THEORY OF COMPUTING, Volume 13 (15), 2017, pp. 1–24
www.theoryofcomputing.org

Tight Hardness of the Non-Commutative Grothendieck Problem

Jop Briët* Oded Regev† Rishi Saket

Received February 2, 2016; Revised January 20, 2017; Published December 2, 2017

Abstract: We prove that for any $\varepsilon > 0$ it is NP-hard to approximate the non-commutative Grothendieck problem to within a factor $1/2+\varepsilon$, which matches the approximation ratio of the algorithm of Naor, Regev, and Vidick (STOC'13). Our proof uses an embedding of ℓ_2 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. We also observe that one can obtain a tight NP-hardness result for the *commutative* Little Grothendieck problem; previously, this was only known based on the Unique Games Conjecture (Khot and Naor, Mathematika 2009).

ACM Classification: G.1.6

AMS Classification: 68Q17, 15A60, 32A70, 03D15

Key words and phrases: hardness of approximation, semidefinite programming, Grothendieck inequality

1 Introduction

The subject of this paper, the *non-commutative Grothendieck problem*, has its roots in celebrated work of Grothendieck [11], 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

A conference version of this paper appeared in the Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2015) [8].

*Supported by a Rubicon grant from the Netherlands Organisation for Scientific Research (NWO).

†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.

 DOI: 10.4086/toc.2017.v013a015

became widely recognized after it was revamped by Lindenstrauss and Pełczyński [27]. 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

$$\mathsf{OPT}(M) = \sup_{\alpha_i, \beta_j} \Big| \sum_{i,j=1}^{d} M_{ij} \alpha_i \overline{\beta_j} \Big| , \tag{1.1}$$

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

$$\mathsf{SDP}(M) = \sup_{a_i, b_j} \Big| \sum_{i,j=1}^{d} M_{ij} \langle a_i, b_j \rangle \Big| , \tag{1.2}$$

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 $\mathsf{SDP}(M) \geq \mathsf{OPT}(M)$. Grothendieck's inequality states that there exists a universal constant $K_G^{\mathbb{C}} < \infty$ such that for any positive integer d and any $d \times d$ matrix M, we also have $\mathsf{SDP}(M) \leq K_G^{\mathbb{C}} \, \mathsf{OPT}(M)$. This result found an enormous number of applications both within and far beyond its original scope and we give some examples below (see [23, 32] for extensive surveys). Despite this, finding the optimal value of $K_G^{\mathbb{C}}$ is the only one of six problems posed in [11] that remains unsolved today; the current best upper and lower bounds are 1.4049 [14] and 1.338 [9], respectively. The situation is similar for the real variant of the problem, where all objects involved are over the real numbers. The constant in that case is denoted K_G and is known to be between 1.6769 and 1.7823 (see [6]).

The non-commutative Grothendieck problem, to which we will refer as the *NCG*, is the optimization problem in which we are asked to maximize a given bilinear form over all pairs of 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

$$\mathsf{OPT}(T) = \sup_{A,B} \Big| \sum_{i,j,k,l=1}^{d} T_{ijkl} A_{ij} \overline{B_{kl}} \Big| , \tag{1.3}$$

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 $\mathsf{OPT}(M)$ as in (1.1) 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 (1.3) and the quantity

$$\mathsf{SDP}(T) = \sup_{\vec{A}, \vec{B}} \Big| \sum_{i,j,k,l=1}^{d} T_{ijkl} \langle \vec{A}_{ij}, \vec{B}_{kl} \rangle \Big| , \tag{1.4}$$

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 $K < \infty$

[1] Namely, 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 $\sum_k \langle \vec{A}_{ik}, \vec{A}_{jk} \rangle$.

such that for every positive integer d and array T as above, we have $\mathsf{OPT}(T) \leq \mathsf{SDP}(T) \leq K\,\mathsf{OPT}(T)$, where the first inequality follows immediately from the definition. Over twenty-five years after being posed, the non-trivial content of Grothendieck's conjecture, $\mathsf{SDP}(T) \leq K\,\mathsf{OPT}(T)$, was finally settled in the positive by Pisier [31]. 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 [13] lowered Pisier's original estimate to $K \leq 2$ and this was later shown to be sharp by Haagerup and Itoh [15].

Algorithmic applications. The importance of Grothendieck's inequality to computer science was pointed out by Alon and Naor [1], who placed it in the context of approximation algorithms for combinatorial optimization problems. They observed that computing $\mathsf{SDP}(M)$ 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 [26] to an efficient rounding scheme that turns SDP vectors into a feasible solution for the real Grothendieck problem (1.1) achieving value at least $\mathsf{SDP}(M)/1.78$. It is known that whatever the value of K_G is, there exists an efficient algorithm achieving value at least $(1/K_G - \varepsilon)\,\mathsf{SDP}(M)$ for any constant $\varepsilon > 0$. (This was first shown in [33], but can also be derived from the results in [7] 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 [1], in statistical physics where it gives ground state energies in the spin glass model [25], and in quantum physics where it is related to Bell inequalities [37].

