

# The Grothendieck constant is strictly smaller than Krivine’s bound

Extended abstract

Mark Braverman  
University of Toronto  
mbraverm@cs.toronto.edu

Konstantin Makarychev  
IBM Research  
konstantin@us.ibm.com

Yury Makarychev  
Toyota Technological Institute at Chicago  
yury@ttic.edu

Assaf Naor  
Courant Institute  
naor@cims.nyu.edu

**Abstract**— The classical Grothendieck constant, denoted  $K_G$ , is equal to the integrality gap of the natural semidefinite relaxation of the problem of computing

$$\max \left\{ \sum_{i=1}^m \sum_{j=1}^n a_{ij} \varepsilon_i \delta_j : \{\varepsilon_i\}_{i=1}^m, \{\delta_j\}_{j=1}^n \subseteq \{-1, 1\} \right\},$$

a generic and well-studied optimization problem with many applications. Krivine proved in 1977 that  $K_G \leq \frac{\pi}{2 \log(1+\sqrt{2})}$  and conjectured that his estimate is sharp. We obtain a sharper Grothendieck inequality, showing that  $K_G < \frac{\pi}{2 \log(1+\sqrt{2})} - \varepsilon_0$  for an explicit constant  $\varepsilon_0 > 0$ . Our main contribution is conceptual: despite dealing with a binary rounding problem, random 2-dimensional projections combined with a careful partition of  $\mathbb{R}^2$  in order to round the projected vectors, beat the random hyperplane technique, contrary to Krivine’s long-standing conjecture.

## 1. INTRODUCTION

In his 1953 Resumé [6], Grothendieck proved a theorem that he called “le théorème fondamental de la théorie métrique des produits tensoriels”. This result is known today as Grothendieck’s inequality. An equivalent formulation of Grothendieck’s inequality, due to Lindenstrauss and Pełczyński [14], states that there exists a universal constant  $K \in (0, \infty)$  such that for every  $m, n \in \mathbb{N}$ , every  $m \times n$  matrix  $(a_{ij})$  with real entries, and every  $m + n$  unit vectors  $x_1, \dots, x_m, y_1, \dots, y_n \in S^{m+n-1}$ , there exist  $\varepsilon_1, \dots, \varepsilon_m, \delta_1, \dots, \delta_n \in \{-1, 1\}$  satisfying

$$\sum_{i=1}^m \sum_{j=1}^n a_{ij} \langle x_i, y_j \rangle \leq K \sum_{i=1}^m \sum_{j=1}^n a_{ij} \varepsilon_i \delta_j. \quad (1)$$

Here  $\langle \cdot, \cdot \rangle$  denotes the standard scalar product on  $\mathbb{R}^{m+n}$ . The infimum over those  $K \in (0, \infty)$  for which (1) holds true is called the Grothendieck constant, and is denoted  $K_G$ .

Grothendieck’s inequality is important to several disciplines, including the geometry of Banach spaces,  $C^*$  algebras, harmonic analysis, operator spaces, quantum mechanics, and computer science. Rather than attempting to

explain the ramifications of Grothendieck’s inequality, we refer to Pisier’s survey [15] and the references therein. The forthcoming survey [9] is devoted to Grothendieck’s inequality in computer science; Section 2 below contains a brief discussion of this topic.

Problem 3 of Grothendieck’s Resumé asks for the determination of the exact value of  $K_G$ . This problem remains open despite major effort by many mathematicians. In fact, even though  $K_G$  occurs in numerous mathematical theorems, and has equivalent interpretations as a key quantity in physics [20], [4] and computer science [1], [17], we currently do not even know what the second digit of  $K_G$  is; the best known bounds [12], [18] are  $K_G \in (1.676, 1.783)$ .

Following the upper bounds on  $K_G$  obtained in [6], [14], [19], progress on this problem halted after a beautiful 1977 theorem of Krivine [12], who proved that

$$K_G \leq \frac{\pi}{2 \log(1 + \sqrt{2})} (= 1.782\dots). \quad (2)$$

One reason for this lack of improvement since 1977 is that Krivine conjectured [12], [11] that his bound is actually the exact value of  $K_G$ . Here we prove that Krivine’s conjecture is false, thus obtaining the best known upper bound on  $K_G$ .

*Theorem 1.1:* There exists  $\varepsilon_0 > 0$  such that

$$K_G < \frac{\pi}{2 \log(1 + \sqrt{2})} - \varepsilon_0.$$

We stress that our proof is effective, and it readily yields a concrete positive lower bound on  $\varepsilon_0$ . We chose not to state an explicit new upper bound on the Grothendieck constant since we know that our estimate is suboptimal. Section 3 below contains a discussion of potential improvements of our bound, based on challenging open problems that conceivably might even lead to an exact evaluation of  $K_G$ .

*Remark 1.1:* There has also been major effort to estimate the complex Grothendieck constant [6], [3], [16]; the best known upper bound in this case is due to Haagerup [8]. We did not investigate this issue here, partly because for complex scalars there is no clean conjectured exact value of the Grothendieck constant in the spirit of Krivine’s conjecture. Nevertheless, it is conceivable that our approach can improve Haagerup’s bound on the complex Grothendieck

M. Braverman was supported in part by an NSERC Discovery Grant.

A. Naor was supported in part by NSF grant CCF-0832795, BSF grant 2006009, and the Packard Foundation.

The full version of this paper is available at <http://arxiv.org/abs/1103.6161>.

constant as well. We leave this research direction open for future investigations.

In our opinion, the interest in the exact value of  $K_G$  does not necessarily arise from the importance of this constant itself, though the reinterpretation of  $K_G$  as a fundamental constant in physics and computer science makes it even more interesting to know at least its first few digits. Rather, we believe that it is very interesting to understand the geometric configuration of unit vectors  $x_1, \dots, x_m, y_1, \dots, y_n \in S^{m+n-1}$  (and matrix  $a_{ij}$ ) which make the inequality (1) “most difficult”. This issue is related to the “rounding problem” in theoretical computer science; see Section 2. With this in mind, Krivine’s conjecture corresponds to a natural geometric intuition about the worst spherical configuration for Grothendieck’s inequality. This geometric picture has been crystalized and cleanly formulated as an extremal analytic/geometric problem due to the works of Haagerup, König, and Tomczak-Jaegermann. We shall now explain this issue, since one of the main conceptual consequences of Theorem 1.1 is that the geometric picture behind Grothendieck’s inequality that was previously believed to be true, is actually false. Along the way, we resolve a conjecture of König [10].

### 1.1. König’s problem

One can reformulate Grothendieck’s inequality using integral operators (see [10]). Given a measure space  $(\Omega, \mu)$  and a kernel  $K \in L_1(\Omega \times \Omega, \mu \times \mu)$ , consider the integral operator  $T_K : L_\infty(\Omega, \mu) \rightarrow L_1(\Omega, \mu)$  induced by  $K$ , i.e.,

$$T_K f(x) \stackrel{\text{def}}{=} \int_{\Omega} f(y) K(x, y) d\mu(y).$$

Grothendieck’s inequality asserts that for every  $f, g \in L_\infty(\Omega, \mu; \ell_2)$ , i.e., two bounded measurable functions with values in Hilbert space,

$$\begin{aligned} & \int_{\Omega} \int_{\Omega} K(x, y) \langle f(x), g(y) \rangle d\mu(x) d\mu(y) \\ & \leq K_G \|T_K\|_{L_\infty(\Omega, \mu) \rightarrow L_1(\Omega, \mu)} \|g\|_{L_\infty(\Omega, \mu; \ell_2)} \|f\|_{L_\infty(\Omega, \mu; \ell_2)}. \end{aligned} \quad (3)$$

König [10], citing unpublished computations of Haagerup, asserts that the assumption  $K_G = \pi / (2 \log(1 + \sqrt{2}))$  suggests that the oscillatory Gaussian kernel  $K : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  given by

$$K(x, y) \stackrel{\text{def}}{=} \exp\left(-\frac{\|x\|_2^2 + \|y\|_2^2}{2}\right) \sin(\langle x, y \rangle) \quad (4)$$

should be extremal for Grothendieck’s inequality in the asymptotic sense, i.e., for  $n \rightarrow \infty$ . In the rest of this paper  $K$  will always stand for the kernel appearing in (4), and the corresponding bilinear form  $B_K : L_\infty(\mathbb{R}^n) \times L_\infty(\mathbb{R}^n) \rightarrow \mathbb{R}$

will be given by

$$B_K(f, g) \stackrel{\text{def}}{=} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x) g(y) K(x, y) dx dy. \quad (5)$$

The above discussion led König to make the following conjecture:

*Conjecture 1.2 (König [10]):* Define  $f_0 : \mathbb{R}^n \rightarrow \{-1, 1\}$  by  $f_0(x_1, \dots, x_n) = \text{sign}(x_1)$ . Then  $B_K(f, g) \leq B_K(f_0, f_0)$  for every  $n \in \mathbb{N}$  and every measurable  $f, g : \mathbb{R}^n \rightarrow \{-1, 1\}$ .

