constant is strictly smaller than Krivine’s bound,” *Forum Math. Pi*, vol. 1, pp. e4, 42, 2013, preliminary version in FOCS’11.
- [8] G. Pisier, “Grothendieck’s theorem for noncommutative C^* -algebras, with an appendix on Grothendieck’s constants,” *J. Funct. Anal.*, vol. 29, no. 3, pp. 397–415, 1978. [Online]. Available: [http://dx.doi.org/10.1016/0022-1236\(78\)90038-1](http://dx.doi.org/10.1016/0022-1236(78)90038-1)
- [9] U. Haagerup, “The Grothendieck inequality for bilinear forms on C^* -algebras,” *Adv. in Math.*, vol. 56, no. 2, pp. 93 – 116, 1985.
- [10] U. Haagerup and T. Itoh, “Grothendieck type norms for bilinear forms on C^* -algebras,” *J. Operator Theory*, vol. 34, no. 2, pp. 263–283, 1995.
- [11] N. Alon and A. Naor, “Approximating the cut-norm via Grothendieck’s inequality,” *SIAM J. Comput.*, vol. 35, no. 4, pp. 787–803 (electronic), 2006, preliminary version in STOC’04. [Online]. Available: <http://dx.doi.org/10.1137/S0097539704441629>
- [12] J.-L. Krivine, “Constantes de Grothendieck et fonctions de type positif sur les sphères,” *Adv. in Math.*, vol. 31, pp. 16–30, 1979.

- [13] P. Raghavendra and D. Steurer, “Towards computing the Grothendieck constant,” in *Proc. 20th ACM-SIAM Symp. Discrete Algorithms (SODA 09)*. SIAM, Philadelphia, 2009, pp. 525–534.
- [14] J. Briët, F. M. de Oliveira Filho, and F. Vallentin, “The Grothendieck problem with rank constraint,” in *Proc. 19th Symp. Mathematical Theory of Networks and Systems (MTNS 10)*, 2010, pp. 111–113.
- [15] G. Kindler, A. Naor, and G. Schechtman, “The UGC hardness threshold of the L_p Grothendieck problem,” *Math. Oper. Res.*, vol. 35, no. 2, pp. 267–283, 2010. [Online]. Available: <http://dx.doi.org/10.1287/moor.1090.0425>
- [16] B. S. Tsirelson, “Quantum analogues of the Bell inequalities. The case of two spatially separated domains,” *J. Soviet Math.*, vol. 36, pp. 557–570, 1987.
- [17] A. Naor, O. Regev, and T. Vidick, “Efficient rounding for the noncommutative Grothendieck inequality,” *Theory of Computing*, vol. 10, no. 11, pp. 257–295, 2014, preliminary version in STOC’13. [Online]. Available: <http://www.theoryofcomputing.org/articles/v010a011>
- [18] O. Regev and T. Vidick, “Quantum XOR games,” *ACM Transactions on Computation Theory (ToCT)*, 2014, to appear.
- [19] A. S. Bandeira, C. Kennedy, and A. Singer, “Approximating the little Grothendieck problem over the orthogonal group,” *CoRR*, vol. abs/1308.5207, 2013. [Online]. Available: <http://arxiv.org/abs/1308.5207>
- [20] J. Håstad, “Some optimal inapproximability results,” *J. ACM*, vol. 48, no. 4, pp. 798–859, 2001. [Online]. Available: <http://dx.doi.org/10.1145/502090.502098>
- [21] S. Khot and R. O’Donnell, “SDP gaps and UGC-hardness for max-cut-gain,” *Theory Comput.*, vol. 5, pp. 83–117, 2009.
- [22] S. Khot, “On the unique games conjecture,” in *Proc. 25th IEEE Conf. Computational Complexity (CCC 10)*. IEEE Press, 2010, pp. 99–121.
- [23] L. Trevisan, “On Khot’s unique games conjecture,” *Bull. Amer. Math. Soc. (N.S.)*, vol. 49, no. 1, pp. 91–111, 2012.
- [24] R. E. Rietz, “A proof of the Grothendieck inequality,” *Israel J. Math.*, vol. 19, pp. 271–276, 1974.
- [25] Y. Nesterov, “Semidefinite relaxation and nonconvex quadratic optimization,” *Optim. Methods Softw.*, vol. 9, no. 1-3, pp. 141–160, 1998. [Online]. Available: <http://dx.doi.org/10.1080/10556789808805690>
- [26] S. Khot and A. Naor, “Approximate kernel clustering,” *Mathematika*, vol. 55, no. 1-2, pp. 129–165, 2009. [Online]. Available: <http://dx.doi.org/10.1112/S002557930000098X>
- [27] B. Barak, F. G. S. L. Brandão, A. W. Harrow, J. Kelner, D. Steurer, and Y. Zhou, “Hypercontractivity, sum-of-squares proofs, and their applications,” in *Proc. 44th ACM Symp. Symposium on Theory of Computing (STOC 12)*. ACM, New York, 2012, pp. 307–326. [Online]. Available: <http://dx.doi.org/10.1145/2213977.2214006>
- [28] V. Guruswami, P. Raghavendra, R. Saket, and Y. Wu, “Bypassing UGC from some optimal geometric inapproximability results,” in *Proc. 23rd ACM-SIAM Symp. Discrete Algorithms (SODA 12)*, 2012, pp. 699–717, eCCC manuscript <http://eccc.hpi-web.de/report/2010/177>.
- [29] S. Khot, G. Kindler, E. Mossel, and R. O’Donnell, “Optimal inapproximability results for MAX-CUT and other 2-variable CSPs?” *SIAM J. Comput.*, vol. 37, no. 1, pp. 319–357, 2007. [Online]. Available: <http://dx.doi.org/10.1137/S0097539705447372>
- [30] U. Feige and G. Schechtman, “On the optimality of the random hyperplane rounding technique for MAX CUT,” *Random Structures Algorithms*, vol. 20, no. 3, pp. 403–440, 2002, probabilistic methods in combinatorial optimization. [Online]. Available: <http://dx.doi.org/10.1002/rsa.10036>
- [31] S. Khot, “Hardness results for coloring 3-colorable 3-uniform hypergraphs,” in *Proc. 43rd IEEE Symp. Foundations of Computer Science (FOCS 02)*. IEEE Press, 2002, pp. 23–32.
- [32] R. O’Donnell, *Analysis of Boolean functions*. Cambridge University Press, 2014.
- [33] V. Bentkus, “A Lyapunov type bound in \mathbb{R}^d ,” *Theory of Probab. Appl.*, vol. 49, no. 2, pp. 311–323, 2005, translated from Russian.
- [34] W. Hoeffding, “Probability inequalities for sums of bounded random variables,” *J. Amer. Statistical Association*, vol. 58, no. 301, pp. 13–30, 1962.
- [35] R. A. Horn and C. R. Johnson, *Matrix analysis*, 2nd ed. Cambridge University Press, Cambridge, 2013.