In the same spirit, Naor, Regev, and Vidick [28] recently translated the *non-commutative* Grothendieck inequality into an efficient SDP-based approximation algorithm for the NCG problem (1.3) that achieves value at least $\mathsf{SDP}(T)/2$. They also considered the real variant and a Hermitian variant, for which they gave analogous algorithms achieving value at least $\mathsf{SDP}(T)/2\sqrt{2}$. This in turn implies efficient constant-factor approximation algorithms for a variety of problems, including the Procrustes problem and two robust versions of Principal Component Analysis [28] and quantum XOR games [34]. In a related result, Bandeira et al. [2] 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. The Grothendieck problem contains MAXCUT as a special case (in fact, it is a special case of the "Little Grothendieck Problem," discussed below, in which M is the positive semidefinite Laplacian matrix of a graph [1]). It therefore follows from Håstad's inapproximability result [16] that it is NP-hard to approximate the value (1.1) 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 [24] proved that (1.1) 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 [33] were able to improve this Unique Games hardness to $1/K_G$. (See for instance [20, 36] 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 1.1. *For any constant $\varepsilon > 0$ it is* NP*-hard to approximate the optimum* (1.3) *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 1.1 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 $\mathcal{F} : \mathbb{R}^n \to \ell_1^d$ (where ℓ_p^d denotes $\mathbb{R}^d$ endowed with the ℓ_p norm), 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 this is a special case of (the real version of) Equation (1.1): for any $\mathcal{F}$ there exists a positive semidefinite $d \times d$ matrix M such that $\mathsf{OPT}(M) = \|\mathcal{F}\|^2$; and vice versa, one can also map any such M into a corresponding operator $\mathcal{F}$ (see, e. g., [32] or Section 6). We wish to highlight that for such instances, the constant K_G may be replaced by the smaller value $\pi/2$ [35] and that this value is known to be optimal [11]. Moreover, Nesterov made this algorithmic, namely, he showed an algorithm that approximates $\|\mathcal{F}\|$ as above to within $\sqrt{2/\pi}$ [29]. Finally, Khot and Naor [22], as part of a more general result, showed that this is tight: the Unique-Games-hardness threshold for the Little Grothendieck problem is exactly $\sqrt{2/\pi}$. As an aside, we note that other operator norms, in particular of operators in $\mathbb{R}^n \to \ell_4^d$, have played an important role recently in theoretical computer science (see, e. g., [3]).

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 $\|A\|_{S_1} = d^{-1}\,\mathsf{Tr}\sqrt{A^*A}$. 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 $\mathcal{F} : \mathbb{C}^n \to S_1^d$. This problem is a special case of the NCG where $\mathsf{OPT}(T) = \|\mathcal{F}\|^2$ (see Section 6). In particular, it follows from [13, 28] 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 1.1.

Theorem 1.2. *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 [28] (see Section 5.1). Finally, as we show in the "warm-up" section of this paper (Section 4), we also get a tight NP-hardness result for the *commutative* Little Grothendieck problem, strengthening the unique-games-based result of [22].

Theorem 1.3. *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 [33, 22], uses the machinery of Fourier analysis over the hypercube, influences, or the *majority is stablest* theorem [30]. Our attempts to apply these techniques here failed. Instead, we use a more direct approach similar to that taken in [12] 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 $\mathcal{F} : \mathbb{C}^n \to S_1^d$ 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 (namely, the image of $\mathcal{F}$) in which the unit ball looks somewhat like the intersection of the Euclidean ball of radius $\sqrt{2}$ with the ℓ_∞ ball $[-1,1]^n$ (since with such a unit ball, the norm of a vector a is given by $\max\{\|a\|_2/\sqrt{2}, \|a\|_\infty\}$).

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 [15] (see also [32, 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 (1.4). 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. [21] obtained tight hardness of MAXCUT by restricting the tight integrality gap instances by Feige and Schechtman [10] 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 [37] (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 5.2) 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 5.1). Even though the proof of this fact is simple, we find it surprising; we are not aware of any previous application of such "complex extensions" of Clifford algebras.

Open questions. For the real and Hermitian cases there is a gap of $\sqrt{2}$ between the guarantee of the [28] 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 2, 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 3, we prove hardness of approximation for the problem of computing the norm of a general class of Banach-space-valued functions, closely following [12]. In

Section 4, as a "warm up," we prove Theorem 1.3 using the generic result of Section 3 and straightforward applications of real and complex versions of the Berry-Esséen Theorem. Section 5 contains our main technical contribution, which we use there to finish the proof of our main result (Theorem 1.2).

Acknowledgements. We thank Steve Heilman and Thomas Vidick for early discussions. We also thank Gilles Pisier and anonymous referees for helpful comments on an earlier version of this manuscript.

2 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 probability with respect to a uniformly distributed random edge incident with v. For a finite set U we denote by $\mathbf{E}_{u\in U}[\cdot]$ the expectation with respect to the uniform distribution over U. For a complex number $c \in \mathbb{C}$, we denote its real and imaginary parts by $\Re(c)$ and $\Im(c)$, respectively. All Banach spaces are assumed to be finite-dimensional (so we can equivalently talk about normed spaces). Recall that for Banach spaces X,Y the operator norm of a linear operator $\mathcal{F} : X \to Y$ is given by

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

For a real number $p \ge 1$, 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 $f : U \to \mathbb{C}$ with the norm

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

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

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

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

$$\Sigma = \big((\pi_{ev},\pi_{ew}) : e = (v,w) \in E\big)$$

of pairs of maps both from $[n]$ to $[k]$ associated with the endpoints of the edges in E. Given an *assignment* $A : V \to [n]$, we say that an edge $e = (v,w) \in E$ is *satisfied* if $\pi_{ev}\big(A(v)\big) = \pi_{ew}\big(A(w)\big)$.

The following hardness result for Smooth Label Cover, given in [12],[2] is a slight variant of the original

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

construction due to Khot [19]. The theorem also describes the various structural properties, including smoothness, that are satisfied by the hard instances.

Theorem 2.1. *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*

$$\Pr_{e \sim v}[\pi_{ev}(i) = \pi_{ev}(j)] \leq \gamma. \tag{2.1}$$

  - *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 $(\delta^2/2)|E|$.*

3 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 3.1. *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 $f : \mathbb{C}^n \to X_n$ 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 $\mathcal{F} : L_2 \to L_1(X_n)$ to within a factor greater than $(\tau/\eta) + \varepsilon'$.*

3.1 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 2.1, 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

$$\sum_{i \in \pi_{ev}^{-1}(j)} a_v(i) = \sum_{i \in \pi_{ew}^{-1}(j)} a_w(i)\,, \tag{3.1}$$

where $a_v(i)$ denotes the ith coordinate of the vector a_v. Notice that if an assignment $A : V \to [n]$ satisfies the edge $e = (v, w)$, then the standard basis vectors $a_v = e_{A(v)}$ and $a_w = e_{A(w)}$ satisfy (3.1); indeed, if $\pi_{ev}\big(A(v)\big) = \pi_{ew}\big(A(w)\big) = j'$ then both sides of (3.1) equal 1 if $j = j'$ and equal zero otherwise.