In [10] the following result of König and Tomczak-Jaegermann is proved:

*Proposition 1.2 (König and Tomczak-Jaegermann [10]):* A positive answer to Conjecture 1.2 would imply that  $K_G = \frac{\pi}{2 \log(1 + \sqrt{2})}$ .

Proposition 1.2 itself can be viewed as motivation for Conjecture 1.2, since it is consistent with Haagerup’s work and Krivine’s conjecture. But, there are additional reasons why Conjecture 1.2 is natural. First of all, we know due to Lieb’s work [13] that general Gaussian kernels, when viewed as operators from  $L_p(\mathbb{R}^n)$  to  $L_q(\mathbb{R}^n)$ , have only Gaussian maximizers provided  $p$  and  $q$  satisfy certain conditions. The kernel  $K$  does not fit into Lieb’s framework, since it is the imaginary part of a Gaussian kernel (the Gaussian Fourier transform) rather than an actual Gaussian kernel, and moreover the range  $p = \infty$  and  $q = 1$  is not covered by Lieb’s theorem. Nevertheless, in light of Lieb’s theorem one might expect that maximizers of kernels of this type have a simple structure, which could be viewed as a weak justification of Conjecture 1.2. A much more substantial justification of Conjecture 1.2 is that in [10] König announced an unpublished result that he obtained jointly with Tomczak-Jaegermann asserting that Conjecture 1.2 is true for  $n = 1$ .

*Theorem 1.3:* For every Lebesgue measurable  $f, g : \mathbb{R} \rightarrow \{-1, 1\}$  we have

$$\begin{aligned} & \int_{\mathbb{R}} \int_{\mathbb{R}} f(x) g(y) e^{-\frac{x^2 + y^2}{2}} \sin(xy) dx dy \\ & \leq \int_{\mathbb{R}} \int_{\mathbb{R}} \text{sign}(x) \text{sign}(y) e^{-\frac{x^2 + y^2}{2}} \sin(xy) dx dy \quad (6) \\ & = 2\sqrt{2} \log(1 + \sqrt{2}). \end{aligned}$$

Moreover, equality in (6) is attained only when  $f(x) = g(x) = \text{sign}(x)$  almost everywhere or  $f(x) = g(x) = -\text{sign}(x)$  almost everywhere.

We believe that it is important to have a published proof of Theorem 1.3, and for this reason we prove it in the full version of this paper. Conceivably our proof is similar to the unpublished proof of König and Tomczak-Jaegermann, though they might have found a different explanation of this phenomenon. Since Theorem 1.1 combined with Proposition 1.2 implies that König’s conjecture is false, and as we

shall see it is false already for  $n = 2$ , Theorem 1.3 highlights special behavior of the one dimensional case.

Our proof of Theorem 1.1 starts by disproving König’s conjecture for  $n = 2$ . This is done in Section 4. Obtaining an improved upper bound on the Grothendieck constant requires a substantial amount of additional work that uses the counterexample to Conjecture 1.2. This is carried out in Section 5. The failure of König’s conjecture shows that the situation is more complicated than originally hoped, and in particular that for  $n > 1$  the maximizers of the kernel  $K$  have a truly high-dimensional behavior. This more complicated geometric picture highlights the availability of high dimensional rounding schemes that are more sophisticated (and better) than “hyperplane rounding”. These issues are discussed in Section 2 and Section 3.

## 2. KRIVINE-TYPE ROUNDING SCHEMES AND ALGORITHMIC IMPLICATIONS

Consider the following optimization problem. Given an  $m \times n$  matrix  $A = (a_{ij})$ , compute in polynomial time the value

$$\text{OPT}(A) \stackrel{\text{def}}{=} \max_{\varepsilon_1, \dots, \varepsilon_m, \delta_1, \dots, \delta_n \in \{-1, 1\}} \sum_{i=1}^m \sum_{j=1}^n a_{ij} \varepsilon_i \delta_j. \quad (7)$$

We refer to [1], [9] for a discussion of the combinatorial significance of this problem. It suffices to say here that it relates to the problem of computing efficiently the Cut Norm of a matrix, which is a subroutine in a variety of applications, starting with the pioneering work of Frieze and Kannan [5]. Special choices of matrices  $A$  in (7) lead to specific problems of interest, including efficient construction of Szemerédi partitions [1].

As shown in [1], there is no PTAS for the problem unless  $P = NP$ . But, since the quantity

$$\text{SDP}(A) \stackrel{\text{def}}{=} \max_{x_1, \dots, x_m, y_1, \dots, y_n \in S^{m+n-1}} \sum_{i=1}^m \sum_{j=1}^n a_{ij} \langle x_i, y_j \rangle$$

can be computed in polynomial time with arbitrarily good precision (it is a semidefinite program [7]), Grothendieck’s inequality tells us that the polynomial time algorithm that outputs the number  $\text{SDP}(A)$  is always within a factor of  $K_G$  of  $\text{OPT}(A)$ .

Remarkably, the work of Raghavendra and Steurer [17] shows that  $K_G$  has a complexity theoretic interpretation: no polynomial time algorithm can approximate  $\text{OPT}(A)$  to within a factor smaller than  $K_G$  assuming the Unique Games Conjecture. Note that Raghavendra and Steurer manage to prove this result despite the fact that the value of  $K_G$  is unknown.

Theorem 1.1 yields the first improved upper bound on the Unique Games hardness threshold of the  $\text{OPT}(A)$  computation problem since Krivine’s 1977 bound. As we shall see, what hides behind Theorem 1.1 is also a new

algorithmic method which is of independent interest. To explain this, note that the above discussion dealt with the problem of computing the *number*  $\text{OPT}(A)$ . But it is actually of greater interest to find in polynomial time signs  $\varepsilon_1, \dots, \varepsilon_m, \delta_1, \dots, \delta_n \in \{-1, 1\}$  from among all such  $2^{m+n}$  choices of signs, for which  $\sum_{i=1}^m \sum_{j=1}^n a_{ij} \varepsilon_i \delta_j$  is at least a constant multiple  $\text{OPT}(A)$ . This amounts to a “rounding problem”: we need to find a procedure that, given vectors  $x_1, \dots, x_m, y_1, \dots, y_n \in S^{m+n-1}$ , produces signs  $\varepsilon_1, \dots, \varepsilon_m, \delta_1, \dots, \delta_n \in \{-1, 1\}$  whose existence is ensured by Grothendieck’s inequality (1).

Krivine’s proof of (2) is based on a clever two-step rounding procedure. We shall now describe a generalization of Krivine’s method.

*Definition 2.1 (Krivine rounding scheme):* Fix  $k \in \mathbb{N}$  and assume that we are given two odd measurable functions  $f, g : \mathbb{R}^k \rightarrow \{-1, 1\}$ . Let  $G_1, G_2 \in \mathbb{R}^k$  be independent random vectors that are distributed according to the standard Gaussian measure on  $\mathbb{R}^k$ , i.e., the measure with density  $x \mapsto e^{-\|x\|_2^2/2}/(2\pi)^{k/2}$ . For  $t \in (-1, 1)$  define

$$H_{f,g}(t) \stackrel{\text{def}}{=} \mathbb{E} \left[ f \left( \frac{1}{\sqrt{2}} G_1 \right) g \left( \frac{t}{\sqrt{2}} G_1 + \frac{\sqrt{1-t^2}}{\sqrt{2}} G_2 \right) \right] = \frac{1}{\pi^k (1-t^2)^{k/2}} \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} f(x) g(y) e^{-\frac{\|x\|_2^2 - \|y\|_2^2 + 2t \langle x, y \rangle}{1-t^2}} dx dy. \quad (8)$$

Then  $H_{f,g}$  extends to an analytic function on the strip  $\{z \in \mathbb{C} : \Re(z) \in (-1, 1)\}$ . We shall call  $\{f, g\}$  a Krivine rounding scheme if  $H_{f,g}$  is invertible on a neighborhood of the origin, and if we consider the Taylor expansion

$$H_{f,g}^{-1}(z) = \sum_{j=0}^{\infty} a_{2j+1} z^{2j+1} \quad (9)$$

then there exists  $c = c(f, g) \in (0, \infty)$  satisfying

$$\sum_{j=0}^{\infty} |a_{2j+1}| c^{2j+1} = 1. \quad (10)$$

(Only odd Taylor coefficients appear in (9) since  $H_{f,g}$ , and therefore also  $H_{f,g}^{-1}$ , is odd.)

