Quadratic Forms on Graphs

Noga Alon * Konstantin Makarychev † Yury Makarychev ‡ Assaf Naor §

Abstract

We introduce a new graph parameter, called the *Grothendieck constant* of a graph $G = (V, E)$, which is defined as the least constant K such that for every $A : E \to \mathbb{R}$,

$$\sup_{f:V\to S^{|V|-1}} \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u), f(v)\rangle \le K \sup_{\varphi:V\to\{-1,+1\}} \sum_{\{u,v\}\in E} A(u,v)\cdot\varphi(u)\varphi(v).$$

The classical Grothendieck inequality corresponds to the case of *bipartite* graphs, but the case of general graphs is shown to have various algorithmic applications. Indeed, our work is motivated by the algorithmic problem of maximizing the quadratic form $\sum_{\{u,v\}\in E} A(u,v)\varphi(u)\varphi(v)$ over all $\varphi : V \to \{-1, 1\}$, which arises in the study of correlation clustering and in the investigation of the spin glass model. We give upper and lower estimates for the integrality gap of this program. We show that the integrality gap is $O(\log \vartheta(\overline{G}))$, where $\vartheta(\overline{G})$ is the Lovász Theta Function of the complement of G, which is always smaller than the *chromatic number of* G. This yields an efficient constant factor approximation algorithm for the above maximization problem for a wide range of graphs G. We also show that the maximum possible integrality gap is always at least $\Omega(\log \omega(G))$, where $\omega(G)$ is the *clique number* of G. In particular it follows that the maximum possible integrality gap for the complete graph on n vertices with no loops is $\Theta(\log n)$. More generally, the maximum possible integrality gap for any perfect graph with chromatic number n is $\Theta(\log n)$. The lower bound for the complete graph improves a result of Kashin and Szarek on Gram matrices of uniformly bounded functions, and settles a problem of Megretski and of Charikar and Wirth.

1 Introduction

An important inequality of A. Grothendieck [14] (see also [21]) states that for every $n \times m$ matrix (a_{ij}) and every choice of unit vectors $x_1, \ldots, x_n, y_1, \ldots, y_m \in S^{n+m-1}$ there exists a choice of signs $\varepsilon_1, \ldots, \varepsilon_n, \delta_1, \ldots, \delta_m \in \{-1, +1\}$ for which

$$\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}\langle x_i, y_j\rangle \le K_G \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}\varepsilon_i\delta_j. \tag{1}$$

*Schools of Mathematics and Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel and IAS, Princeton, NJ 08540, USA. Email: nogaa@tau.ac.il. Research supported in part by the Israel Science Foundation, by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University and by the Von Neumann Fund.

†Department of Computer Science, Princeton University. Email: kmakaryc@cs.princeton.edu. Research supported by NSF grants CCR-0205594, CCR-0237113, DOE award DE-FG02-02ER25594 and a Gordon Wu fellowship.

‡Department of Computer Science, Princeton University. Email: ymakaryc@cs.princeton.edu. Research supported by NSF grants CCR-0205594, CCR-0237113, DOE award DE-FG02-02ER25594 and a Gordon Wu fellowship.

§Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399, USA. Email: anaor@microsoft.com.

Here K_G is a universal constant, the best value of which is unknown (in [20] it is shown that $K_G \le \pi/[2\log(1+\sqrt{2})] = 1.782...$). In [3] Grothendieck's inequality was shown to have various algorithmic applications including the efficient construction of Szemerédi partitions of graphs and the estimation of the cutnorm of matrices, which yields efficient approximation algorithms for dense graph problems, using the methods of [12].

A quadratic variant of Grothendieck's inequality was studied in [10], where it was shown that there is a universal constant $C > 0$ such that for every $n \times n$ matrix (a_{ij}) and every $x_1, \ldots, x_n \in S^{n-1}$ there are signs $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}$ for which

$$\sum_{\substack{i,j \in \{1,\ldots,n\} \\ i \neq j}} a_{ij} \langle x_i, x_j \rangle \le C \log n \sum_{\substack{i,j \in \{1,\ldots,n\} \\ i \neq j}} a_{ij} \varepsilon_i \varepsilon_j. \tag{2}$$

This was also proved by Megretski in [25] and by Nemirovski, Roos and Terlaky in [26]. Various algorithmic applications of (2) were presented in [10], including an $O(\log n)$ approximation algorithm for the maximum correlation clustering problem suggested in [6] (and discussed in greater detail below). The authors of [10] (as well as the author of [25]) asked whether the logarithmic upper bound in (2) can be improved to a constant. In this paper we show that the $\log n$ term in (2) is in fact optimal. Our construction is based on a refinement of a recent result of Kashin and Szarek [19], who established an $\Omega(\sqrt{\log n})$ lower bound. Moreover, unlike in the classical functional analytic setting, inequalities (1) and (2) belong to a more general family of inequalities, all of which have algorithmic significance. Specifically, given a graph $G = (V, E)$ we define the *Grothendieck constant of* G, denoted $K(G)$, to be the least constant K such that for every matrix $A : V \times V \to \mathbb{R}$:

$$\sup_{f: V \to S^{|V|-1}} \sum_{\{u,v\} \in E} A(u,v) \cdot \langle f(u), f(v) \rangle \le K \sup_{\varphi: V \to \{-1,+1\}} \sum_{\{u,v\} \in E} A(u,v) \cdot \varphi(u)\varphi(v).$$

Grothendieck's inequality (1) simply states that if G is bipartite then $K(G) = O(1)$. On the other hand, inequality (2) states that for the complete graph K_n, $K(K_n) = O(\log n)$.

In this paper we initiate a systematic study of the parameter $K(G)$. We show that there is a natural way to interpolate between the bipartite case (1) and the case of the complete graph (2). Namely, we show that for every loop-free graph G,

$$\Omega(\log \omega(G)) = K(G) = O(\log \vartheta(\overline{G})),$$

where $\omega(G)$ is the clique number of G (i.e., the size of the largest clique in G), and $\vartheta(\overline{G})$ is the *Lovász theta function* of the complement of G, denoted $\overline{G}$ (see Section 3 for a definition of the Lovász theta function). As shown in [23], for every graph G, $\omega(G) \le \vartheta(\overline{G}) \le \chi(G)$, where $\chi(G)$ is the *chromatic number* of G. It follows that $K(G) = \Theta(\log \vartheta(\overline{G})) = \Theta(\log \chi(G))$ for every perfect graph G. The proofs are algorithmic, and provide efficient randomized algorithms for finding a function φ that approximates the maximum possible value of a given quadratic form

$$\sum_{\{u,v\} \in E} A(u,v) \cdot \varphi(u)\varphi(v) \quad \text{over all} \quad \varphi : V \to \{-1, +1\} \tag{3}$$

up to a factor of $K(G)$ for any loop-free graph G.

Various related results are also obtained. We hope that this will initiate future investigation of the interesting graph parameter $K(G)$ and its algorithmic applications.

The rest of the paper is organized as follows. After presenting a few basic results and definitions in Section 2, we prove, in Section 3, that $K(G) \leq O(\log \vartheta(\overline{G}))$, and obtain some related results. The fact that $K(G) \geq \Omega(\log \omega(G))$ is established in Section 4. Sections 5 and 6 contain some algorithmic consequences, and various additional remarks including several new Grothendieck-type inequalities. We end in Section 7, with several open problems.

2 Definitions and basic facts

Let $G = (V, E)$ be a graph on n vertices, which may have loops. We introduce the following parameters:

Definition 2.1 (The Grothendieck constant of G). *Denote by $K(G)$ the least constant K such that for every matrix $A : V \times V \to \mathbb{R}$:*

$$\sup_{f:V\to S^{n-1}} \sum_{\{u,v\}\in E} A(u,v) \cdot \langle f(u), f(v)\rangle \leq K \sup_{\varphi:V\to\{-1,+1\}} \sum_{\{u,v\}\in E} A(u,v) \cdot \varphi(u)\varphi(v),$$

The following definition plays a central role in the ensuing arguments. It is a formalization, and a generalization to the graph theoretical setting, of a concept studied in [19].