Now let η, τ and f be as in Theorem 3.1. We associate with the Smooth Label Cover instance $(G, [n], [k], \Sigma)$ from above the linear operator $\mathcal{F} : \mathcal{H} \to L_1(V, X_n)$ given by

$$\big(\mathcal{F}(\mathbf{a})\big)(v) = f(a_v)\,. \tag{3.2}$$

The operator $\mathcal{F}$ thus maps a $\mathbb{C}^n$-valued assignment $\mathbf{a} = (a_v)_{v \in V}$ satisfying (3.1) to an X_n-valued assignment given by $f(a_v)$ for each $v \in V$. Theorem 3.1 follows from the following two lemmas, which we prove in Sections 3.2 and 3.3, respectively.

Lemma 3.2 (Completeness)**.** *Suppose that there exists an assignment $A : V \to [n]$ that satisfies all the edges in E. Then, $\|\mathcal{F}\| \geq \eta$.*

Lemma 3.3 (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 3.1. Let $\varepsilon > 0$ be arbitrary, let ζ, γ be as in Lemma 3.3, and let $n = n(\zeta, \gamma)$ and $k = k(\zeta, \gamma)$ be as in Theorem 2.1. We use the reduction described above, which maps a Smooth Label Cover instance $(G, [n], [k], \Sigma)$ to the linear operator $\mathcal{F} : L_2 \to L_1(X_n)$ specified in (3.2). By Lemma 3.2, YES instances are mapped to $\mathcal{F}$ satisfying $\|\mathcal{F}\| \geq \eta$, whereas by Lemma 3.3, 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. □

3.2 Completeness

Here we prove Lemma 3.2.

Proof of Lemma 3.2. Let $A : V \to [n]$ 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 (3.1) for every $e \in E$ and $j \in [n]$, and thus $\mathbf{a}$ lies in the domain $\mathcal{H}$ of $\mathcal{F}$. Moreover, by the second property of f given in Theorem 3.1,

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

Hence, $\|\mathcal{F}\| \geq \eta$. □

3.3 Soundness

Here we prove Lemma 3.3 and show that among the family of operators $\mathcal{F} = \mathcal{F}_{\zeta,\gamma}$ as in (3.2), 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

$$\mathbf{E}_{v \in V}\left[\|f(b_v)\|\right] = \|\mathcal{F}(\mathbf{b})\|_{L_1} \geq \tau + 4\varepsilon . \tag{3.3}$$

The weak expansion property in Theorem 2.1 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 3.1, we will consider set of vertices

$$V_0 = \{v \in V \mid \|b_v\|_{\ell_4} > \delta\varepsilon \;\text{ and }\; \|b_v\|_{\ell_2} \leq 1/\varepsilon\} . \tag{3.4}$$

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

Lemma 3.4. *For $V_0 \subseteq V$ defined as in (3.4), 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 (3.3), we have

$$\sum_{v \in V_0} \|f(b_v)\|_{X_n} + \sum_{v \in V_1} \|f(b_v)\|_{X_n} + \sum_{v \in V_2} \|f(b_v)\|_{X_n} + \sum_{v \in V_3} \|f(b_v)\|_{X_n} \geq (\tau + 4\varepsilon)|V| . \tag{3.5}$$

We bound the four sums on the left-hand side of (3.5) individually. Since (by the first item in Theorem 3.1) 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 (3.5) can be bounded by

$$\sum_{v \in V_0} \|f(b_v)\|_{X_n} \leq |V_0|/\varepsilon . \tag{3.6}$$

Similarly, using the definition of V_1 the second sum in (3.5) is at most $\varepsilon|V|$. Next, from the third property of f in Theorem 3.1, 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 (3.5) is bounded as

$$\begin{aligned}
\sum_{v\in V_2}\|f(b_v)\|_{X_n} &\le (\tau+\varepsilon)\sum_{v\in V_2}\|b_v\|_{\ell_2} \\
&\le (\tau+\varepsilon)|V_2|^{\frac{1}{2}}\Big(\sum_{v\in V_2}\|b_v\|_{\ell_2}^2\Big)^{\frac{1}{2}} && \text{(by Cauchy-Schwartz)} \\
&\le (\tau+\varepsilon)|V|^{\frac{1}{2}}\Big(\sum_{v\in V}\|b_v\|_{\ell_2}^2\Big)^{\frac{1}{2}} \\
&= (\tau+\varepsilon)|V|\,, && (3.7)
\end{aligned}$$

where the last inequality uses $\|\mathbf{b}\|_{L_2}=1$. Finally, the fourth sum in (3.5) is bounded by

$$\begin{aligned}
\sum_{v\in V_3}\|f(b_v)\|_{X_n} &\le \sum_{v\in V_3}\|b_v\|_{\ell_2} && \text{(by the property of } f\text{)} \\
&< \sum_{v\in V_3}\varepsilon\|b_v\|_{\ell_2}^2 \\
&\le \varepsilon\sum_{v\in V}\|b_v\|_{\ell_2}^2 \\
&= \varepsilon|V|\|\mathbf{b}\|_{L_2}^2 = \varepsilon|V|\,. && (3.8)
\end{aligned}$$

Combining the above with Equation (3.5) yields, $|V_0|/\varepsilon \ge \varepsilon|V|$, which proves Lemma 3.4. □