*Definition 2.2 (Alternating Krivine rounding scheme):* A Krivine rounding scheme  $\{f, g\}$  is called an alternating Krivine rounding scheme if the coefficients  $\{a_{2j+1}\}_{j=0}^{\infty} \subseteq \mathbb{R}$  in (9) satisfy  $\text{sign}(a_{2j+1}) = (-1)^j$  for all  $j \in \mathbb{N} \cup \{0\}$ . Note that in this case equation (10) becomes  $H_{f,g}^{-1}(ic)/i = 1$ , or

$$c(f, g) = \frac{H_{f,g}(i)}{i} \stackrel{(4) \wedge (5) \wedge (8)}{=} \frac{B_K(f, g)}{(\sqrt{2}\pi)^k}. \quad (11)$$

Given a Krivine rounding scheme  $f, g : \mathbb{R}^k \rightarrow \{-1, 1\}$  and  $x_1, \dots, x_m, y_1, \dots, y_n \in S^{m+n-1}$ , the (generalized) Krivine rounding method proceeds via the following two steps.

**Step 1 (preprocessing the vectors).** Consider the Hilbert space

$$\mathcal{H} = \bigoplus_{j=0}^{\infty} (\mathbb{R}^{m+n})^{\otimes(2j+1)}.$$

For  $x \in S^{m+n-1}$  we can then define two vectors  $I(x), J(x) \in \mathcal{H}$  by

$$I(x) \stackrel{\text{def}}{=} \sum_{j=0}^{\infty} |a_{2j+1}|^{1/2} c^{(2j+1)/2} x^{\otimes(2j+1)} \quad (12)$$

and

$$J(x) \stackrel{\text{def}}{=} \sum_{j=0}^{\infty} \text{sign}(a_{2j+1}) |a_{2j+1}|^{1/2} c^{(2j+1)/2} x^{\otimes(2j+1)}, \quad (13)$$

where  $c = c(f, g)$ . The choice of  $c$  was made in order to ensure that  $I(x)$  and  $J(x)$  are unit vectors in  $\mathcal{H}$ . Moreover, the definitions (12) and (13) were made so that the following identity holds:

$$\langle I(x), J(y) \rangle_{\mathcal{H}} = H_{f,g}^{-1}(c\langle x, y \rangle), \quad (14)$$

for all  $x, y \in S^{m+n-1}$ . The preprocessing step of the Krivine rounding method transforms the initial unit vectors  $\{x_r\}_{r=1}^m, \{y_s\}_{s=1}^n \subseteq S^{m+n-1}$  to vectors  $\{u_r\}_{r=1}^m, \{v_s\}_{s=1}^n \subseteq S^{m+n-1}$  satisfying the identities

$$\langle u_r, v_s \rangle = \langle I(x_r), J(y_s) \rangle_{\mathcal{H}} \stackrel{(14)}{=} H_{f,g}^{-1}(c\langle x_r, y_s \rangle), \quad (15)$$

for all  $r \in \{1, \dots, m\}$  and  $s \in \{1, \dots, n\}$ . As explained in [1], these new vectors can be computed efficiently provided  $H_{f,g}^{-1}$  can be computed efficiently; this simply amounts to computing a Cholesky decomposition.

**Step 2 (random projection).** Let  $G : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^k$  be a random  $k \times (m+n)$  matrix whose entries are i.i.d. standard Gaussian random variables. Define random signs  $\sigma_1, \dots, \sigma_m, \tau_1, \dots, \tau_n \in \{-1, 1\}$  by

$$\sigma_r \stackrel{\text{def}}{=} f\left(\frac{1}{\sqrt{2}} G u_r\right) \text{ and } \tau_s \stackrel{\text{def}}{=} g\left(\frac{1}{\sqrt{2}} G v_s\right), \quad (16)$$

for all  $r \in \{1, \dots, m\}$  and  $s \in \{1, \dots, n\}$ .

Having obtained the random signs  $\sigma_1, \dots, \sigma_m, \tau_1, \dots, \tau_n \in \{-1, 1\}$  as in (16), for every  $m \times n$  matrix  $(a_{rs})$  we have

$$\begin{aligned} \max_{\substack{\varepsilon_1, \dots, \varepsilon_m \in \{-1, 1\} \\ \delta_1, \dots, \delta_n \in \{-1, 1\}}} \sum_{r=1}^m \sum_{s=1}^n a_{rs} \varepsilon_r \delta_s &\geq \mathbb{E} \left[ \sum_{r=1}^m \sum_{s=1}^n a_{rs} \sigma_r \tau_s \right] \\ &\stackrel{(\clubsuit)}{=} \mathbb{E} \left[ \sum_{r=1}^m \sum_{s=1}^n a_{rs} H_{f,g}(\langle u_r, v_s \rangle) \right] \\ &\stackrel{(15)}{=} c(f, g) \sum_{r=1}^m \sum_{s=1}^n a_{rs} \langle x_r, y_s \rangle, \end{aligned}$$

where  $(\clubsuit)$  follows by rotation invariance from (16) and (8). We have thus proved the following corollary, which yields

a systematic way to bound the Grothendieck constant from above.

*Corollary 2.3:* Assume that  $f, g : \mathbb{R}^k \rightarrow \{-1, 1\}$  is a Krivine rounding scheme. Then

$$K_G \leq \frac{1}{c(f, g)}.$$

Krivine's proof of (2) corresponds to Corollary 2.3 when  $k = 1$  and  $f(x) = g(x) = \text{sign}(x)$ . In this case  $\{f, g\}$  is an alternating Krivine rounding scheme with  $H_{f,g}(t) = \frac{2}{\pi} \arcsin(t)$  (Grothendieck's identity). By (11) we have  $c(f, g) = \frac{2}{\pi} \arcsin(i) = \frac{2}{\pi} \log(1 + \sqrt{2})$ , so that Corollary 2.3 does indeed correspond to Krivine's bound (2).

One might expect that, since we want to round vectors  $x_1, \dots, x_m, y_1, \dots, y_n \in S^{m+n-1}$  to signs  $\varepsilon_1, \dots, \varepsilon_m, \delta_1, \dots, \delta_n \in \{-1, 1\}$ , the best possible Krivine rounding scheme occurs when  $k = 1$  and  $f(x) = g(x) = \text{sign}(x)$ . This is the intuition leading to König's conjecture. The following simple corollary of Theorem 1.3 says that among all *one dimensional* Krivine rounding schemes  $f, g : \mathbb{R} \rightarrow \{-1, 1\}$  we indeed have  $c(f, g) \leq c(\text{sign}, \text{sign})$ , so it does not pay off to take partitions of  $\mathbb{R}$  which are more complicated than the half-line partitions.

*Lemma 2.4:* Let  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  be a Krivine rounding scheme. Then  $c(f, g) \leq \frac{2}{\pi} \log(1 + \sqrt{2})$ .

*Proof:* Denote  $c = c(f, g)$  and assume for contradiction that  $c > \frac{2}{\pi} \log(1 + \sqrt{2})$ . Let  $r$  be the radius of convergence of the power series of  $H_{f,g}^{-1}$  given in (9). Due to (10) we know that  $r \geq c > \frac{2}{\pi} \log(1 + \sqrt{2})$ . Denote

$$\alpha \stackrel{\text{def}}{=} \frac{H_{f,g}(i)}{i} \stackrel{(4) \wedge (5) \wedge (8)}{=} \frac{B_K(f, g)}{\pi \sqrt{2}}. \quad (17)$$

By Theorem 1.3 we have  $|\alpha| \leq \frac{2}{\pi} \log(1 + \sqrt{2}) < r$ , and therefore  $H_{f,g}^{-1}$  is well defined at the point  $i\alpha \in \mathbb{C}$ . Thus,

$$\begin{aligned} 1 &\stackrel{(17)}{=} \frac{H_{f,g}^{-1}(i\alpha)}{i} \stackrel{(9)}{=} \sum_{j=0}^{\infty} (-1)^j a_{2j+1} \alpha^{2j+1} \\ &\leq \sum_{j=0}^{\infty} |a_{2j+1}| \cdot |\alpha|^{2j+1}. \end{aligned}$$

By the definition of  $c$  in (10) we deduce that  $c \leq |\alpha| \leq \frac{2}{\pi} \log(1 + \sqrt{2})$ , as required.  $\blacksquare$