Definition 2.2 (The Gram representation constant of G). *Denote by $R(G)$ the infimum over constants R such that for every $f : V \to S^{n-1}$ there is a function $F : V \to L_\infty[0,1]$ so that for every $v \in V$ we have $\|F(v)\|_\infty \leq R$ and for every $\{u,v\} \in E$,*

$$\langle f(u), f(v)\rangle = \langle F(u), F(v)\rangle \equiv \int_0^1 F(u)(t)F(v)(t)dt.$$

The finiteness of the constants $R(G)$ and $K(G)$ for loopless graphs follows from Lemma 3.1 below, and the following simple lemma which relates these two parameters:

Lemma 2.3. *Let G be a graph without loops. Then $K(G) = R(G)^2$.*

Proof. To prove that $K(G) \leq R(G)^2$ fix $R > R(G)$ and $f : V \to S^{n-1}$. There is a function $F : V \to L_\infty[0,1]$ such that for every $v \in V$ we have $\|F(v)\|_\infty \leq R$ and for every $\{u,v\} \in E$, $\langle f(u), f(v)\rangle = \langle F(u), F(v)\rangle$. Then for every $A : V \times V \to \mathbb{R}$

$$\begin{aligned}
\sum_{\{u,v\}\in E} A(u,v) \cdot \langle f(u), f(v)\rangle &= \int_0^1 \left(\sum_{\{u,v\}\in E} A(u,v) \cdot \langle F(u)(t), F(v)(t)\rangle \right) dt \\
&\leq \sup_{g:V\to[-R,R]} \sum_{\{u,v\}\in E} A(u,v) \cdot g(u)g(v) \\
&= R^2 \sup_{\varphi:V\to\{-1,+1\}} \sum_{\{u,v\}\in E} A(u,v) \cdot \varphi(u)\varphi(v),
\end{aligned}$$

where we have used the fact that since G has no loops, the functional $\sum_{\{u,v\}\in E} A(u,v) \cdot g(u)g(v)$ is linear in each of the variables $\{g(v)\}_{v\in V}$.

In the reverse direction, for each $f : V \to S^{n-1}$ define $M(f) := (\langle f(u), f(v)\rangle)_{\{u,v\}\in E} \subseteq \mathbb{R}^E$. Similarly, each $\varphi : V \to \{-1,1\}$ defines $M(\varphi) := (\varphi(u)\varphi(v))_{\{u,v\}\in E} \subseteq \mathbb{R}^E$. Clearly

$$\operatorname{conv}\{M(\varphi) :\ \varphi : V \to \{-1,1\}\} \subseteq \operatorname{conv}\{M(f) :\ f : V \to S^{n-1}\}.$$

On the other hand, by the separation theorem, the validity of the inequality

$$\sup_{f:V\to S^{n-1}} \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u), f(v)\rangle \le K(G)\cdot \sup_{\varphi:V\to\{-1,+1\}} \sum_{\{u,v\}\in E} A(u,v)\cdot\varphi(u)\varphi(v)$$

for all $A : V\times V \to \mathbb{R}$, implies that

$$\operatorname{conv}\{M(f) :\ f : V \to S^{n-1}\} \subseteq K(G)\cdot \operatorname{conv}\{M(\varphi) :\ \varphi : V \to \{-1,1\}\}.$$

Thus for every $f : V \to S^{n-1}$ there are weights $\{\lambda_g :\ g : V \to \{-1,+1\}\}$ which satisfy $\sum_{g:V\to\{-1,+1\}} \lambda_g = 1$, $\lambda_g \ge 0$, and for every $\{u,v\} \in E$

$$\langle f(u), f(v)\rangle = K(G) \sum_{g:V\to\{-1,+1\}} \lambda_g \cdot g(u)g(v).$$

Partition the interval $[0,1]$ into subintervals $\{I_g :\ g : V \to \{-1,+1\}\}$ such that the length of I_g is λ_g. For every $v \in V$ let $F(v)$ be the function which takes the value $\sqrt{K(G)}g(v)$ on I_g. The above identity becomes simply: $\langle f(u), f(v)\rangle = \langle F(u), F(v)\rangle$, as required. $\square$

Remark 2.1. The proof above easily implies that for every graph G (that may contain loops), $R(G)^2$ is equal to the following slight modification of $K(G)$, denoted $K'(G)$, which is defined to be the least constant K' such that for every matrix $A : V \times V \to \mathbb{R}$:

$$\sup_{f:V\to S^{n-1}} \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u), f(v)\rangle \le K' \sup_{\varphi:V\to[-1,+1]} \sum_{\{u,v\}\in E} A(u,v)\cdot\varphi(u)\varphi(v),$$

3 Upper bounds

Observe that if H is a subgraph of G then $R(H) \le R(G)$. In [10] Charikar and Wirth show that for the complete graph on n-vertices, denoted K_n, $K(K_n) = O(\log n)$. See also [25], [26] for similar proofs. This fact is a corollary of the following estimate:

Lemma 3.1. *Let $K_n^{\circlearrowleft}$ denote the complete graph on n vertices with loops. Then*

$$R(K_n^{\circlearrowleft}) = O\left(\sqrt{\log n}\right).$$

Proof. The simple argument we present is taken from [19]. Denote by σ the normalized surface measure on S^{n-1}. By a straightforward computation, for some universal constant c,

$$\sigma\left\{x \in S^{n-1} : \|x\|_\infty \le c\sqrt{\frac{\log n}{n}}\right\} \ge 1 - \frac{1}{2n}.$$

By rotation invariance, for every $x \in S^{n-1}$, the random variable on the orthogonal group $O(n)$ given by $U \mapsto Ux$ is uniformly distributed on S^{n-1}. It follows that for every $f : V \to S^{n-1}$ there is a rotation $U \in O(n)$ such that for all $v \in V$, $\|U(f(v))\|_\infty \le c\sqrt{\frac{\log n}{n}}$. Let $F(v)$ be the function on $[0,1]$ such that $F(v)(t) \equiv \sqrt{n}\cdot(U(f(v)))_j$ for $(j-1)/n \le t < j/n$. Then $\|F(v)\|_\infty = O(\sqrt{\log n})$ and $\langle F(u), F(v)\rangle = \langle U(f(u)), U(f(v))\rangle = \langle f(u), f(v)\rangle$. This shows that $R(K_n^{\circlearrowleft}) = O\left(\sqrt{\log n}\right)$. $\square$

Remark 3.1. For any $n \geq 3$, $K(K_n^{\circlearrowleft}) = \infty$. Indeed, suppose that $1/3 > \epsilon > 0$, and consider the 3 by 3 matrix A_{ij} defined by $A_{ij} = -1$ for all $i \neq j$ and $A_{ii} = -1 + \epsilon$ for all i. It is easy to see that the maximum of the sum $\sum_{ij} A_{ij} \langle u_i, u_j \rangle$ for unit vectors u_i is at least 3ϵ, whereas the maximum of the quadratic form $\sum_{ij} A_{ij} x_i x_j$ over $\{-1, 1\}^3$ is $3\epsilon - 1$, which is negative. In fact, for any graph G that has at least one loop, if $K(G)$ exists, then it must be 1 (as we can consider a matrix A in which the only nonzero entry is -1, at the loop, or 1, at the loop).