Lemma 3.4 and the weak expansion property implies that the set $E(V_0)$ of edges induced by V_0 has cardinality

$$|E(V_0)| \ge (\varepsilon^4/2)|E|\,. \tag{3.9}$$

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

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

$$\delta^4\varepsilon^4 \le \|b_v\|_{\ell_4}^4 \le \|b_v\|_{\ell_\infty}^2\|b_v\|_{\ell_2}^2 \le \|b_v\|_{\ell_\infty}^2/\varepsilon^2\,, \tag{3.10}$$

giving the claim. □

For the to-be-determined value of ζ let $t=t(\zeta)$ be as in Theorem 2.1 and for each $v\in V_0$ let

$$A_1^v = \Big\{ i\in[n] \,\Big|\, |b_v(i)| \ge \frac{\beta}{4} \Big\} \qquad \text{and} \qquad A_2^v = \Big\{ i\in[n] \,\Big|\, |b_v(i)| \ge \frac{\beta}{4t} \Big\}\,.$$

By Proposition 3.5 these sets are nonempty and clearly $A_1^v\subseteq A_2^v$. Moreover, since $\|b_v\|_{\ell_2}\le 1/\varepsilon$, we have,

$$|A_1^v| \le \frac{16}{\varepsilon^2\beta^2} \qquad \text{and} \qquad |A_2^v| \le \frac{16t^2}{\varepsilon^2\beta^2}\,. \tag{3.11}$$

Now consider a random assignment $A : V_0 \to [n]$ 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 3.6. *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 3.3; 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 3.6. 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,

$$E' = \left\{ e = (v,w) \in E(V_0) \,\middle|\, |\pi_{ev}(A_2^v)| = |A_2^v|, \text{ and } |\pi_{ew}(A_2^w)| = |A_2^w| \right\}. \tag{3.12}$$

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

Proof. Consider any vertex $v \in V_0$. By the smoothness property of Theorem 2.1 and a union bound over all distinct pairs $i, j \in A_v^2$, the fraction of edges $e \in E$ incident on v that *do not* satisfy

$$|\pi_{ev}(A_2^v)| = |A_2^v| \tag{3.13}$$

is at most

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

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 (3.13) is at most

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

Thus,

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

by Equation (3.9). □

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

Proof. From Proposition 3.5, 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 (3.1), in particular,

$$\left| \sum_{i \in \pi_{ev}^{-1}(j^*)} b_v(i) \right| = \left| \sum_{i \in \pi_{ew}^{-1}(j^*)} b_w(i) \right| . \tag{3.14}$$

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 2.1 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 (3.14) is at least

$$|b_v(i^*)| - \sum_{\substack{i \in \pi_{ev}^{-1}(j^*) \\ i \neq i^*}} |b_v(i)| \geq \beta - t \cdot \left(\frac{\beta}{4t} \right) = \frac{3\beta}{4} . \tag{3.15}$$

Combining (3.14), (3.15), and the triangle inequality lets us bound the right-hand side of (3.14) by

$$\begin{aligned} \frac{3\beta}{4} &\leq \left| \sum_{i \in \pi_{ew}^{-1}(j^*)} b_w(i) \right| \\ &\leq \sum_{i \in \pi_{ew}^{-1}(j^*) \cap A_2^w} |b_w(i)| + \sum_{i \in \pi_{ew}^{-1}(j^*) \setminus A_2^w} |b_w(i)| \\ &\leq \sum_{i \in \pi_{ew}^{-1}(j^*) \cap A_2^w} |b_w(i)| + t \frac{\beta}{4t} , \end{aligned} \tag{3.16}$$

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 (3.16) consists of at most one term. We see that sum must is 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. □

Proof of Lemma 3.6. By Proposition 3.8 and Equation (3.11) 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 3.7, we have $|E'| \geq (\varepsilon^4/4)|E|$, the expected fraction of satisfied edges is at least $\varepsilon^8\beta^4/1024$. □

4 The commutative case

Recall that the *commutative* Little Grothendieck problem asks for the norm of a linear operator $\mathcal{F} : L_2 \to L_1$. In this section we use Theorem 3.1 to prove Theorem 1.3, the tight hardness result for this problem. We first consider the real case of Theorem 1.3, and then the complex case in Section 4.2.

4.1 The real case

The real case of Theorem 1.3 follows easily by combining Theorem 3.1 with the following simple lemma.

Lemma 4.1. *For every positive integer n there exists a map $f : \mathbb{R}^n \to L_1$ with the following properties:*

- *For any vector $a \in \mathbb{R}^n$, we have $\|f(a)\|_{L_1} \le \|a\|_{\ell_2}$.*
- *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 3.1 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 $\mathcal{F} : L_2 \to L_1(L_1)$ over $\mathbb{R}$ to a factor $\sqrt{2/\pi}+\varepsilon$ for any $\varepsilon > 0$. The real case of Theorem 1.3 then follows from the fact that $L_1(L_1)$ is isometrically isomorphic to L_1 (i. e., there is a bijective isometry between the two).

The proof of Lemma 4.1 uses the following version of the Berry–Esséen Theorem (see for example [30, Chapter 5.2, Theorem 5.16]).

Theorem 4.2 (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 $\varepsilon > 0$ and for any vector $a \in \mathbb{R}^n$ such that $\|a\|_{\ell_\infty} \le \varepsilon\|a\|_{\ell_2}$, we have*

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

Proof of Lemma 4.1. Endow $\{-1,1\}^n$ with the uniform probability measure and define the function $f : \mathbb{R}^n \to L_1(\{-1,1\}^n)$ by

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

The first property follows since

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

The second property is trivial. The third property follows from Theorem 4.2. Indeed, the theorem implies that if for some $\varepsilon > 0$, we have

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

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

4.2 The complex case

A similar argument to the one above shows the complex case of Theorem 1.3. This follows from the following complex analogue of Lemma 4.1.

Lemma 4.3. *For every positive integer n there exists a map $f : \mathbb{C}^n \to L_1$ with the following properties:*

- *For any vector $a \in \mathbb{C}^n$, we have $\|f(a)\|_{L_1} \le \|a\|_{\ell_2}$.*
- *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 3.1 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 $\mathcal{F} : L_2 \to L_1(L_1)$ over $\mathbb{C}$ to a factor $\sqrt{\pi/4} + \varepsilon$ for any $\varepsilon > 0$. The complex case of Theorem 1.3 then follows as before from the fact that $L_1(L_1)$ is isometrically isomorphic to L_1.