The conceptual message behind Theorem 1.1 is that, despite the above satisfactory state of affairs in the one dimensional case, it does pay off to use more complicated higher dimensional partitions. Specifically, our proof of Theorem 1.1 uses the following rounding procedure. Let  $c, p \in (0, 1)$  be small enough absolute constants. Given  $\{x_r\}_{r=1}^m, \{y_s\}_{s=1}^n \subseteq S^{m+n-1}$ , we preprocess them to obtain new vectors  $\{u_r = u_r(p, c)\}_{r=1}^m, \{v_s = v_s(p, c)\}_{s=1}^n \subseteq S^{m+n-1}$ . Due to certain technical complications, these new vectors are obtained via a procedure that is similar to the preprocessing step (Step 1) described above,

but is not identical to it. We refer to Section 5 for a precise description of the preprocessing step that we use (we conjecture that this complication is unnecessary; see Conjecture 5.5). Once the new vectors  $\{u_r\}_{r=1}^m, \{v_s\}_{s=1}^n \subseteq S^{m+n-1}$  have been constructed, we take an  $2 \times (m+n)$  matrix  $G$  with entries that are i.i.d. standard Gaussian random variables, and we consider the random vectors  $\{Gu_r = ((Gu_r)_1, (Gu_r)_2)\}_{r=1}^m, \{Gv_s = ((Gv_s)_1, (Gv_s)_2)\}_{s=1}^n \subseteq \mathbb{R}^2$ . Having thus obtained new vectors in  $\mathbb{R}^2$ , with probability  $(1-p)$  we “round” our initial vectors to the signs  $\{\text{sign}((Gu_r)_2)\}_{r=1}^m, \{\text{sign}((Gv_s)_2)\}_{s=1}^n \subseteq \mathbb{R}$ , while with probability  $p$  we round  $x_r$  to  $+1$  if

$$(Gu_r)_2 \geq c(((Gu_r)_1)^5 - 10((Gu_r)_1)^3 + 15(Gu_r)_1). \quad (18)$$

and we round  $x_r$  to  $-1$  if

$$(Gu_r)_2 < c(((Gu_r)_1)^5 - 10((Gu_r)_1)^3 + 15(Gu_r)_1). \quad (19)$$

For concreteness, at this juncture it suffices to describe our rounding procedure without explaining how it was derived — the origin of the fifth degree polynomial appearing in (18) and (19) will become clear in Section 4 and Section 5. The rounding procedure for  $y_s$  is identical to (18) and (19), with  $(Gv_s)_1, (Gv_s)_2$  replacing  $(Gu_r)_1, (Gu_r)_2$ , respectively.

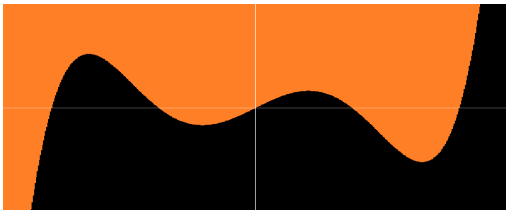


Figure 1. The rounding procedure used in the proof of Theorem 1.1 relies on the partition of  $\mathbb{R}^2$  depicted above. After a preprocessing step, high dimensional vectors are projected randomly onto  $\mathbb{R}^2$  using a matrix with i.i.d. standard Gaussian entries. With a certain fixed probability, if the projected vector falls above the graph  $y = c(x^5 - 10x^3 + 15x)$  then it is assigned the value  $+1$ , and otherwise it is assigned the value  $-1$ .

### 3. THE TIGER PARTITION AND DIRECTIONS FOR FUTURE RESEARCH

The partition of the plane described in Figure 1 leads to a proof of Theorem 1.1, but it is not the optimal partition for this purpose. It makes more sense to use the partitions corresponding to maximizers  $f_{\max}, g_{\max} : \mathbb{R}^2 \rightarrow \{-1, 1\}$  of Krivine’s bilinear form  $B_K$  as defined in (5), i.e.,

$$B_K(f_{\max}, g_{\max}) = \max_{f, g: \mathbb{R}^2 \rightarrow \{-1, 1\}} B_K(f, g) = \max_{f, g: \mathbb{R}^2 \rightarrow \{-1, 1\}} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} f(x)g(y) e^{-\frac{\|x\|_2^2 + \|y\|_2^2}{2}} \sin(\langle x, y \rangle) dx dy. \quad (20)$$

A straightforward weak compactness argument shows that the maximum in (20) is indeed attained (see Section 4).

Given  $f : \mathbb{R}^2 \rightarrow \{-1, 1\}$  define  $\sigma(f) : \mathbb{R}^2 \rightarrow \{-1, 1\}$  by

$$\sigma(f)(y) \stackrel{\text{def}}{=} \text{sign} \left( \int_{\mathbb{R}^2} f(x) e^{-\|x\|_2^2/2} \sin(\langle x, y \rangle) dx \right).$$

Then

$$\sigma(f_{\max}) = g_{\max} \quad \text{and} \quad \sigma(g_{\max}) = f_{\max}. \quad (21)$$

Given  $f : \mathbb{R}^2 \rightarrow \{-1, 1\}$  we can then hope to approach  $f_{\max}$  by considering the iterates  $\{\sigma^{2j}(f)\}_{j=1}^{\infty}$ . If these iterates converge to  $f_{\infty}$  then the pair of functions  $\{f_{\infty}, \sigma(f_{\infty})\}$  would satisfy the equations (21). One can easily check that  $\sigma(f_0) = f_0$  when  $f_0 : \mathbb{R}^2 \rightarrow \{-1, 1\}$  is given by  $f_0(x_1, x_2) = \text{sign}(x_2)$ . But, we have experimentally applied the above iteration procedure to a variety of initial functions  $f \neq f_0$  (both deterministic and random choices), and in all cases the numerical computations suggest that the iterates  $\{\sigma^{2j}(f)\}_{j=1}^{\infty}$  converge to the function  $f_{\infty}$  that is depicted in Figure 2 and Figure 3 (the corresponding function  $g_{\infty} = \sigma(f_{\infty})$  is different from  $f_{\infty}$ , but has a similar structure).



Figure 2. The “tiger partition”: a depiction of the limiting function  $f_{\infty}$  restricted to the square  $[-7, 7] \times [-7, 7] \subseteq \mathbb{R}^2$ , based on numerical computations. The two shaded regions correspond to the points where  $f_{\infty}$  takes the values  $+1$  and  $-1$ .



Figure 3. A zoomed-out view of the tiger partition: a depiction of the limiting function  $f_{\infty}$  restricted to the square  $[-20, 20] \times [-20, 20] \subseteq \mathbb{R}^2$ , based on numerical computations.

*Question 3.1:* Find an analytic description of the function  $f_\infty$  from Figure 2 and Figure 3. Our numerical computations suggest that the iterates  $\{\sigma^{2j}(f)\}_{j=1}^\infty$  converge to  $f_\infty$  for (almost?) all initial data  $f : \mathbb{R}^2 \rightarrow \{-1, 1\}$ . Can this statement be made rigorous? If so, is it the case the  $\{f_\infty, \sigma(f_\infty)\}$  are maximizers of the bilinear form  $B_K$ ? We conjecture that the answer to this question is positive.

*Question 3.2:* Analogously to the above planar computations, can one find an analytic description of the maximizers  $f_{\max}^{(n)}, g_{\max}^{(n)} : \mathbb{R}^n \rightarrow \{-1, 1\}$  of the  $n$ -dimensional version of König's bilinear form  $B_K$ ? If so, does  $\{f_{\max}^{(n)}, g_{\max}^{(n)}\}$  form an alternating Krivine rounding scheme (recall Definition 2.2)?

We do not have sufficient data to conjecture whether the answer to Question 3.2 is positive or negative. But, we note that if  $\{f_{\max}^{(n)}, g_{\max}^{(n)}\}$  were an alternating Krivine rounding scheme then

$$K_G = \sup_{n \in \mathbb{N}} \frac{(\sqrt{2}\pi)^n}{B_K(f_{\max}^{(n)}, g_{\max}^{(n)})} = \sup_{n \in \mathbb{N}} \frac{(\sqrt{2}\pi)^n}{\|T_K\|_{L_\infty(\mathbb{R}^n) \rightarrow L_1(\mathbb{R}^n)}}. \quad (22)$$

Indeed, assuming that  $\{f_{\max}^{(n)}, g_{\max}^{(n)}\}$  is an alternating Krivine rounding scheme the upper bound in (22) follows from Corollary 2.3 and the identity (11). For the reverse inequality in (22) we proceed as in [10]. Using (3) with  $f, g : \mathbb{R}^n \rightarrow S^{n-1}$  given by  $f(x) = g(x) = x/\|x\|_2$ , we see that

$$K_G \geq \frac{\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} K(x, y) \frac{\langle x, y \rangle}{\|x\|_2 \|y\|_2} dx dy}{\|T_K\|_{L_\infty(\mathbb{R}^n) \rightarrow L_1(\mathbb{R}^n)}}, \quad (23)$$

and we conclude that (22) is true since by equation (2.3) in [10] the integral in the numerator of (23) equals  $2^{n/2} \pi^n (1 - 1/n + O(1/n^2))$ .

#### 4. A COUNTEREXAMPLE TO KÖNIG'S CONJECTURE

In this section, we present a counterexample to König's conjecture. We construct a pair of functions  $f, g : \mathbb{R}^2 \rightarrow \{-1, 1\}$  such that  $B_K(f, g) > B_K(f_0, g_0)$ . In the conference version of this paper, we omit the proof of this result.