Remark 3.2. When allowing negative entries on the diagonal, the integrality gap in the quadratic program may be infinite, as observed above, since the maximum over the discrete cube may be negative while the one over Gram matrices is positive. In fact, a stronger (simple) assertion holds. If P and NP differ, there is no polynomial time algorithm that approximates the maximum of the quadratic form above over the discrete cube up to any factor (even one that grows with n arbitrarily fast). This is because by defining the matrix A as one corresponding to the maximum cut of a graph (see the construction in [3] for more details), and by putting an appropriate constant in, say, A_{11}, letting all other diagonal entries of A be zero, we can, by an obvious binary search, find the value of the maximum cut of a graph precisely, if we can determine if the maximum of the quadratic form over the discrete cube is positive or negative.

We proceed to prove two theorems which strengthen the upper bound $K(K_n) = O(\log n)$. In Section 4 we show that in fact $K(K_n) = \Omega(\log n)$. This improves upon the lower bound $\Omega\left(\sqrt{\log n}\right)$ which was proved by Szarek and Kashin in [19].

Definition 3.2. *A strict vector k-coloring of a graph $G = (V, E)$ is a mapping $s : V \to \ell_2$, such that all vectors $s(u)$ are unit vectors, and for every two adjacent vertices u and v $\langle s(u), s(v) \rangle = -1/(k-1)$. Observe that we can always assume that s takes values in a $|V|$-dimensional Euclidean space.*

A graph is strictly vector k-colorable if it has a strict vector k-coloring. The strict vector chromatic number of a graph is the smallest real k for which the graph is strictly vector k-colorable.

The notion of strict vector colorability was introduced by Karger, Motwani, and Sudan in [18]. They showed that the strict vector chromatic number of a graph G is equal to the Lovász theta function of the complement of G, $\vartheta(\overline{G})$, which is bounded by the chromatic number of G. For the sake of completeness we recall that the original definition in [23] of the Lovász theta function of a graph on the vertices $\{1, \ldots, n\}$ is the minimum of $\max_{1 \leq i \leq n} \frac{1}{\langle x_i, y \rangle^2}$, where the minimum is taken over all choices of unit vectors x_i and y such that x_i and x_j are orthogonal for every pair of non-adjacent vertices i, j (we will not use this definition in what follows).

Theorem 3.3. *For any loop-free graph $G = (V, E)$, $K(G) = O\left(\log \vartheta(\overline{G})\right)$.*

The proof of Theorem 3.3 is based on the proof of Grothendieck's inequality appearing in [15], but requires several additional ideas as well. Before passing to the proof we introduce some notation, and motivate the ensuing arguments by recalling the reasoning in [15].

Let $g_1, g_2, \ldots$ be i.i.d. standard Gaussian random variables, defined on some probability space (Ω, P). Consider the *Gaussian Hilbert space* H, which is defined as the following closed subspace of $L_2(\Omega)$:

$$H = \left\{ \sum_{i=1}^{\infty} a_i g_i : \sum_{i=1}^{\infty} a_i^2 < \infty \right\}.$$

The scalar product on $L_2(\Omega)$ (and hence also on H) is defined for $X, Y \in L_2(\Omega)$ by $\langle X, Y\rangle = \mathbb{E}(X \cdot Y)$. We denote by $B(H)$ and $B(L_2(\Omega))$ the unit balls of H and $L_2(\Omega)$, respectively.

Define:

$$\Gamma \quad = \quad \sup_{f:V\to B(H)} \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u), f(v)\rangle, \tag{4}$$

and

$$\Delta = \sup_{\varphi:V\to[-1,+1]} \sum_{\{u,v\}\in E} A(u,v)\cdot\varphi(u)\varphi(v). \tag{5}$$

Our goal is to bound the ratio Γ/Δ.

For every $M > 0$ and $\psi \in L_2(\Omega)$, denote the truncation of ψ at M by

$$\psi^M(x) = \begin{cases} \psi(x) & \text{if } |\psi(x)| \le M \\ M & \text{if } \psi(x) > M \\ -M & \text{if } \psi(x) < -M. \end{cases} \tag{6}$$

Fix $f : V \to B(H)$ for which

$$\Gamma = \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u), f(v)\rangle.$$

The existence of f follows from a straightforward compactness argument.

The basic idea is to relate Γ to the scalar case via truncation. Specifically, fix $M > 0$ and assume that there exists a mapping $h : V \to H'$, for some Hilbert space H', for which it is possible to write

$$\Gamma = \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u)^M, f(v)^M\rangle + \sum_{\{u,v\}\in E} A(u,v)\cdot\langle h(u), h(v)\rangle. \tag{7}$$

The first term in (7) is bounded by relating to the scalar case:

$$\sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u)^M, f(v)^M\rangle = \mathbb{E}\sum_{\{u,v\}\in E} A(u,v)\cdot f(u)^M\cdot f(v)^M \le M^2\Delta. \tag{8}$$

When G is a bipartite graph with sides R, L (i.e. in the setting of the classical Grothendieck inequality), the proof in [15] proceeds by simply taking in (7) $h : V \to L_2(\Omega)\oplus L_2(\Omega)$ given by

$$h(v) = \begin{cases} \left(f(v) - f(v)^M\right)\oplus f(v)^M & v \in L, \\ f(v)\oplus\left(f(v) - f(v)^M\right) & v \in R. \end{cases}$$

Since all the elements of $B(H)$ have the same distribution as a Gaussian random variable with mean zero and variance at most one, we see that for $M \ge 1$ and every $v \in V$,

$$\|f(v) - f(v)^M\|_2^2 = \sqrt{\frac{2}{\pi}}\int_M^\infty x^2 e^{-x^2/2}dx \le 2Me^{-M^2/2}. \tag{9}$$

Thus, in this case, by the definition of Γ,

$$\sum_{\substack{u\in L\\ v\in R}} A(u,v)\cdot\langle h(u),h(v)\rangle \le 2\sqrt{2M}e^{-M^2/4}\cdot\Gamma. \tag{10}$$

Plugging (8) and (10) into (7) we see that $\Gamma \le M^2\Delta + 2\sqrt{2M}e^{-M^2/4}\cdot\Gamma$. Choosing M large enough we deduce that $\Gamma = O(\Delta)$, which concludes the proof of Grothendieck's inequality in [15].

When the graph $G=(V,E)$ is not necessarily bipartite, we can still bound the second term in (7) as follows:

$$\sum_{\{u,v\}\in E} A(u,v)\cdot\langle h(u),h(v)\rangle \le \left(\max_{v\in V}\|h(v)\|_{H'}\right)^2\cdot\Gamma.$$

We thus have the following

Corollary 3.4. *Fix $M>0$ and assume that (7) holds. Assume also that for every $v\in V$, $\|h(v)\|_{H'}^2 \le \frac{1}{2}$, then $\Gamma \le 2M^2\Delta$.*

We are now in position to conclude the proof of Theorem 3.3.

Proof of Theorem 3.3. Denote $k=\vartheta(\overline{G})$. Let $s:V\to\ell_2$ be a strict vector k-coloring of G. Let $U=\ell_2\oplus\mathbb{R}$. Define two mappings $t,\hat{t}:V\to U$ as follows

$$t(u)=\left(\sqrt{\frac{k-1}{k}}s(u)\right)\oplus\left(\sqrt{\frac{1}{k}}\cdot e_1\right),$$

and

$$\hat{t}(u)=\left(-\sqrt{\frac{k-1}{k}}s(u)\right)\oplus\left(\sqrt{\frac{1}{k}}\cdot e_1\right).$$

Then, $t(u)$ and $\hat{t}(u)$ are unit vectors. Moreover, for every adjacent vertices u and v we have $\langle t(u),t(v)\rangle = \langle \hat{t}(u),\hat{t}(v)\rangle = 0$, and $\langle t(u),\hat{t}(v)\rangle = \frac{2}{k}$.