The proof of Lemma 4.3 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 4.4 (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 $\|a\|_{\ell_\infty} \le \varepsilon \|a\|_{\ell_2}$, we have*

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

The proof is based on the following multi-dimensional version of the Berry–Esséen theorem due to Bentkus [4, Theorem 1.1].

Theorem 4.5 (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 $S = X_1 + \cdots + X_n$ and assume that the covariance matrix of S equals $\mathbb{1}_d$. Let $g \sim \mathcal{N}(0, \mathbb{1}_d)$ be a 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*

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

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

We also use the following standard tail bounds.

Lemma 4.6 (Gaussian tail bound [5])**.** *Let $g \sim \mathcal{N}(0, \mathbb{1}_d)$ be a standard Gaussian vector in $\mathbb{R}^d$. Then, for any $t > 0$, we have*

$$\Pr\big[\big|\|g\|_{\ell_2} - \sqrt{d}\big| > t\big] \le 2e^{-t^2/2}.$$

Lemma 4.7 (Hoeffding's inequality [17]). *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 $S = X_1+\cdots+X_n$. Then, for any $t>0$,*

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

Proof of Lemma 4.4. 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 $X_j = \sqrt{2}[\Re(Z_j a_j), \Im(Z_j a_j)]^{\mathsf{T}}$ and note that $\|X_j\|_{\ell_2} = \sqrt{2}|a_j| \le \sqrt{2}\varepsilon$. Let $S = X_1+\cdots+X_n$, and let $T \ge \sqrt{8}$ be some number to be set later. We have

$$\mathbf{E}[\|S\|_{\ell_2}] = \int_0^\infty \Pr[\|S\|_{\ell_2}>t]\,dt = \int_0^T \Pr[\|S\|_{\ell_2}>t]\,dt + \int_T^\infty \Pr[\|S\|_{\ell_2}>t]\,dt\,. \tag{4.1}$$

We now analyze each integral separately. Notice that $\mathbf{E}[X_j]=0$, and $\mathbf{E}[X_jX_j^{\mathsf{T}}] = |a_j|^2\mathbb{1}_2$. It follows that the covariance matrix of S equals $\mathbb{1}_2$. If we let $g \sim \mathcal{N}(0,\mathbb{1}_2)$ be a standard Gaussian vector in $\mathbb{R}^2$, then it follows from Theorem 4.5 (for $d=2$) that for any $t>0$, we have

$$\begin{aligned}\big|\Pr\big[\|S\|_{\ell_2}>t\big] - \Pr\big[\|g\|_{\ell_2}>t\big]\big| &\le c\sum_{j=1}^n \mathbf{E}\big[\|X_j\|_{\ell_2}^3\big]\\ &\le \sqrt{2}c\varepsilon\sum_{j=1}^n \mathbf{E}\big[\|X_j\|_{\ell_2}^2\big]\\ &\le 2\sqrt{2}c\varepsilon\,. \end{aligned}\tag{4.2}$$

Therefore, the first integral in (4.1) satisfies

$$\begin{aligned}&\Big|\int_0^T \Pr\big[\|S\|_{\ell_2}>t\big]\,dt - \int_0^T \Pr\big[\|g\|_{\ell_2}>t\big]\,dt\Big|\\ &\quad\le \int_0^T \Big|\Pr\big[\|S\|_{\ell_2}>t\big] - \Pr\big[\|g\|_{\ell_2}>t\big]\Big|\,dt\\ &\quad\le 2\sqrt{2}c\varepsilon T\,. \end{aligned}\tag{4.3}$$

Since $\|g\|_{\ell_2}$ is distributed according to a χ_2 distribution, we have

$$\int_0^\infty \Pr\big[\|g\|_{\ell_2}>t\big]\,dt = \mathbf{E}[\|g\|_{\ell_2}] = \sqrt{\pi/2}\,. \tag{4.4}$$

Moreover, it follows from Lemma 4.6 (for $d=2$) and our assumption on T that

$$\begin{aligned}\int_T^\infty \Pr\big[\|g\|_{\ell_2}-\sqrt{2}>t-\sqrt{2}\big]\,dt &\le \int_T^\infty 2e^{-(t-\sqrt{2})^2/2}\,dt\\ &\le \int_T^\infty 2te^{-t^2/8}\,dt\\ &= 8e^{-T^2/8}\,, \end{aligned}\tag{4.5}$$

where we used that $(t-\sqrt{2})^2 \geq t^2/4$ for $t \geq \sqrt{8}$. Combining (4.3), (4.4), and (4.5), we obtain that the first integral in (4.1) satisfies

$$\Big| \int_0^T \Pr\big[\|S\|_{\ell_2} > t\big]\, dt - \sqrt{\pi/2}\Big| \leq 2\sqrt{2}c\varepsilon T + 8e^{-T^2/8}\,. \tag{4.6}$$

We now bound the second integral in (4.1), which is clearly nonnegative. 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 4.7 therefore gives,

$$\begin{aligned}
\int_T^\infty \Pr[\|S\|_{\ell_2} > t]\, dt &\leq \int_T^\infty \Pr[|S_1| > t/\sqrt{2}]\, dt + \int_T^\infty \Pr[|S_2| > t/\sqrt{2}]\, dt \\
&\leq 4\int_T^\infty e^{-t^2/8}\, dt \\
&\leq 4\int_T^\infty t e^{-t^2/8}\, dt \\
&= 16 e^{-T^2/8}\,,
\end{aligned} \tag{4.7}$$

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

Now set $T = \sqrt{8/\varepsilon}$. Combining (4.1), (4.6), and (4.7), we get

$$\left| \mathbf{E}\Big[\Big|\sum_{j=1}^n Z_j a_j\Big|\Big] - \sqrt{\frac{\pi}{4}} \right| = \left| \frac{1}{\sqrt{2}}\mathbf{E}\big[\|S\|_{\ell_2}\big] - \sqrt{\frac{\pi}{4}} \right| \leq \frac{1}{\sqrt{2}}\Big(2\sqrt{2}c\varepsilon T + 24e^{-T^2/8}\Big) \leq K\sqrt{\varepsilon}\,. \qquad \square$$

The proof of Lemma 4.3 is nearly identical to that of Lemma 4.1, now based on Lemma 4.4 and the function $f : \mathbb{C}^n \to L_1(\{1,i,-1,-i\}^n)$ given by $\big(f(a)\big)(Z_1,\ldots,Z_n) = a_1 Z_1 + \cdots + a_n Z_n$, where $\{1,i,-1,-i\}^n$ is endowed with the uniform probability measure.