In our construction, we use Hermite polynomials (see [2, Sec. 6.1]). We let  $\{h_m : \mathbb{R} \rightarrow \mathbb{R}\}_{m=0}^\infty$  denote the sequence of Hermite polynomials normalized so that they form an orthonormal basis with respect to the measure on  $\mathbb{R}$  whose density is  $x \mapsto e^{-x^2}$ . Explicitly,

$$h_m(x) \stackrel{\text{def}}{=} \frac{(-1)^m}{\sqrt{2^m m! \sqrt{\pi}}} \cdot e^{x^2} \frac{d^m}{dx^m} (e^{-x^2}), \quad (24)$$

so that  $\int_{\mathbb{R}} h_m(x) h_k(x) e^{-x^2} dx = \delta_{mk}$ . We consider the fifth Hermite polynomial  $h_5$

$$h_5(x) = \frac{4x^5 - 20x^3 + 15x}{2\sqrt[4]{\pi}\sqrt{15}}.$$

(we discuss the reason why we consider  $h_5$  in the full version of this paper; see also Remark 4.1).

For  $\eta \in (0, 1)$  let  $f_\eta : \mathbb{R}^2 \rightarrow \{-1, 1\}$  be given by

$$f_\eta(x_1, x_2) \stackrel{\text{def}}{=} \begin{cases} 1 & x_2 \geq \eta h_5(x_1), \\ -1 & x_2 < \eta h_5(x_1). \end{cases} \quad (25)$$

Note that since  $h_5$  is odd, so is  $f_\eta$  (almost surely). For  $z \in \mathbb{C}$  with  $|\Re(z)| < 1$  we define

$$H_\eta(z) \stackrel{\text{def}}{=} \frac{\pi}{2} H_{f_\eta, f_\eta}(z) = \frac{\int_{\mathbb{R}^2 \times \mathbb{R}^2} f_\eta(x) f_\eta(y) e^{\frac{-\|x\|_2^2 - \|y\|_2^2 + 2z \langle x, y \rangle}{1 - z^2}} dx dy}{2\pi(1 - z^2)}. \quad (26)$$

*Lemma 4.1:*  $H_\eta$  is analytic on the strip

$$\mathbb{S} \stackrel{\text{def}}{=} \{z \in \mathbb{C} : |\Re(z)| < 1\}. \quad (27)$$

Moreover, for all  $a + bi \in \mathbb{S}$  we have

$$|H_\eta(a + bi)| \leq \frac{\pi((1+a)^2 + b^2)((1-a)^2 + b^2)}{2(1-a^2)\sqrt{(1-a^2)^2 + b^4 + 2(1+a^2)b^2}}. \quad (28)$$

*Lemma 4.2:* For every  $z \in \mathbb{C}$  with  $|\Re(z)| < 1$  we have  $H_0(z) = \arcsin(z)$ .

*Theorem 4.3:* There exists  $\eta_0 > 0$  such that for all  $\eta \in (0, \eta_0)$  we have

$$\frac{H_\eta(i)}{i} \in \left(\log(1 + \sqrt{2}), \infty\right).$$

Theorem 4.3 implies that the answer to König's problem is negative. Indeed,

$$\begin{aligned} \frac{H_\eta(i)}{i} &\stackrel{(26)}{=} \frac{1}{4\pi} \int_{\mathbb{R}^2 \times \mathbb{R}^2} f_\eta(x) f_\eta(y) e^{-\frac{\|x\|_2^2 + \|y\|_2^2}{2}} \sin(\langle x, y \rangle) dx dy \\ &= \frac{B_K(f_\eta, f_\eta)}{4\pi}. \end{aligned}$$

Since  $\arcsin(i) = i \log(1 + \sqrt{2})$ , it follows from Lemma 4.2 and Theorem 4.3 that for every  $\eta \in (0, \eta_0)$  we have  $B_K(f_\eta, f_\eta) > B_K(f_0, f_0)$ . Since  $f_0(x_1, x_2) = \text{sign}(x_2)$ , the claimed negative answer to König's problem follows.

In the proof of Theorem 4.3 presented in the full version of the paper, we consider the function  $\varphi(\eta) = 4\pi H_\eta(i)/i$  and show that  $\varphi''(0) = 0$  and  $\varphi''''(0) = 38400\sqrt{2}$ . We conclude that  $\varphi(\eta) = \varphi(0) + 1600\sqrt{2}\eta^4 + O(\eta^6)$  as  $\eta \rightarrow 0$ . Therefore,

$$\frac{H_\eta(i)}{i} = \frac{\varphi(\eta)}{4\pi} > \frac{\varphi(0)}{4\pi} = \log(1 + \sqrt{2}),$$

when  $\eta$  is small enough.

*Remark 4.1:* Clearly, we did not arrive at the above construction by guessing that the fifth Hermite polynomial  $h_5$  is the correct choice in (25). We arrived at this choice as the simplest member of a general family of ways to perturb the function  $(x_1, x_2) \mapsto \text{sign}(x_2)$ . We discuss our choice in the full version of this paper.

## 5. PROOF THAT $K_G < \frac{\pi}{2 \log(1+\sqrt{2})}$

We will fix from now on some  $\eta \in (0, \eta_0)$ , where  $\eta_0$  is as in Theorem 4.3. For  $p \in [0, 1]$  define

$$F_p \stackrel{\text{def}}{=} (1-p)H_0 + pH_\eta,$$

where  $H_\eta$  is as in (26). In what follows we will denote the unit disc in  $\mathbb{C}$  by

$$\mathbb{D} \stackrel{\text{def}}{=} \{z \in \mathbb{C} : |z| < 1\}.$$

*Theorem 5.1:* The exists  $p_0 > 0$  such that for all  $p \in (0, p_0)$  we have  $F_p(\mathbb{S}) \supseteq \frac{9}{10}\mathbb{D}$  and  $F_p^{-1}$  is well defined and analytic on  $\frac{9}{10}\mathbb{D}$ . Moreover, if we write  $F_p^{-1}(z) = \sum_{k=1}^{\infty} a_k(p)z^k$  then there exists  $\gamma = \gamma_p \in [0, \infty)$  satisfying

$$\sum_{k=1}^{\infty} |a_k(p)|\gamma^k = 1, \quad (29)$$

and

$$\gamma > \log(1 + \sqrt{2}) = 0.88137\dots \quad (30)$$

Assuming Theorem 5.1 for the moment, we will now deduce Theorem 1.1.

*Proof of Theorem 1.1:* Fix  $p \in (0, p_0)$  and let  $\gamma > 0$  be the constant from Theorem 5.1. Due to (29),  $\sum_{k=1}^{\infty} a_k(p)\gamma^k$  converges absolutely, and therefore  $F_p^{-1}$  is analytic and well defined on  $\gamma\mathbb{D}$ . For small enough  $p$  some of the coefficients  $\{a_k(p)\}_{k=1}^{\infty}$  are negative (since the third Taylor coefficient of  $H_0^{-1}(z) = \sin z$  is negative), implying that for every  $r \in [0, 1]$  we have

$$F_p^{-1}(r\gamma) = \sum_{k=1}^{\infty} a_k(p)r^k\gamma^k \in (-1, 1) \subseteq \mathbb{S}. \quad (31)$$

Let  $\mathcal{H}$  be a Hilbert space. Define two mappings  $L_p, R_p : \mathcal{H} \rightarrow \bigoplus_{k=1}^{\infty} \mathcal{H}^{\otimes k} \stackrel{\text{def}}{=} \mathcal{K}$  by

$$L_p(x) \stackrel{\text{def}}{=} \sum_{k=1}^{\infty} \sqrt{|a_k(p)|}\gamma^{k/2}x^{\otimes k};$$

$$R_p(x) \stackrel{\text{def}}{=} \sum_{k=1}^{\infty} \text{sign}(a_k(p))\sqrt{|a_k(p)|}\gamma^{k/2}x^{\otimes k}.$$

By (29), if  $\|x\|_{\mathcal{H}} = 1$  then  $\|L_p(x)\|_{\mathcal{K}} = \|R_p(x)\|_{\mathcal{K}} = 1$ . Moreover, if  $\|x\|_{\mathcal{H}} = \|y\|_{\mathcal{H}} = 1$  then

$$\langle L_p(x), R_p(y) \rangle = \sum_{k=1}^{\infty} a_k(p)\gamma^k \langle x, y \rangle^k = F_p^{-1}(\gamma \langle x, y \rangle) \stackrel{(31)}{\in} \mathbb{S}. \quad (32)$$

For  $N \in \mathbb{N}$  let  $G : \mathbb{R}^N \rightarrow \mathbb{R}^2$  be a  $2 \times N$  random matrix with i.i.d. standard Gaussian entries. Let  $g_1, g_2 \in \mathbb{R}^2$  be the first two columns of  $G$  (i.e.,  $g_1, g_2$  are i.i.d. standard two dimensional Gaussian vectors). If  $x, y \in \mathbb{R}^N$  are unit

vectors satisfying  $\langle x, y \rangle \in \mathbb{S}$  then by rotation invariance we have

$$\mathbb{E} \left[ f_\eta \left( \frac{1}{\sqrt{2}}Gx \right) f_\eta \left( \frac{1}{\sqrt{2}}Gy \right) \right] = \quad (33)$$

$$\mathbb{E} \left[ f_\eta \left( \frac{g_1}{\sqrt{2}} \right) f_\eta \left( \langle x, y \rangle \frac{g_1}{\sqrt{2}} + \sqrt{1 - \langle x, y \rangle^2} \frac{g_2}{\sqrt{2}} \right) \right] = \frac{1}{(2\pi)^2}$$

$$\times \int_{\mathbb{R}^2 \times \mathbb{R}^2} f_\eta \left( \frac{u}{\sqrt{2}} \right) f_\eta \left( \langle x, y \rangle \frac{u}{\sqrt{2}} + \sqrt{1 - \langle x, y \rangle^2} \frac{v}{\sqrt{2}} \right)$$

$$\cdot e^{-\frac{\|u\|_2^2 + \|v\|_2^2}{2}} dudv = \frac{2}{\pi} H_\eta(\langle x, y \rangle),$$

where we made the change of variable  $u = \sqrt{2}u'$  and  $v = (\sqrt{2}v' - \sqrt{2}\langle x, y \rangle u') / \sqrt{1 - \langle x, y \rangle^2}$ , whose Jacobian is  $4/(1 - \langle x, y \rangle^2)$ .