Let Γ and Δ be as in (4) and (5), respectively. Our goal is to show that $\Gamma = O(\log k\cdot\Delta)$. As before, we fix $f:V\to B(H)$ for which

$$\Gamma=\sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u),f(v)\rangle.$$

Consider the function $h:V\to U\otimes L_2(\Omega)$ defined as follows:

$$h(u)=\frac{1}{4}t(u)\otimes(f(u)+f(u)^M)+k\cdot\hat{t}(u)\otimes(f(u)-f(u)^M).$$

With this definition one readily checks that identity (7) holds true. Moreover, for every $v\in V$,

$$\|h(v)\|^2_{U\otimes L_2(\Omega)} \le \left(\frac{1}{2}+k\|f(u)-f(u)^M\|_{L_2(\Omega)}\right)^2 \le \left(\frac{1}{2}+k\sqrt{2M}e^{-M^2/4}\right)^2,$$

where we have used the bound (9). Choosing, say, $M=8\sqrt{\log k}$ yields that for every $v\in V$, $\|h(v)\|^2_{U\otimes L_2(\Omega)}\le\frac{1}{2}$, so that by Corollary 3.4 $\Gamma\le(128\log k)\Delta$. $\square$

Remark 3.3. Since the strict vector chromatic number is less than or equal to the chromatic number $\chi(G)$, the result above implies that for any loop-free graph G, $K(G) = O(\log \chi(G))$.

We sketch below an alternative proof of Theorem 3.3. Although this proof is similar in spirit to the one above, we believe it sheds a different light on the problem. The argument is based on the following new characterization of the theta function, which may be interesting in its own right.

Theorem 3.5. *For a loop-free graph $G = (V, E)$ on $|V| = n$ vertices, let $Z = Z(G)$ denote the least constant Z such that for every matrix $A : V \times V \to \mathbb{R}$:*

$$-\inf_{g:V\to S^{n-1}} \sum_{\{u,v\}\in E} A(u,v)\cdot\langle g(u), g(v)\rangle \le Z \sup_{f:V\to S^{n-1}} \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u), f(v)\rangle. \qquad (11)$$

Then $Z = \vartheta(\overline{G}) - 1$.

Proof. Put $k = \vartheta(\overline{G})$. We first show that $Z(G) \le k - 1$. Fix a function $g : V \to S^{n-1}$. Let $s : V \to \ell_2$ be a strict vector k-coloring of G. Consider the function $t(v) : V \to S^{n-1} \otimes \ell_2$ defined as follows:

$$t(v) = g(v) \otimes s(v).$$

Then for any $v \in V$, $t(v)$ is a unit vector. Thus

$$\sum_{\{u,v\}\in E} A(u,v)\cdot\langle t(u), t(v)\rangle \le \sup_{f:V\to S^{n-1}} \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u), f(v)\rangle.$$

On the other hand,

$$\begin{aligned}\sum_{\{u,v\}\in E} A(u,v)\cdot\langle t(u), t(v)\rangle &= \sum_{\{u,v\}\in E} A(u,v)\cdot\langle g(u), g(v)\rangle\cdot\langle s(u), s(v)\rangle \\ &= -\frac{1}{k-1}\sum_{\{u,v\}\in E} A(u,v)\cdot\langle g(u), g(v)\rangle.\end{aligned}$$

Hence

$$-\sum_{\{u,v\}\in E} A(u,v)\cdot\langle g(u), g(v)\rangle \le (k-1) \sup_{f:V\to S^{n-1}} \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u), f(v)\rangle.$$

As this holds for every g as above, it follows that indeed $Z(G) \le k - 1$.

Next, we show that $k - 1 \le Z = Z(G)$. By the separation theorem, the validity of the inequality (11) for all matrices A implies that for every function $g : V \to S^{n-1}$, the vector $-(\langle g(u), g(v)\rangle)_{\{u,v\}\in E}$ lies in the convex hull of the set of all vectors in $\mathbb{R}^E$ of the form $Z \cdot (\langle f(u), f(v)\rangle)_{\{u,v\}\in E}$, as f ranges over all functions $f : V \to S^{n-1}$. As the latter set is convex it follows that $-(\langle g(u), g(v)\rangle)_{\{u,v\}\in E}$ lies in this set. In particular, this holds for the vector obtained by taking a constant $g : V \to S^{n-1}$, implying that there is an $s : V \to S^{n-1}$ such that $\langle s(u), s(v)\rangle = -\frac{1}{Z}$ for all $\{u, v\} \in E$. By the definition of the theta function, this implies that $k - 1 \le Z$, as needed. $\square$

It is worth noting that one of the known characterizations of the theta function, obtained in [23], implies that $\vartheta(\overline{G})-1$ is the least constant Z' such that for every matrix $A : V \times V \to \mathbb{R}$

$$-\inf_{g:V\to\mathbb{R}} \sum_{\{u,v\}\in E} A(u,v)\cdot g(u)g(v) \le Z' \sup_{f:V\to\mathbb{R}} \sum_{\{u,v\}\in E} A(u,v)\cdot f(u)f(v).$$

This can be viewed as a one-dimensional analogue of the theorem above.

An immediate consequence of Theorem 3.5 is the following.

Corollary 3.6. *For any loop-free graph $G=(V,E)$, for every matrix $A : V\times V \to \mathbb{R}$ and for every $g,h : V \to S^{n-1}$:*

$$\left|\sum_{\{u,v\}\in E} A(u,v)\cdot\langle g(u),h(v)\rangle\right| \le \vartheta(\overline{G}) \sup_{f:V\to S^{n-1}} \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u),f(v)\rangle.$$

Proof. This follows from the following simple observation:

$$\begin{aligned}
\sum_{\{u,v\}\in E} A(u,v)\cdot\langle g(u),h(v)\rangle &= \sum_{\{u,v\}\in E} A(u,v)\cdot\left\langle \frac{g(u)+h(u)}{2}, \frac{g(v)+h(v)}{2}\right\rangle \\
&\quad - \sum_{\{u,v\}\in E} A(u,v)\cdot\left\langle \frac{g(u)-h(u)}{2}, \frac{g(v)-h(v)}{2}\right\rangle \\
&\le \sup_{f:V\to S^{n-1}} \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u),f(v)\rangle - \\
&\quad \inf_{f':V\to S^{n-1}} \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f'(u),f'(v)\rangle \\
&\le \vartheta(\overline{G}) \sup_{f:V\to S^{n-1}} \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u),f(v)\rangle,
\end{aligned}$$

where in the last line we used Theorem 3.5. □

We can now sketch another proof of Theorem 3.3. Let $\Gamma, \Delta, B(H), f$ and f^M be as before. Now,

$$\begin{aligned}
\Gamma &= \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u),f(v)\rangle \\
&= \sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u)^M,f(v)^M\rangle \\
&\quad +\frac{1}{2}\sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u)+f(u)^M, f(v)-f(v)^M\rangle \\
&\quad +\frac{1}{2}\sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u)-f(u)^M, f(v)+f(v)^M\rangle. \qquad (12)
\end{aligned}$$

To bound the two last terms in (12) consider the functions $g,h :\to L_2(\Omega)$ given for $u,v\in V$ by

$$g(u) = \frac{f(u)+f(u)^M}{2} \quad \text{and} \quad h(v) = \frac{f(v)-f(v)^M}{\max_{w\in V}\|f(w)-f(w)^M\|_{L_2(\Omega)}}.$$

Then by Corollary 3.6 and (9) we get that

$$\begin{aligned}
&\frac{1}{2}\sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u)+f(u)^M, f(v)-f(v)^M\rangle \\
&= \left(\max_{w\in V}\|f(w)-f(w)^M\|_{L_2(\Omega)}\right)\cdot\sum_{\{u,v\}\in E} A(u,v)\cdot\langle g(u),h(v)\rangle \\
&\le \sqrt{2M}e^{-M^2/4}\cdot\sup_{g',h':V\to B(L_2(\Omega))}\sum_{\{u,v\}\in E} A(u,v)\cdot\langle g'(u),h'(v)\rangle \\
&= \sqrt{2M}e^{-M^2/4}\cdot\sup_{g',h':V\to S^{n-1}}\sum_{\{u,v\}\in E} A(u,v)\cdot\langle g'(u),h'(v)\rangle \\
&\le k\sqrt{2M}e^{-M^2/4}\cdot\Gamma.
\end{aligned}$$

Similarly,

$$\frac{1}{2}\sum_{\{u,v\}\in E} A(u,v)\cdot\langle f(u)-f(u)^M, f(v)+f(v)^M\rangle \le k\sqrt{2M}e^{-M^2/4}\cdot\Gamma.$$

Plugging these estimates, together with (8), into (12) we conclude that

$$\Gamma \le M^2\Delta + 2k\sqrt{2M}e^{-M^2/4}\cdot\Gamma.$$

Choosing, as before, $M = 8\sqrt{\log k}$, and simplifying, yields the assertion of Theorem 3.3.