5 The non-commutative case

In this section we complete the proof of our main theorem (Theorem 1.2). The following lemma gives the linear matrix-valued map f mentioned in the introduction.

Lemma 5.1. *Let n be a positive integer and let $d = 2^{2n+\lceil n/2\rceil}$. Then, there exists a linear operator $f : \mathbb{C}^n \to \mathbb{C}^{d\times d}$ 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 1.2 now follows easily by combining the above lemma with Theorem 3.1. Indeed, Lemma 5.1 shows that the conditions of Theorem 3.1 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 $\mathcal{F} : L_2 \to L_1(S_1)$ 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 5.1. For a complex vector $a \in \mathbb{C}^n$ define

$$\Lambda(a) = \sqrt{\|\Re(a)\|_{\ell_2}^2 \|\Im(a)\|_{\ell_2}^2 - \langle \Re(a), \Im(a) \rangle^2}\,. \tag{5.1}$$

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

Lemma 5.2. *Let n be a positive integer and let $d' = 2^{\lceil n/2 \rceil}$. Then, there exists a operator $C : \mathbb{C}^n \to \mathbb{C}^{d' \times d'}$ such that for any vector $a \in \mathbb{C}^n$, we have*

$$\|C(a)\|_{S_1} = \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a)} + \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a)}\,. \tag{5.2}$$

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

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

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 $C(x)C(z)\big(C(x)C(z)\big)^* = \|x\|_{\ell_2}^2\|z\|_{\ell_2}^2 I$, 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 (5.2). Let $x = \Re(a)$ and $y = \Im(a)$ so that $C(a) = C(x) + iC(y)$. Write $y = y^{\|} + y^{\perp}$ where $y^{\|}$ is parallel to x and $y^{\perp}$ is orthogonal to x. Then,

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

where in the last line we used the fact that $C(y^{\|})$ 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 $a \circ b = (a_1b_1, \ldots, a_nb_n)$.

Proposition 5.3. *Let ω be a vector chosen uniformly from $\{1, i, -1, -i\}^n$. Then, for any $a \in \mathbb{C}^n$, we have*

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

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 $\psi_1, \ldots, \psi_n$ 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

$$\begin{aligned} \|x_\omega\|_{\ell_2}^2\|y_\omega\|_{\ell_2}^2 - \langle x_\omega, y_\omega\rangle^2 = &\sum_{j\neq k} \alpha_j^2\alpha_k^2 \cos^2(\phi_j+\psi_j)\sin^2(\phi_k+\psi_k) - \\ &\sum_{j\neq k} \alpha_j^2\alpha_k^2 \cos(\phi_j+\psi_j)\sin(\phi_j+\psi_j)\cos(\phi_k+\psi_k)\sin(\phi_k+\psi_k). \end{aligned} \tag{5.3}$$

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 (5.3) equals

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

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 5.1 with a smaller parameter d.

Proof of Lemma 5.1. Let C be the map given by Lemma 5.2. Define the map

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

where ω ranges over 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

$$\begin{aligned}
\|f(a)\|_{S_1}^2 &= \Big(\mathbf{E}_\omega\left[\|C(a \circ \omega)\|_{S_1}\right]\Big)^2 \\
&\overset{\text{Lemma 5.2}}{=} \left(\mathbf{E}_\omega\left[\frac{1}{2}\sqrt{\|a \circ \omega\|_{\ell_2}^2 + 2\Lambda(a \circ \omega)} + \frac{1}{2}\sqrt{\|a \circ \omega\|_{\ell_2}^2 - 2\Lambda(a \circ \omega)}\right]\right)^2 \\
&\leq \mathbf{E}_\omega\left[\left(\frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 + 2\Lambda(a \circ \omega)} + \frac{1}{2}\sqrt{\|a\|_{\ell_2}^2 - 2\Lambda(a \circ \omega)}\right)^2\right] \\
&= \frac{\|a\|_{\ell_2}^2}{2} + \frac{1}{2}\mathbf{E}_\omega\left[\sqrt{\|a\|_{\ell_2}^4 - 4\Lambda(a \circ \omega)^2}\right].
\end{aligned} \tag{5.4}$$

Concavity of the square-root function, Jensen's inequality and Proposition 5.3 gives that the expectation in (5.4) is at most

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

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 5.2,

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

5.1 The real and Hermitian variants

We end this section by showing that our hardness result of Theorem 1.2 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 [28], 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 3.1. Let us denote by $\mathbb{S}^{d\times d} \subseteq \mathbb{R}^{d\times d}$ the space of real symmetric matrices.

Lemma 5.4. *Let n be a positive integer and let d be as in Lemma 5.1. Then, there exists a linear operator $f : \mathbb{R}^n \to \mathbb{S}^{4d\times 4d}$ satisfying the conditions stated in Lemma 5.1 (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 5.1 to $\mathbb{R}^n$.