Fix an  $m \times n$  matrix  $A = (a_{ij})$  and let  $x_1, \dots, x_m, y_1, \dots, y_n \in \mathcal{H}$  be unit vectors satisfying

$$\sum_{i=1}^m \sum_{j=1}^n a_{ij} \langle x_i, y_j \rangle = M \stackrel{\text{def}}{=} \max \sum_{i=1}^m \sum_{j=1}^n a_{ij} \langle u_i, v_j \rangle, \quad (34)$$

where the maximum is taken over all unit vectors  $u_1, \dots, u_m, v_1, \dots, v_n$  in  $\mathcal{H}$ . Consider the unit vectors  $\{L_p(x_i)\}_{i=1}^m \cup \{R_p(y_j)\}_{j=1}^n$ , which we can think of as residing in  $\mathbb{R}^N$  for  $N = m + n$ . By (32) we have  $\langle L_p(x_i), R_p(y_j) \rangle \in \mathbb{S}$  for all  $i \in \{1, \dots, m\}$  and  $j \in \{1, \dots, n\}$ , so that we may use the identity (33) for these vectors. Let  $\lambda$  be a random variable satisfying  $\Pr[\lambda = 1] = p$ ,  $\Pr[\lambda = 0] = 1 - p$ . Assume that  $\lambda$  is independent of  $G$ . Define random variables  $\varepsilon_1, \dots, \varepsilon_m, \delta_1, \dots, \delta_n \in \{-1, 1\}$  by

$$\varepsilon_i = (1 - \lambda)f_0 \left( \frac{1}{\sqrt{2}}GL_p(x_i) \right) + \lambda f_\eta \left( \frac{1}{\sqrt{2}}GL_p(x_i) \right)$$

and

$$\delta_j = (1 - \lambda)f_0 \left( \frac{1}{\sqrt{2}}GR_p(y_j) \right) + \lambda f_\eta \left( \frac{1}{\sqrt{2}}GR_p(y_j) \right).$$

Then,

$$\max_{\substack{\sigma_1, \dots, \sigma_m, \tau_1, \dots, \tau_n \in \{-1, 1\}}} \sum_{i=1}^m \sum_{j=1}^n a_{ij} \sigma_i \tau_j \geq \mathbb{E} \left[ \sum_{i=1}^m \sum_{j=1}^n a_{ij} \varepsilon_i \delta_j \right]$$

$$\stackrel{(33)}{=} \frac{2}{\pi} \sum_{i=1}^m \sum_{j=1}^n a_{ij} \left( (1-p)H_0(\langle L_p(x_i), R_p(y_j) \rangle) \right.$$

$$\left. + pH_\eta(\langle L_p(x_i), R_p(y_j) \rangle) \right)$$

$$= \frac{2}{\pi} \sum_{i=1}^m \sum_{j=1}^n a_{ij} F_p(\langle L_p(x_i), R_p(y_j) \rangle)$$

$$\stackrel{(32)}{=} \frac{2}{\pi} \sum_{i=1}^m \sum_{j=1}^n a_{ij} F_p(F_p^{-1}(\gamma \langle x_i, y_j \rangle)) \stackrel{(34)}{=} \frac{2\gamma}{\pi} M.$$

This gives the bound  $K_G \leq \frac{\pi}{2\gamma} \stackrel{(30)}{<} \frac{\pi}{2 \log(1+\sqrt{2})}$ , as required.  $\blacksquare$

Our goal from now on will be to prove Theorem 5.1.

*Lemma 5.2:*  $H_0$  is one-to-one on  $\mathbb{S}$  and  $H_0(\mathbb{S}) \supseteq \mathbb{D}$ .

*Proof:* The fact that  $H_0$  is one-to-one on  $\mathbb{S}$  is a consequence of Lemma 4.2. To show that  $H_0(\mathbb{S}) \supseteq \mathbb{D}$  we need to prove that if  $a, b \in \mathbb{R}$  and  $a^2 + b^2 < 1$  then  $|\Re(\sin(a + bi))| < 1$ . Now,

$$|\Re(\sin(a + bi))| = \frac{e^b + e^{-b}}{2} |\sin a|. \quad (35)$$

Using the inequality  $|\sin a| \leq |a|$ , we see that it suffices to show that for all  $x \in (0, 1)$  we have

$$\frac{e^x + e^{-x}}{2} \sqrt{1 - x^2} < 1 \quad (36)$$

By Taylor's formula we know that there exists  $y \in [0, x]$  such that

$$\begin{aligned} \frac{e^x + e^{-x}}{2} &= 1 + \frac{x^2}{2} + \frac{x^4}{24} \cdot \frac{e^y + e^{-y}}{2} \\ &\leq 1 + \frac{x^2}{2} + \frac{x^4}{24} \cdot \frac{e + e^{-1}}{2} < 1 + \frac{x^2}{2} + \frac{x^4}{12}. \end{aligned} \quad (37)$$

Note that

$$\left(1 + \frac{x^2}{2} + \frac{x^4}{12}\right)^2 (1 - x^2) = 1 - \frac{7x^4}{12} - \frac{x^6}{3} - \frac{11x^8}{144} - \frac{x^{10}}{144} < 1,$$

which together with (37) implies (36).  $\blacksquare$

*Lemma 5.3:* For every  $r \in (0, 1)$  there exists  $p_r \in (0, 1)$  and

a bounded open subset  $\Omega_r \subseteq \mathbb{S}$  with  $\overline{\Omega_r} \subseteq \mathbb{S}$  such that for all  $p \in (0, p_r)$  the function  $F_p$  is one-to-one on  $\Omega_r$  and  $F_p(\Omega_r) = r\mathbb{D}$ . Thus  $F_p^{-1}$  is well defined and analytic on  $r\mathbb{D}$ .

*Proof:* For  $n \in \mathbb{N}$  consider the set

$$E_n = \left\{ z \in \mathbb{C} : |\Re(z)| < 1 - \frac{1}{n} \wedge |\Im(z)| < n \right\}.$$

Using Lemma 5.2, fix a large enough  $n \in \mathbb{N}$  so that  $H_0(E_n) \supseteq r\mathbb{D}$ . The bound (28) implies that there exists  $M > 0$  such that  $|H_\eta(z)| \leq M$  for all  $\eta > 0$  and  $z \in \partial E_{n+1}$ . By Lemma 5.2,  $H_0$  takes a value  $\zeta \in r\mathbb{D}$  exactly once on  $E_{n+1}$ , and this occurs at some point in  $E_n$ . Hence,

$$m \stackrel{\text{def}}{=} \min_{\substack{\zeta \in r\mathbb{D} \\ z \in \partial E_{n+1}}} |H_0(z) - \zeta| > 0.$$

Define  $p_r = m/(2M)$ .