Our next upper bound shows that for matrices with zeros on the diagonal whose entries are highly non-uniform, a better bound can be obtained for the constant K in the inequality appearing in Definition 2.1. This upper bound implies the $O(\log n)$ upper bound by the Cauchy-Schwartz inequality.

Theorem 3.7. *There exists a universal constant $C > 0$ such that for every $n \times n$ matrix (a_{ij}) with $a_{ii} = 0$ for all i:*

$$\max_{\|x_i\|_2=1}\sum_{i,j=1}^n a_{ij}\langle x_i,x_j\rangle \le C\log\left(\frac{\sum_{i,j=1}^n |a_{ij}|}{\sqrt{\sum_{i,j=1}^n a_{ij}^2}}\right)\cdot\max_{\varepsilon_i\in\{-1,+1\}}\sum_{i,j=1}^n a_{ij}\varepsilon_i\varepsilon_j.$$

Proof. By the argument of [10], it is enough to show that the maximum of the quadratic form $\langle Ax,x\rangle$ over the discrete cube $\{-1,1\}^n$ is at least $\Omega(\sigma)$, where $\sigma^2 = \sum_{ij} a_{ij}^2$. Let X be the random variable $X = \langle Ax,x\rangle$, where $x = (x_1,\dots,x_n)$ and x_i are independent, identically distributed, uniform random variables on $\{-1,1\}$. As $a_{ii} = 0$ for all i, the expectation of X is zero. Moreover, $E(X^2) = \sum_{ij} a_{ij}^2 = \sigma^2$. Since X is a polynomial of degree two, the Bonamie-Beckner inequality [8] implies that $E(X^4) = O(\sigma^4)$ (see also Lemma 3.3 in [4] for a direct argument). By part (ii) of Lemma 3.2 in [4], if the expectation of a random variable X is 0, $E(X^2) = \sigma^2 > 0$ and $E(X^4) \le b\sigma^4$, then X exceeds $\frac{\sigma}{4\sqrt{b}}$ with positive probability. This completes the proof. $\square$

4 A tight lower bound for the complete graph

In this section we refine the method of [19] and show that $K(K_n) = \Theta(\log n)$.

Lemma 4.1. *Let $\mathcal{N}$ be a 1-net in S^{d-1} (i.e. a maximal 1-separated subset of S^{n-1}). Assume that there is a function $F : \mathcal{N} \to L_\infty[0,1]$ such that for every distinct $x, y \in \mathcal{N}$ we have $\langle x, y\rangle = \langle F(x), F(y)\rangle$, and for every $x \in \mathcal{N}$ we have $1 \le \|F(x)\|_2^2 \le 1 + \frac{d}{3^{d+4}}$. Then*

$$\max_{x\in\mathcal{N}} \|F(x)\|_\infty \ge \frac{\sqrt{d}}{8}.$$

Proof. By standard volume estimates, $|\mathcal{N}| \le 3^d$. Let $\{e_x\}_{x\in\mathcal{N}}$ be the standard unit basis of $\mathbb{R}^{\mathcal{N}}$. For every $x \in \mathcal{N}$ define $w_x \in \mathbb{R}^d \oplus \mathbb{R}^{\mathcal{N}}$ by $w_x = x + \sqrt{\|F(x)\|_2^2 - 1}\cdot e_x$. Then $\|w_x\|_2 = \|F(x)\|_2$ and for every distinct $x, y \in \mathcal{N}$ we have $\langle w_x, w_y\rangle = \langle x, y\rangle = \langle F(x), F(y)\rangle$. It follows that the mapping $T : \mathrm{span}\{w_x : x \in \mathcal{N}\} \to \mathrm{span}\{F(\mathcal{N})\}$ given by $T(w_x) = F(x)$ is a linear isometry. Let $\tilde{T} : \mathbb{R}^d \oplus \mathbb{R}^{\mathcal{N}} \to L_2[0,1]$ be an arbitrary linear isometry which extends T.

Let $\psi_1, \dots, \psi_d$ be an arbitrary orthonormal basis of $\tilde{T}(\mathbb{R}^d)$ and for every $t \in [0,1]$ define

$$\rho_t = \left(\sum_{i=1}^d \frac{\psi_i(t)}{\sqrt{\sum_{i=1}^d \psi_i^2(t)}}\psi_i\right)\cdot \mathbf{1}_{\{\sum_{i=1}^d \psi_i^2(t)>0\}} \in \tilde{T}(B),$$

where B denotes the unit ball of $\mathbb{R}^d$. Since $\mathcal{N}$ is a 1-net in S^{d-1}, $\mathrm{conv}(\mathcal{N}) \supseteq \frac{1}{2}B$, so that $\mathrm{conv}(\tilde{T}(\mathcal{N})) \supseteq \frac{1}{2}\tilde{T}(B)$. It follows that for every $t \in [0,1]$ there is some $x(t) \in \mathcal{N}$ such that

$$\tilde{T}(x(t))(t) \ge \frac{1}{2}\rho_t(t) = \frac{1}{2}\sqrt{\sum_{i=1}^d \psi_i^2(t)}. \tag{13}$$

Define $A = \left\{t \in [0,1] : \sum_{i=1}^d \psi_i^2(t) \ge \frac{d}{2}\right\}$ and for every $x \in \mathcal{N}$ put $A_x = \{t \in A : x(t) = x\}$. Since $\sum_{i=1}^d \int_0^1 \psi_i^2(t)dt = d$ and $\sum_{i=1}^d \int_{[0,1]\setminus A} \psi_i^2(t)dt < d/2$ it follows that $\sum_{i=1}^d \int_A \psi_i^2(t)dt > d/2$. It follows from (13) that $\sum_{x\in\mathcal{N}} \int_{A_x} \tilde{T}(x)^2(t)dt = \int_A \tilde{T}(x(t))^2(t)dt > d/8$. So, there is some $x \in \mathcal{N}$ for which $\int_{A_x} \tilde{T}(x)^2(t)dt \ge d/(8|\mathcal{N}|) > d/(8\cdot 3^d)$.