Claim 5.5. *For every positive integer d there exists a map $\rho : \mathbb{C}^{d\times d} \to \mathbb{S}^{4d\times 4d}$ 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 $\sigma_1 \geq \cdots \geq \sigma_d$. The first transformation is given by $A \mapsto \left[\begin{smallmatrix} 0 & A \\ A^* & 0 \end{smallmatrix}\right]$. By [18, Theorem 7.3.3], the last matrix has eigenvalues $\sigma_1 \geq \cdots \geq \sigma_d \geq -\sigma_d \geq \cdots \geq -\sigma_1$. 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 $\lambda_1 \geq \cdots \geq \lambda_d$. 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 [18, 1.30.P20 (g), p. 71], that matrix has the same eigenvalues as B but with doubled multiplicities, that is, the matrix has eigenvalues $\lambda_1 \geq \lambda_1 \geq \cdots \geq \lambda_d \geq \lambda_d$. 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. □

6 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., [32, Section 5]. This discussion clarifies how to derive Theorem 1.1 from Theorem 1.2.

Consider a linear map $\mathcal{F} : \mathbb{C}^n \to S_1^d$. A standard and easy-to-prove fact is that for two finite-dimensional Banach spaces X, Y, the operator norm of a linear map $\mathcal{G} : X \to Y$ equals the norm of its adjoint $\mathcal{G}^* : Y^* \to X^*$. 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 $\mathcal{F}^* : S_\infty^d \to \mathbb{C}^n$. 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 $T : \mathbb{C}^{d\times d} \times \mathbb{C}^{d\times d} \to \mathbb{C}$, and asked to compute $\mathsf{OPT}(T) = \sup_{A,B} \left|T(A,B)\right|$, 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,

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

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

References

[1] NOGA ALON AND ASSAF NAOR: Approximating the cut-norm via Grothendieck's inequality. *SIAM J. Comput.*, 35(4):787–803, 2006. Preliminary version in STOC'04. [doi:10.1137/S0097539704441629] 3

[2] AFONSO S. BANDEIRA, CHRISTOPHER KENNEDY, AND AMIT SINGER: Approximating the little Grothendieck problem over the orthogonal group. *Math. Program.*, 160(1-2):433–475, 2016. [doi:10.1007/s10107-016-0993-7, arXiv:1308.5207] 3

[3] BOAZ BARAK, FERNANDO G. S. L. BRANDÃO, ARAM W. HARROW, JONATHAN A. KELNER, DAVID STEURER, AND YUAN ZHOU: Hypercontractivity, sum-of-squares proofs, and their applications. In *Proc. 44th STOC*, pp. 307–326. ACM Press, 2012. [doi:10.1145/2213977.2214006, arXiv:1205.4484] 4

[4] VIDMANTAS BENTKUS: A Lyapunov type bound in $\mathbb{R}^d$. *Theory Probab. Appl.*, 49(2):311–323, 2005. Translated from Russian. [doi:10.1137/S0040585X97981123] 14

[5] STÉPHANE BOUCHERON, GÁBOR LUGOSI, AND PASCAL MASSART: *Concentration Inequalities*. Oxford Univ. Press, 2013. 14

[6] MARK BRAVERMAN, KONSTANTIN MAKARYCHEV, YURY MAKARYCHEV, AND ASSAF NAOR: The Grothendieck constant is strictly smaller than Krivine's bound. *Forum Math. Pi*, 1:e4, 42, 2013. Preliminary version in FOCS'11. [doi:10.1017/fmp.2013.4] 2

[7] JOP BRIËT, FERNANDO MÁRIO DE OLIVEIRA FILHO, AND FRANK VALLENTIN: The Grothendieck problem with rank constraint. In *Proc. 19th Symp. Mathem. Theory of Networks and Systems (MTNS'10)*, pp. 111–113, 2010. MTNS. 3

[8] JOP BRIËT, ODED REGEV, AND RISHI SAKET: Tight hardness of the non-commutative Grothendieck problem. In *Proc. 56th FOCS*, pp. 1108–1122. IEEE Comp. Soc. Press, 2015. [doi:10.1109/FOCS.2015.72] 1

[9] ALEXANDER DAVIE: Lower bound for K_G. Unpublished, 1984. 2

[10] URIEL FEIGE AND GIDEON SCHECHTMAN: On the optimality of the random hyperplane rounding technique for MAX CUT. *Random Structures Algorithms*, 20(3):403–440, 2002. [doi:10.1002/rsa.10036] 5

[11] ALEXANDER GROTHENDIECK: Résumé de la théorie métrique des produits tensoriels topologiques (French). *Bol. Soc. Mat. São Paulo*, 8:1–79, 1953. Available from Instituto de Matemática e Estatística da Universidade de São Paulo. 1, 2, 4

[12] VENKATESAN GURUSWAMI, PRASAD RAGHAVENDRA, RISHI SAKET, AND YI WU: Bypassing UGC from some optimal geometric inapproximability results. *ACM Trans. Algor.*, 12(1):6:1–6:25, 2016. Preliminary version in SODA'12. [doi:10.1145/2737729] 5, 6

[13] UFFE HAAGERUP: The Grothendieck inequality for bilinear forms on C^*-algebras. *Adv. in Math.*, 56(2):93 – 116, 1985. [doi:10.1016/0001-8708(85)90026-X] 3, 4

[14] UFFE HAAGERUP: A new upper bound for the complex Grothendieck constant. *Israel J. Math.*, 60(2):199–224, 1987. [doi:10.1007/BF02790792] 2

[15] UFFE HAAGERUP AND TAKASHI ITOH: Grothendieck type norms for bilinear forms on C^*-algebras. *J. Operator Theory*, 34(2):263–283, 1995. Available from JSTOR. 3, 5

[16] JOHAN HÅSTAD: Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001. Preliminary version in STOC'97. [doi:10.1145/502090.502098] 3

[17] WASSILY HOEFFDING: Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58(301):13–30, 1963. [doi:10.1007/978-1-4612-0865-5_26] 15

[18] ROGER A. HORN AND CHARLES R. JOHNSON: *Matrix Analysis*. Cambridge Univ. Press, 1990. 20

[19] SUBHASH KHOT: Hardness results for coloring 3-colorable 3-uniform hypergraphs. In *Proc. 43rd FOCS*, pp. 23–32. IEEE Comp. Soc. Press, 2002. [doi:10.1109/SFCS.2002.1181879] 7