Fix  $\zeta \in r\mathbb{D}$ . If  $p \in (0, p_r)$  then for every  $z \in \partial E_{n+1}$  we have  $|p(H_\eta(z) - H_0(z))| < \frac{m}{2M} (|H_\eta(z)| + |H_0(z)|) \leq m \leq |H_0(z) - \zeta|$ . Rouché's theorem now implies that the number of zeros of  $H_0 - \zeta$  in  $E_{n+1}$  is the same as the number of zeros of  $H_0 - \zeta + p(H_\eta - H_0) = F_p - \zeta$  in  $E_{n+1}$ . Hence  $F_p$  takes the value  $\zeta$  exactly once in  $E_{n+1}$ . Since  $\zeta$  was an arbitrary point in  $r\mathbb{D}$ , we can define  $\Omega_r = F_p^{-1}(r\mathbb{D})$ .  $\blacksquare$

*Lemma 5.4:* For every  $r \in (0, 1)$  there exists  $C_r \in (0, \infty)$  such that, using the notation of Lemma 5.3, for every  $p \in (0, p_r)$  and  $z \in r\mathbb{D}$  we have

$$\left| F_p^{-1}(z) - \sin z - p(z - H_\eta(\sin z)) \cos z \right| \leq C_r p^2. \quad (38)$$

*Proof:* Note that

$$z = F_p(F_p^{-1}(z)) = (1-p)H_0(F_p^{-1}(z)) + pH_\eta(F_p^{-1}(z)). \quad (39)$$

By differentiating (39) with respect to  $p$ , we see that

$$\begin{aligned} 0 &= H_\eta(F_p^{-1}(z)) - H_0(F_p^{-1}(z)) + \left( \frac{d}{dp} F_p^{-1}(z) \right) \\ &\quad \times \left( (1-p) \frac{dH_0}{dz}(F_p^{-1}(z)) + p \frac{dH_\eta}{dz}(F_p^{-1}(z)) \right) \\ &= H_\eta(F_p^{-1}(z)) - H_0(F_p^{-1}(z)) \\ &\quad + \left( \frac{d}{dp} F_p^{-1}(z) \right) \frac{dF_p}{dz}(F_p^{-1}(z)) \\ &= H_\eta(F_p^{-1}(z)) - H_0(F_p^{-1}(z)) + \frac{\frac{d}{dp} F_p^{-1}(z)}{\frac{d}{dz}(F_p^{-1}(z))}. \end{aligned}$$

Hence,

$$\frac{d}{dp} F_p^{-1}(z) = (H_0(F_p^{-1}(z)) - H_\eta(F_p^{-1}(z))) \frac{d}{dz}(F_p^{-1}(z)). \quad (40)$$

If we now differentiate (40) with respect to  $p$ , while using (40) whenever the term  $\frac{d}{dp} F_p^{-1}(z)$  appears, we obtain the following identity.

$$\begin{aligned} \frac{d^2}{dp^2} F_p^{-1}(z) &= \left[ \frac{dH_0}{dz}(F_p^{-1}(z)) - \frac{dH_\eta}{dz}(F_p^{-1}(z)) \right. \\ &\quad \left. + (H_0(F_p^{-1}(z)) - H_\eta(F_p^{-1}(z))) \frac{d^2}{dz^2}(F_p^{-1}(z)) \right] \\ &\quad \cdot (H_0(F_p^{-1}(z)) - H_\eta(F_p^{-1}(z))) \frac{d}{dz}(F_p^{-1}(z)). \end{aligned} \quad (41)$$

Take  $M = M_r > 0$  such that for all  $w \in \Omega_r$  we have

$$\max \left\{ |H_0(w)|, |H_\eta(w)|, \left| \frac{dH_0}{dz}(w) \right|, \left| \frac{dH_\eta}{dz}(w) \right| \right\} \leq M. \quad (42)$$

Note that (42) applies to  $w = F_p^{-1}(z)$  for  $z \in r\mathbb{D}$ . We also define  $R = R_r = \max_{w \in \partial \Omega_r} |w|$ . Then for  $\zeta \in \frac{1+r}{2} \mathbb{D}$  we have  $|F_p^{-1}(\zeta)| \leq R$ . If  $z \in r\mathbb{D}$  then by the Cauchy formula we have

$$\begin{aligned} \left| \frac{d}{dz}(F_p^{-1}(z)) \right| &= \left| \frac{1}{\pi(r+1)i} \oint_{\frac{1+r}{2} \partial \mathbb{D}} \frac{F_p^{-1}(\zeta)}{(\zeta - z)^2} d\zeta \right| \\ &\leq \max_{\zeta \in \frac{1+r}{2} \partial \mathbb{D}} \frac{|F_p^{-1}(\zeta)|}{(|\zeta| - |z|)^2} \leq \frac{4R}{(1-r)^2}. \end{aligned}$$

Similarly,

$$\left| \frac{d^2}{dz^2} F_p^{-1}(z) \right| = \left| \frac{2}{\pi(r+1)i} \oint_{\frac{1+r}{2} \partial \mathbb{D}} \frac{F_p^{-1}(\zeta)}{(\zeta - z)^3} d\zeta \right| \leq \frac{16R}{(1-r)^3}.$$



These estimates, in conjunction with the identity (41), imply the following bound:

$$\left| \frac{d^2}{dp^2} F_p^{-1}(z) \right| \leq \left( 2M + 2M \frac{16R}{(1-r)^3} \right) 2M \frac{4R}{(1-r)^2}.$$

By the Taylor formula we deduce that

$$\left| F_p^{-1}(z) - F_0^{-1}(z) - p \frac{d}{dp} F_p^{-1}(z) \Big|_{p=0} \right| \leq C_r p^2,$$

where  $C_r = \frac{8M^2 R}{(1-r)^2} \left( 1 + \frac{16R}{(1-r)^3} \right)$ . It remains to note that due to Lemma 4.2 and the identity (40), we have  $\frac{d}{dp} F_p^{-1}(z) \Big|_{p=0} = (z - H_\eta(\sin z)) \cos z$ . ■

*Proof of Theorem 5.1:* We will fix from now on some  $r \in (9/10, 1)$ . Note that since the Hermite polynomial  $h_5$  is odd, so is  $H_\eta$ . Hence also  $F_p$  is odd, and therefore  $a_k(p) = 0$  for even  $k$ . For  $z \in r\mathbb{D}$  write  $\phi(z) = (z - H_\eta(\sin z)) \cos z$ . Consider the power series expansions

$$\sin z = \sum_{k=0}^{\infty} b_{2k+1} z^{2k+1} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} z^{2k+1},$$

and

$$\phi(z) = \sum_{k=0}^{\infty} c_{2k+1} z^{2k+1}. \quad (43)$$

By the Cauchy formula we have for every  $k \in \mathbb{N} \cup \{0\}$ ,

$$|a_{2k+1}(p) - b_{2k+1} - pc_{2k+1}| = \left| \frac{1}{2\pi i r} \oint_{r\partial\mathbb{D}} \frac{F_p^{-1}(z) - \sin z - p\phi(z)}{z^{2k+2}} dz \right| \stackrel{(38)}{\leq} \frac{C_r p^2}{r^{2k+2}}. \quad (44)$$

Note that by Lemma 4.1 the radius of convergence of the series in (43) is at least 1, and therefore  $\sum_{k=0}^{\infty} |c_{2k+1}| (9/10)^{2k+1} < \infty$ . Hence,

$$\begin{aligned} \sum_{k=0}^{\infty} |a_{2k+1}(p)| \left( \frac{9}{10} \right)^{2k+1} &\stackrel{(44)}{\geq} \sum_{k=0}^{\infty} \frac{(9/10)^{2k+1}}{(2k+1)!} \\ &- p \sum_{k=0}^{\infty} |c_{2k+1}| \left( \frac{9}{10} \right)^{2k+1} - \frac{C_r p^2}{r} \sum_{k=0}^{\infty} \left( \frac{9}{10r} \right)^{2k+1} \\ &= \frac{e^{9/10} - e^{-9/10}}{2} - O(p) > 1.02 - O(p). \end{aligned} \quad (45)$$

By continuity, it follows from (45) that provided  $p$  is small enough there exists  $\gamma > 0$  satisfying the identity (29). Our goal is to prove (30), so assume for contradiction that  $\gamma \leq \log(1 + \sqrt{2}) < 9/10$ . Note that since  $r \in (9/10, 1)$  we have

$$\frac{\gamma}{r} \leq \frac{10 \log(1 + \sqrt{2})}{9} < \frac{49}{50}. \quad (46)$$

Fix  $\varepsilon > 0$  that will be determined later. We have seen in Lemma 5.2 that  $\sin(\frac{9}{10}\mathbb{D}) \subseteq \mathbb{S}$ . Since  $H_\eta$  is analytic on  $\mathbb{S}$ ,

it follows that  $\phi$  is analytic on  $\frac{9}{10}\mathbb{D}$ . Since  $\gamma < 9/10$ , there exists  $n \in \mathbb{N}$  satisfying

$$\sum_{k=n+1}^{\infty} |c_{2k+1}| \gamma^{2k+1} < \frac{\varepsilon}{2}. \quad (47)$$

There exists  $p = p(\varepsilon)$  such that for all  $p \in (0, p(\varepsilon))$  we have  $p|c_{2k+1}| < \frac{1}{2}|b_{2k+1}|$  for all  $k \in \{0, \dots, n\}$ . In particular, we have  $\text{sign}(b_{2k+1} + pc_{2k+1}) = \text{sign}(b_{2k+1}) = (-1)^k$ . Now,

$$\begin{aligned} \left| 1 - \frac{F_p^{-1}(i\gamma)}{i} \right| &\stackrel{(29)}{=} \left| \sum_{k=0}^{\infty} (|a_{2k+1}(p)| - (-1)^k a_{2k+1}(p)) \cdot \gamma^{2k+1} \right| \\ &\stackrel{(44)}{\leq} \sum_{k=0}^{\infty} \left| |b_{2k+1} + pc_{2k+1}| - (-1)^k (b_{2k+1} + pc_{2k+1}) \right| \cdot \gamma^{2k+1} \\ &\quad + 2 \sum_{k=0}^{\infty} \frac{C_r p^2}{r^{2k+2}} \gamma^{2k+1}. \end{aligned} \quad (48)$$