We claim that $\|F(x)\|_\infty > \frac{\sqrt{d}}{8}$. Indeed, assuming the contrary we have that for every $t \in A_x$, $\tilde{T}(x)(t) - F(x)(t) > \frac{1}{2}\tilde{T}(x)(t)$, since for every $t \in A_x$, $\tilde{T}(x)(t) \ge \frac{1}{2}\rho_t(t) \ge \frac{\sqrt{d}}{4}$. But

$$\|F(x)\|_2^2 - 1 = \|w_x - x\|_2^2 = \|\tilde{T}(w_x) - \tilde{T}(x)\|_2^2 \ge \int_{A_x} [F(x) - \tilde{T}(x)]^2(t)dt \ge \int_{A_x} \frac{\tilde{T}(x)^2(t)}{4}dt > \frac{d}{3^{d+4}},$$

contradicting our assumption. □

Theorem 4.2. *There exists a universal constant $c > 0$ such that for every integer n,*

$$K(K_n) \ge c\log n.$$

Proof. Fix integers d, k and let $\mathcal{N}$ be a minimal 1-net in S^{d-1}. By the definition of the Gram representation constant, applied to the multi-set in which each element of $\mathcal{N}$ appears exactly k times, there are $\{f_{x,i} : x \in \mathcal{N},\ i \in \{1, \ldots, k\}\} \subseteq L_\infty[0,1]$ such that $\|f_{x,i}\|_\infty \le R(K_{k3^d})$ for every x and i, and for every distinct $(x,i),(y,j) \in \mathcal{N} \times \{1,\ldots,k\}$:

$$\langle f_{x,i}, f_{y,j}\rangle = \begin{cases} \langle x, y\rangle & x \neq y \\ 1 & x = y. \end{cases}$$

Define $F : \mathcal{N} \to L_\infty[0,1]$ by $F(x) = \frac{1}{k}\sum_{i=1}^k f_{x,i}$. Then for every distinct $x, y \in \mathcal{N}$ we have that $\langle F(x), F(y)\rangle = \langle x, y\rangle$. Moreover, for every $x \in \mathcal{N}$, using Lemma 3.1, we have that

$$\|F(x)\|_2^2 = \frac{k(k-1)}{k^2} + \frac{1}{k^2}\sum_{i=1}^k \|f_{x,i}\|_2^2 \le 1 - \frac{1}{k} + \frac{R(K_{k3^d})^2}{k} = 1 + O\left(\frac{\log(k3^d)}{k}\right) \le 1 + \frac{d}{3^{d+4}}, \quad (14)$$

where we have chosen $k = 6^d$, and d is assumed to be large enough. Now, for every distinct $i, j \in \{1, \ldots, k\}$ we have that $\|f_{x,i}\|_2^2 + \|f_{x,j}\|_2^2 \ge 2\langle f_{x,i}, f_{x,j}\rangle = 2$. Averaging this inequality over all such i, j we find that $\frac{1}{k}\sum_{i=1}^k \|f_{x,i}\|_2^2 \ge 1$. By (14) we deduce that $1 \le \|F(x)\|_2^2 \le 1 + \frac{d}{3^{d+4}}$, so that by Lemma 4.1, $R(K_{18^d}) = \Omega(\sqrt{d})$. Since $R(K_n) = \sqrt{K(K_n)}$ is increasing in n, it follows that $K(K_n) = \Omega(\log n)$. $\square$

5 Applications and examples

5.1 Algorithmic consequences

Our results yield polynomial time randomized algorithms for the following problem: Given a graph $G = (V, E)$ and $A : V \times V \to \mathbb{R}$, we wish to approximate up to a small multiplicative factor the value of

$$\beta := \max_{\varphi : V \to \{-1,1\}} \sum_{\{u,v\} \in E} A(u,v) \cdot \varphi(u)\varphi(v). \quad (15)$$

As we shall see presently, this combinatorial optimization problem encompasses several interesting algorithmic questions. A natural approach to the efficient evaluation of β is via *semidefinite relaxation*, namely we consider the value β^* given by

$$\beta^* := \max_{f : V \to S^{|V|-1}} \sum_{\{u,v\} \in E} A(u,v) \cdot \langle f(u), f(v)\rangle. \quad (16)$$

The continuous optimization problem in (16) is called the *natural semidefinite relaxation* of the combinatorial optimization problem in (15). Clearly $\beta^* \ge \beta$.

The advantage of passing to β^* is that such a semidefinite optimization problem can be solved in polynomial time (up to an arbitrarily small additive error) using the ellipsoid algorithm (see, e.g., [13] for details on semidefinite programming). Hence, we can efficiently produce a mapping $f : V \to S^{|V|-1}$ such that $\sum_{\{u,v\}\in E} A(u,v) \cdot \langle f(u), f(v)\rangle \ge (1 - o(1))\beta^*$. The proofs of all of our upper bounds (namely Theorem 3.3 and Theorem 3.7) are algorithmic in the sense that they efficiently produce $\varphi : V \to \{-1,1\}$ for which $\sum_{\{u,v\}\in E} A(u,v) \cdot \langle f(u), f(v)\rangle$ is at most the constant ensured by these theorems times $\sum_{\{u,v\}\in E} A(u,v) \cdot \varphi(u)\varphi(v)$. Indeed, the proof of Theorem 3.3 shows that the algorithm simply has to generate $|V|$ independent standard Gaussian random variables $\{g_u\}_{u\in V}$ and consider the random scalar-valued function $h : V \to [-1,1]$ given by

$$h(v) = \frac{1}{M}\left(\sum_{u\in V} f(v)_u g_u\right)^M,$$

where $M = 8\sqrt{\log\vartheta(\overline{G})}$ (observe that $\vartheta(\overline{G})$ itself can be computed in polynomial time using senidefinite programming). As G has no loops, $\sum_{\{u,v\}\in E} A(u,v)\cdot h(u)h(v)$ is a linear function in each entry $h(u)$, and we can thus shift each value of $h(u)$ that lies in the open interval $(-1,1)$, in its turn, to the boundary, without decreasing the value of the quadratic form. Thus $\beta^* \le O(\log\vartheta(\overline{G}))\cdot\beta$, and in fact our randomized algorithm produces a distribution over functions $\varphi : V :\to \{-1,1\}$ such that in expectation $\sum_{\{u,v\}\in E} A(u,v)\cdot\varphi(u)\varphi(v) \ge \beta/[O(\log\vartheta(\overline{G}))]$. A similar algorithm can be designed by replacing the Gaussian random variables by independent, uniform ±1 Bernoulli random variables- see the argument in [3] for more details.

As observed in [3], the problem of finding the maximum of (3) is MAX-SNP hard even when the graph G is bipartite, and hence it is interesting to find efficiently an approximation of it. Some algorithmic applications (of the bipartite case) are given in [3]. Here we describe two additional natural applications.

The first application arises in the study of the *spin glass model* in mathematical physics. In this model, there is a system of atoms each of which has a spin that can point either up or down. Some of the pairs of atoms have non-negligible interactions. The system can be described by a graph $G = (V,E)$, with a (not necessarily positive) weight $A(u,v)$ for each edge $uv \in E$. The atoms are the vertices, the spin of the atom v is $f(v) \in \{-1,1\}$, where 1 represents the "up" position, the weights $A(u,v)$ represent the interactions between u and v, and the energy of the system (when there is no external field) is given by its *Hamiltonian* $H = -\sum_{uv\in E} A(u,v)f(u)f(v)$. A *ground state* is a state that minimizes the energy, and thus the problem of finding a ground state is precisely that of finding the maximum of the integer program (3). See, e.g., [24], pp. 352-355, [27], and the references therein for more details.

It is known that if the graph G is planar, one can find a ground state in polynomial time using matching algorithms (see [7]), but in general this problem (which is equivalent to that of maximizing (3)) is NP-hard (and in fact even hard to approximate). Our technique here thus supplies an efficient way to find a configuration that approximates the minimum energy, and in many cases this approximation is up to a constant factor.

An additional algorithmic application of the fact one can find efficiently an approximation of (3) arises in the problem of maximizing correlation in correlation clustering.