[20] SUBHASH KHOT: On the unique games conjecture. In *Proc. 25th IEEE Conf. on Computational Complexity (CCC'10)*, pp. 99–121. IEEE Comp. Soc. Press, 2010. [doi:10.1109/CCC.2010.19] 3

[21] 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(1):319–357, 2007. Preliminary version in FOCS'04. [doi:10.1137/S0097539705447372] 5

[22] SUBHASH KHOT AND ASSAF NAOR: Approximate kernel clustering. *Mathematika*, 55(1-2):129–165, 2009. Preliminary version in FOCS'08. [doi:10.1112/S002557930000098X, arXiv:0807.4626] 4, 5

[23] SUBHASH KHOT AND ASSAF NAOR: Grothendieck-type inequalities in combinatorial optimization. *Comm. Pure Appl. Math.*, 65(7):992–1035, 2012. [doi:10.1002/cpa.21398, arXiv:1108.2464] 2

[24] SUBHASH KHOT AND RYAN O'DONNELL: SDP gaps and UGC-hardness for Max-Cut-Gain. *Theory of Computing*, 5(4):83–117, 2009. Preliminary version in FOCS'06. [doi:10.4086/toc.2009.v005a004] 3

[25] GUY KINDLER, ASSAF NAOR, AND GIDEON SCHECHTMAN: The UGC hardness threshold of the L_p Grothendieck problem. *Math. Oper. Res.*, 35(2):267–283, 2010. Preliminary version in SODA'08. [doi:10.1287/moor.1090.0425] 3

[26] JEAN-LOUIS KRIVINE: Constantes de Grothendieck et fonctions de type positif sur les sphères. *Adv. Math.*, 31(1):16–30, 1979. [doi:10.1016/0001-8708(79)90017-3] 3

[27] JORAM LINDENSTRAUSS AND ALEKSANDER PEŁCZYŃSKI: Absolutely summing operators in L_p-spaces and their applications. *Studia Math.*, 29(3):275–326, 1968. Available from EuDML. 2

[28] ASSAF NAOR, ODED REGEV, AND THOMAS VIDICK: Efficient rounding for the noncommutative Grothendieck inequality. *Theory of Computing*, 10(11):257–295, 2014. Preliminary version in STOC'13. [doi:10.4086/toc.2014.v010a011] 3, 4, 5, 20

[29] YURII NESTEROV: Semidefinite relaxation and nonconvex quadratic optimization. *Optim. Methods Softw.*, 9(1-3):141–160, 1998. [doi:10.1080/10556789808805690] 4

[30] RYAN O'DONNELL: *Analysis of Boolean Functions*. Cambridge Univ. Press, 2014. Available at Semantic Scholar. [doi:10.1017/CBO9781139814782] 5, 13

[31] GILLES PISIER: Grothendieck's theorem for noncommutative C^*-algebras, with an appendix on Grothendieck's constants. *J. Funct. Anal.*, 29(3):397–415, 1978. [doi:10.1016/0022-1236(78)90038-1] 3

[32] GILLES PISIER: Grothendieck's theorem, past and present. *Bull. Amer. Math. Soc.*, 49(2):237–323, 2012. [doi:10.1090/S0273-0979-2011-01348-9, arXiv:1101.4195] 2, 4, 5, 20

[33] PRASAD RAGHAVENDRA AND DAVID STEURER: Towards computing the Grothendieck constant. In *Proc. 20th ACM-SIAM Symp. on Discrete Algorithms (SODA'09)*, pp. 525–534. ACM Press, 2009. ACM DL. 3, 5

[34] ODED REGEV AND THOMAS VIDICK: Quantum XOR games. *ACM Trans. Comput. Theory*, 7(4):15:1–15:43, 2015. Preliminary version in CCC'13. [doi:10.1145/2799560, arXiv:1207.4939] 3

[35] RONALD E. RIETZ: A proof of the Grothendieck inequality. *Israel J. Math.*, 19(3):271–276, 1974. [doi:10.1007/BF02757725] 4

[36] LUCA TREVISAN: On Khot's unique games conjecture. *Bull. Amer. Math. Soc. (N.S.)*, 49(1):91–111, 2012. [doi:10.1090/S0273-0979-2011-01361-1] 3

[37] BORIS S. TSIREL'SON: Quantum analogues of the Bell inequalities. The case of two spatially separated domains. *J. Soviet Math.*, 36(4):557–570, 1987. [doi:10.1007/BF01663472] 3, 5

AUTHORS

Jop Briët
Assistant professor
CWI, Amsterdam
The Netherlands
j.briet@cwi.nl
http://homepages.cwi.nl/~jop/

Oded Regev
Professor
Courant Institute of Mathematical Sciences
New York University
New York, N.Y.
regev@cims.nyu.edu
http://www.cims.nyu.edu/~regev/

Rishi Saket
Researcher
IBM Research
Bangalore, India
rissaket@in.ibm.com
http://researcher.ibm.com/researcher/view.php?person=in-rissaket

ABOUT THE AUTHORS

JOP BRIËT graduated from CWI in 2011; his advisor was Harry Buhrman. He enjoys problems at the intersection of theoretical computer science and pure mathematics. After a stint as a yogi during a postdoc at the Courant Institute in New York City, he returned to rock climbing in the flattest country in the world.

ODED REGEV graduated from Tel Aviv University in 2001 under the supervision of Yossi Azar. He spent two years as a postdoc at the Institute for Advanced Study, Princeton, and one year at the University of California, Berkeley. He is currently with the Courant Institute of Mathematical Sciences, and enjoys life in NYC. His research interests include computational and mathematical aspects of lattices, quantum computation, and other topics in theoretical computer science.

RISHI SAKET completed his Ph. D. from Georgia Tech in 2009 under the supervision of Subhash Khot. After post-doctoral stints at CMU, Princeton University, and IBM T. J. Watson, he joined IBM Research, Bangalore, India in 2013 where he is currently a researcher. His interests are in hardness of approximation, approximation algorithms, optimization, learning theory, and related areas of theoretical and applied computer sciences.