To estimate the two terms on the right hand side on (48), note first that

$$2 \sum_{k=0}^{\infty} \frac{C_r p^2}{r^{2k+2}} \gamma^{2k+1} \stackrel{(46)}{\leq} \frac{2C_r}{r} p^2 \sum_{k=0}^{\infty} \left( \frac{49}{50} \right)^{2k+1} \leq C'_r p^2, \quad (49)$$

where  $C'_r$  depends only on  $r$ . Since  $p \in (0, p(\varepsilon))$  we know that for all  $k \in \{0, \dots, n\}$  we have  $|b_{2k+1} + pc_{2k+1}| = (-1)^k (b_{2k+1} + pc_{2k+1})$ . Hence the first  $n$  terms of the first sum in the right hand side of (48) vanish. Therefore,

$$\begin{aligned} &\sum_{k=0}^{\infty} \left| |b_{2k+1} + pc_{2k+1}| - (-1)^k (b_{2k+1} + pc_{2k+1}) \right| \gamma^{2k+1} \\ &= \sum_{k=n+1}^{\infty} \left| |b_{2k+1} + pc_{2k+1}| - |b_{2k+1}| - (-1)^k pc_{2k+1} \right| \gamma^{2k+1} \\ &\leq 2p \sum_{k=n+1}^{\infty} |c_{2k+1}| \gamma^{2k+1} \stackrel{(47)}{<} p\varepsilon. \end{aligned} \quad (50)$$

By substituting (49) and (50) into (48), we see that if we define  $\beta = F_p^{-1}(i\gamma) - i$  then

$$|\beta| \leq C'_r p^2 + p\varepsilon. \quad (51)$$

Let  $L_0$  be the Lipschitz constant of  $H_0$  on  $i + \frac{1}{2}\mathbb{D} \subseteq \mathbb{S}$  (the disc of radius  $\frac{1}{2}$  centered at  $i$ ). Similarly let  $L_\eta$  be the Lipschitz constant of  $H_\eta$  on  $i + \frac{1}{2}\mathbb{D}$ , and set  $L = \max\{L_0, L_\eta\}$ . It follows that  $F_p = (1-p)H_0 + pH_\eta$  is  $L$ -Lipschitz on  $i + \frac{1}{2}\mathbb{D}$ . Due to (51), if  $p$  is small enough then  $i + \beta \in i + \frac{1}{2}\mathbb{D}$ , and therefore,

$$\begin{aligned} \log(1 + \sqrt{2}) &\geq \gamma = \frac{F_p(\beta + i)}{i} \geq \frac{F_p(i)}{i} - L|\beta| \\ &\stackrel{(51)}{\geq} \frac{(1-p)H_0(i) + pH_\eta(i)}{i} - Lp(C'_r p + \varepsilon) \\ &= (1-p) \log(1 + \sqrt{2}) + p \frac{H_\eta(i)}{i} - Lp(C'_r p + \varepsilon). \end{aligned}$$

This simplifies to give the following estimate:

$$\frac{H_\eta(i)}{i} \leq \log(1 + \sqrt{2}) + LC'_r p + L\varepsilon.$$

Since this is supposed to hold for all  $\varepsilon > 0$  and  $p \in (0, p(\varepsilon))$ , we arrive at a contradiction to Theorem 4.3. ■

*Remark 5.1:* An inspection of the proof of Theorem 1.1 shows that the only property of  $H_\eta$  that was used is that it is a counterexample to König's problem. In other words, assume that  $f, g : \mathbb{R}^2 \rightarrow \{-1, 1\}$  are measurable functions and consider the function  $H : \mathbb{S} \rightarrow \mathbb{C}$  given by  $H(z) = \frac{1}{2\pi(1-z^2)} \int_{\mathbb{R}^2 \times \mathbb{R}^2} f(x)g(y) \exp\left(\frac{-\|x\|_2^2 - \|y\|_2^2 + 2z\langle x, y \rangle}{1-z^2}\right) dx dy$ . Assume that  $B_K(f, g) > 4\pi \log(1 + \sqrt{2})$ , where  $B_K$  is König's bilinear form given in (5). Then one can repeat the proof of Theorem 1.1 with  $H_\eta$  replaced by  $H$ , arriving at the same conclusion.

*Conjecture 5.5:* Recalling Definition 2.1 and Corollary 2.3, we conjecture that for small enough  $\eta \in (0, 1)$ , the pair of functions  $f = g = f_\eta : \mathbb{R}^2 \rightarrow \{-1, 1\}$  is a Krivine rounding scheme for which we have  $c(f_\eta, f_\eta) > \frac{2}{\pi} \log(1 + \sqrt{2})$ . In other words, we conjecture that in order to prove Theorem 1.1 we do not need to use a convex combination of  $H_0$  and  $H_\eta$  as we did above, but rather use only  $H_\eta$  itself.

#### REFERENCES

- [1] 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. [Online]. Available: <http://dx.doi.org/10.1137/S0097539704441629>
- [2] G. E. Andrews, R. Askey, and R. Roy, *Special functions*, ser. Encyclopedia of Mathematics and its Applications. Cambridge: Cambridge University Press, 1999, vol. 71.
- [3] A. M. Davie, "Matrix norms related to Grothendieck's inequality," in *Banach spaces (Columbia, Mo., 1984)*, ser. Lecture Notes in Math. Berlin: Springer, 1985, vol. 1166, pp. 22–26. [Online]. Available: <http://dx.doi.org/10.1007/BFb0074689>
- [4] P. C. Fishburn and J. A. Reeds, "Bell inequalities, Grothendieck's constant, and root two," *SIAM J. Discrete Math.*, vol. 7, no. 1, pp. 48–56, 1994. [Online]. Available: <http://dx.doi.org/10.1137/S0895480191219350>
- [5] A. Frieze and R. Kannan, "Quick approximation to matrices and applications," *Combinatorica*, vol. 19, no. 2, pp. 175–220, 1999. [Online]. Available: <http://dx.doi.org/10.1007/s004930050052>
- [6] A. Grothendieck, "Résumé de la théorie métrique des produits tensoriels topologiques," *Bol. Soc. Mat. São Paulo*, vol. 8, pp. 1–79, 1953.
- [7] M. Grötschel, L. Lovász, and A. Schrijver, "The ellipsoid method and its consequences in combinatorial optimization," *Combinatorica*, vol. 1, no. 2, pp. 169–197, 1981. [Online]. Available: <http://dx.doi.org/10.1007/BF02579273>
- [8] 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>
- [9] S. Khot and A. Naor, "Grothendieck-type inequalities in combinatorial optimization," 2011, to appear in *Comm. Pure Appl. Math.*, preprint available at <http://arxiv.org/abs/1108.2464>.
- [10] H. König, "On an extremal problem originating in questions of unconditional convergence," in *Recent progress in multivariate approximation (Witten-Bommerholz, 2000)*, ser. Internat. Ser. Numer. Math. Basel: Birkhäuser, 2001, vol. 137, pp. 185–192.
- [11] J.-L. Krivine, "Constantes de Grothendieck et fonctions de type positif sur les sphères," *Adv. in Math.*, vol. 31, no. 1, pp. 16–30, 1979. [Online]. Available: [http://dx.doi.org/10.1016/0001-8708\(79\)90017-3](http://dx.doi.org/10.1016/0001-8708(79)90017-3)
- [12] —, "Sur la constante de Grothendieck," *C. R. Acad. Sci. Paris Sér. A-B*, vol. 284, no. 8, pp. A445–A446, 1977.
- [13] E. H. Lieb, "Gaussian kernels have only Gaussian maximizers," *Invent. Math.*, vol. 102, no. 1, pp. 179–208, 1990. [Online]. Available: <http://dx.doi.org/10.1007/BF01233426>
- [14] J. Lindenstrauss and A. Pełczyński, "Absolutely summing operators in  $L_p$ -spaces and their applications," *Studia Math.*, vol. 29, pp. 275–326, 1968.
- [15] G. Pisier, "Grothendieck's theorem, past and present," 2011, to appear in *Bull. Amer. Math. Soc. (N.S.)*, preprint available at <http://arxiv.org/abs/1101.4195>.
- [16] —, "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)
- [17] P. Raghavendra and D. Steurer, "Towards computing the Grothendieck constant," in *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2009, pp. 525–534.
- [18] J. A. Reeds, "A new lower bound on the real Grothendieck constant," 1991, unpublished manuscript, available at <http://www.dtc.umn.edu/reedsj/bound2.dvi>.
- [19] R. E. Rietz, "A proof of the Grothendieck inequality," *Israel J. Math.*, vol. 19, pp. 271–276, 1974.
- [20] B. S. Tsirelson, "Quantum analogues of Bell's inequalities. The case of two spatially divided domains," *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, vol. 142, pp. 174–194, 200, 1985, problems of the theory of probability distributions, IX.