Typical clustering problems involve the partitioning of a data set into classes which are small in some quantitative (typically metric) sense. In contrast, in correlation clustering, first considered in [6] (and referred to as *clustering with qualitative information* in [9]), we are given a judgement graph $G = (V,E)$ and for every $\{u,v\} \in E$ a real number $A(u,v)$ which is interpreted as a judgement of the similarity of u and v. In the simplest case $A(u,v) \in \{1,-1\}$, where if $A(u,v) = 1$ then u,v are said to be similar, and if $A(u,v) = -1$ then u,v are said to be dissimilar. Given a partition of V into clusters, a pair is called an agreement if it is a similar pair within one cluster or a dissimilar pair across two distinct clusters. Analogously, a disagreement is a similar pair across two different clusters or a dissimilar pair in one cluster. In the maximum correlation problem (MAXCORR) the goal is to partition V so that the correlation is maximized, where the correlation of a partition is the difference of the number of agreements and the number of disagreements. In the case of general weights, for a partition P of V into pairwise disjoint clusters, the value of the partition, denoted by $k(P)$, is the sum of all positive entries $A(u,v)$ for u,v that lie in the same cluster, minus the sum of all positive entries $A(u,v)$ for u,v that lie in distinct clusters, minus the sum of all values $|A(u,v)|$ for negative entries $A(u,v)$, where u,v lie in the same cluster, plus the sum of all values $|A(u,v)|$ for negative entries $A(u,v)$, where u,v lie in distinct clusters. The objective is to find a partition P whose value approximates the maximum possible value over all partitions.

The authors of [10] show that (2) can be used to provide an $\Omega(1/\log n)$ approximation algorithm for the maximum correlation in the correlation clustering problem when the judgement graph is the complete graph on n vertices. They prove that the maximum possible value of $k(P)$ is at most 3 times the maximum between the value k_n of the trivial partition into clusters of size 1, and the maximum possible value v_2 of a partition into two clusters. The latter is easily seen to be equivalent to the maximum of (3), and hence any algorithm that finds a solution giving at least an α-fraction of the maximum of (3) provides an $\alpha/3$-approximation algorithm for the maximum correlation clustering problem (this can be slightly improved to $\frac{\alpha}{2+\alpha}$ by being a bit more careful- see [10]). The authors of [10] conclude that there is an efficient $\Omega(1/\log n)$ approximation algorithm for the maximum correlation clustering problem for any graph G with n vertices. By the results here, the approximation guarantee can be improved to $\Omega(1/K(G)) = \Omega(1/\log(\vartheta(\overline{G})))$. In particular, this is $\Omega(1)$ for any bounded degree graph or any graph with a bounded chromatic number or genus - in all these cases and some additional ones considered in the next subsection no constant factor approximation algorithm for MAXCORR was previously known.

5.2 Restricted families of graphs

By Theorems 3.3 and 4.2, if G is a graph containing a clique of size whose logarithm is proportional to the logarithm of the theta function of $\overline{G}$, then $K(G) = \Theta(\log \omega(G))$. In particular, this holds for any perfect graph G, as in this case $\omega(G) = \chi(G) = \vartheta(\overline{G})$. Note that certain classes of perfect graphs, like comparability graphs or chordal graphs, occur in various applications, and therefore it might be desirable to optimize quadratic forms as in (3) on such graphs. Since the chromatic number of perfect graphs can be determined in polynomial time, it follows that in such cases we can determine the value of $K(G)$ up to a constant factor, and in case it is smaller than n^ϵ where n is the number of vertices, we can obtain a guaranteed improved approximation of (3).

Moreover, by the remark above and the fact that $\vartheta(\overline{G}) \leq \chi(G)$ for every G, $K(G) = \Theta(\log(\chi(G)))$ for any graph G whose chromatic number is bounded by a polynomial of the size of its largest clique. There are various known classes of such graphs, including complements of intersection graphs of the edges of hypergraphs with certain properties. See, for example, [2], [11]. Another interesting family of classes of graphs for which this holds is obtained as follows. Let $k \geq 2$ be an integer, and let $\mathcal{G}_k$ denote the family of all graphs which contain no induced star on k vertices. In particular, $\mathcal{G}_2$ is the family of all unions of pairwise vertex disjoint cliques, and $\mathcal{G}_3$ is the family of all claw-free graphs. An easy application of Ramsey Theorem implies that the maximum degree Δ of any graph $G \in \mathcal{G}_k$ whose largest clique is of size ω, is at most ω^{k-1}. This implies that the chromatic number χ of any such graph is at most $\omega^{k-1}+1$, showing that (for fixed k) its Grothendieck constant is $\Theta(\log \Delta) = \Theta(\log \omega) = \Theta(\log \chi)$. Since the maximum degree (as well as the fact that $G \in \mathcal{G}_k$ for some fixed k) can be computed efficiently, this is another case in which we can compute the value of $K(G)$ up to a constant factor.

A graph G is d-degenerate if any subgraph of it contains a vertex of degree at most d. It is easy and well known that any such graph is $d+1$-colorable, and that there is a linear time algorithm that finds, for a given graph G, the smallest number d such that G is d-degenerate. In particular, graphs of genus g are $O(\sqrt{g})$-degenerate, implying that their Grothendieck constant is $O(\log g)$.

Other classes of graphs for which the clique number is proportional (with a universal constant) to the chromatic number (and hence also to the theta function of the complement), are intersection graphs of a family of homothetic copies of a fixed convex set in the plane- see [17]. A few additional examples appear in the next subsection.

5.3 New Grothendieck-type inequalities

Theorem 3.3 enables us to generate new Grothendieck type Inequalities, and Theorem 4.2 can be used to show that in some cases these are essentially tight. We list here several examples that seem interesting.

Let $m > 2k > 1$ be integers. The *Kneser graph* $S(m,k)$ is the graph whose $n = \binom{m}{k}$ vertices are all k-subsets of an m-element set, where two vertices are adjacent iff the corresponding subsets are disjoint. The clique number of $S(m,k)$ is clearly $\lfloor m/k \rfloor$, and as shown in [22], its chromatic number is $m-2k+2$. Its theta function, computed in [23], is $\binom{m-1}{k-1}$, and as this graph is vertex transitive, the product of its theta function and that of its complement is the number of its vertices (see [23]). It thus follows that the theta-function of the complement of $S(m,k)$ is m/k. We conclude that the Grothendieck constant of $S(m,k)$ is $\Theta(\log(m/k))$. This gives the following Grothendieck-type inequality.

Proposition 5.1. *There exists an absolute constant c such that the following holds. Let $m > 2k > 1$ be positive integers. Put $M = \{1, 2, \ldots, m\}$, $n = \binom{m}{k}$, and let $A(I,J)$ be a real number for each pair of disjoint k-subsets I, J of M. Then for every vectors x_I of Euclidean norm 1, there are signs $\epsilon_I \in \{-1, 1\}$ such that*

$$\sum_{\substack{I,J\subset M\\ |I|=|J|=k\\ I\cap J=\emptyset}} A(I,J)\langle x_I, x_J\rangle \le c\log\left(\frac{m}{k}\right) \sum_{\substack{I,J\subset M\\ |I|=|J|=k\\ I\cap J=\emptyset}} A(I,J)\epsilon_I\epsilon_J.$$

Moreover, the above inequality is tight, up to the constant factor c, for all admissible values of m and k.

Let D_m denote the line graph of the directed complete graph on m vertices. This is the graph whose vertices are all ordered pairs (i,j) with $i,j \in M = \{1,2,\ldots,m\}$, $i \neq j$, in which (i,j) is adjacent to (j,k) for all for all admissible $i,j,k \in M$. It is known that the chromatic number of this graph is $(1+o(1))\log_2 m$, and it is not difficult to see that its clique number is 2 and its fractional chromatic number is at most 4, (see, for example, [5]). As shown in [23], the theta function of the complement of any graph is bounded by its fractional chromatic number. This gives the following inequality.

Proposition 5.2. *There exists an absolute constant c such that the following holds. Let m be a positive integer, and let $A(i,j,k)$ be a real for each $i,j,k \in M$, $i \neq j, j \neq k$. Then for every $m(m-1)$ vectors $x_{i,j}, (i,j \in M, i \neq j)$ of Euclidean norm 1, there are signs $\epsilon_{i,j} \in \{-1,1\}$ such that*

$$\sum_{\substack{i,j,k\in M\\ i\neq j\neq k}} A(i,j,k)\langle x_{i,j}, x_{j,k}\rangle \le c \sum_{\substack{i,j,k\in M\\ i\neq j\neq k}} A(i,j,k)\epsilon_{i,j}\epsilon_{j,k}.$$

Let $n = 2^m$, and let M be as before. Let G be the comparability graph of all subsets of M, ordered by inclusion. Then G is a perfect graph, and its clique number is $m + 1 = \log n + 1$. As G is perfect, this is also its chromatic number (and the theta function of its complement), providing the following inequality.

Proposition 5.3. *There exists an absolute constant c such that the following holds. Let m be a positive integer, put $n = 2^m$, and let $A(I, J)$ be a real for each $I \subsetneq J \subseteq M$. Then for every n vectors $x_I, (I \subseteq M)$ of Euclidean norm 1, there are signs $\epsilon_I \in \{-1, 1\}$ such that*

$$\sum_{I \subsetneq J \subseteq M} A(I, J)\langle x_I, x_J \rangle \leq c \log\log n \sum_{I \subsetneq J \subseteq M} A(I, J)\epsilon_I \epsilon_J.$$

This is tight, up to the constant c.

6 Further inequalities

6.1 Tensor products and analytic functions

Let $G = (V, E)$ be an n-vertex graph and $f_1, \ldots, f_k : V \to S^{n-1}$. Consider the function $f : V \to \ell_2$ given by the tensor product $f(v) = f_1(v) \otimes f_2(v) \otimes \cdots \otimes f_k(v)$. By the definition of $K(G)$, for every $A : V \times V \to \mathbb{R}$ we have that

$$\sum_{\{u,v\} \in E} A(u, v) \cdot \prod_{i=1}^{k} \langle f_i(u), f_i(v) \rangle \leq K(G) \max_{g:V \to \{-1,+1\}} \sum_{\{u,v\} \in E} A(u, v) \cdot g(u)g(v).$$

It follows that for every real analytic function $\Psi : \mathbb{R}^k \to \mathbb{R}$, all of whose Taylor coefficients are non-negative, the following inequality holds true:

$$\begin{aligned} &\sum_{\{u,v\} \in E} A(u, v) \cdot \Psi\left(\langle f_1(u), f_1(v) \rangle, \ldots, \langle f_k(u), f_k(v) \rangle\right) \\ &\qquad \leq \quad K(G) \cdot \Psi(1, \ldots, 1) \max_{g:V \to \{-1,+1\}} \sum_{\{u,v\} \in E} A(u, v) \cdot g(u)g(v). \end{aligned}$$

Similarly, there exists a function $F : V \to L_\infty[0, 1]$ such that for every $v \in V$, $\|F(v)\|_\infty \leq R(G) \cdot \sqrt{\Psi(1, \ldots, 1)}$ and

$$\{u, v\} \in E \Longrightarrow \Psi\left(\langle f_1(u), f_1(v) \rangle, \ldots, \langle f_k(u), f_k(v) \rangle\right) = \langle F(u), F(v) \rangle.$$

6.2 The homogenous Grothendieck inequality

The classical Grothendieck inequality has the following equivalent homogenous formulation: For every $n \times m$ matrix (a_{ij}),

$$\sup_{x_i, y_j \in \ell_2} \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} \langle x_i, y_j \rangle}{\max_{i,j}(\|x_i\|_2 \cdot \|y_j\|_2)} \leq K_G \cdot \sup_{x_i, y_j \in \mathbb{R}} \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} x_i y_j}{\max_{i,j}(|x_i| \cdot |y_j|)},$$

where K_G is Grothendieck's constant. Indeed, we clearly have that

$$\sup_{x_i,y_j\in\mathbb{R}} \frac{\sum_{i=1}^n\sum_{j=1}^m a_{ij}x_iy_j}{\max_{i,j}(|x_i|\cdot|y_j|)} = \sup_{\varepsilon_i,\delta_j\in\{-1,1\}} \sum_{i=1}^n\sum_{j=1}^m a_{ij}\varepsilon_i\delta_j,$$

and an analogous equality holds in the vector valued case.

For non-bipartite graphs the above phenomenon is no longer valid. To see this fix an integer m and consider the graph $G=(V,E)$ where $V=\{u_i\}_{i=1}^{m^2}\cup\{v_j\}_{j=1}^m$ and

$$E=\Big\{\{u_i,u_j\}:\ i\neq j\Big\}\cup\Big\{\{u_i,v_j\}:\ 1\le i\le m^2,\ 1\le j\le m\Big\}.$$

Define a symmetric matrix $A:V\times V\to\mathbb{R}$ via $A(u_i,u_j)=-1$ for all $i\neq j$ and $A(u_i,v_j)=A(v_j,u_i)=1$ for all $i\in\{1,\dots,m^2\}$ and $j\in\{1,\dots,m\}$. We claim that

$$\max_{\varepsilon_u\in\{-1,1\}}\sum_{\{u,v\}\in E} A(u,v)\varepsilon_u\varepsilon_v\le 2m^2.$$

Indeed, fix $\{\varepsilon_u\}_{u\in V}\subseteq\{-1,1\}^V$ and define $S=\sum_{i=1}^{m^2}\varepsilon_{u_i}$. Then

$$\begin{aligned}\sum_{\{u,v\}\in E} A(u,v)\varepsilon_u\varepsilon_v &= -\sum_{\substack{i,j\in\{1,\dots,m^2\}\\ i\neq j}}\varepsilon_{u_i}\varepsilon_{u_j}+2\sum_{i=1}^{m^2}\sum_{j=1}^m\varepsilon_{u_i}\varepsilon_{v_j}\\ &= m^2-S^2+2S\sum_{j=1}^m\varepsilon_{v_j}\\ &\le m^2-S^2+2m|S|\\ &\le 2m^2.\end{aligned}$$

On the other hand,

$$\sup_{x_u\in\mathbb{R}}\frac{\sum_{\{u,v\}\in E}A(u,v)x_ux_v}{\max_{\{u,v\}\in E}(|x_u|\cdot|x_v|)}\ge m^3+m.$$

Indeed, let $x_{u_i}=\frac{1}{\sqrt{m}}$ and $x_{v_j}=\sqrt{m}$. Then

$$\frac{\sum_{\{u,v\}\in E}A(u,v)x_ux_v}{\max_{\{u,v\}\in E}(|x_u|\cdot|x_v|)}=-\frac{m^4-m^2}{m}+2m^3=m^3+m.$$

In view of this example, it is of interest to study the natural homogenous extension of Grothendieck's inequality in the case of arbitrary graphs. We do not attempt to study this independently interesting notion here, except for the following variant of Theorem 3.3:

Proposition 6.1. *There is a universal constant $C>0$ such that for every loop-free graph $G=(V,E)$ and for every $A:V\times V\to\mathbb{R}$,*

$$\sup_{f:V\to\ell_2}\frac{\sum_{\{u,v\}\in E}A(u,v)\langle f(u),f(v)\rangle}{\max_{\{u,v\}\in E}(\|f(u)\|_2\cdot\|f(v)\|_2)}\le\big(C\log\vartheta(\overline{G})\big)\cdot\sup_{f:V\to\mathbb{R}}\frac{\sum_{\{u,v\}\in E}A(u,v)f(u)f(v)}{\max_{\{u,v\}\in E}(|f(u)|\cdot|f(v)|)}.$$

Proof. We shall use the notation of the proof of Theorem 3.3. Let Φ be the least constant $\phi > 0$ such that for every $f : V \to L_2(\Omega)$

$$\sum_{\{u,v\}\in E} A(u,v) \cdot \langle f(u), f(v)\rangle \le \phi \cdot \max_{\{u,v\}\in E} \left(\|f(u)\|_{L_2(\Omega)} \cdot \|f(v)\|_{L_2(\Omega)}\right).$$

We clearly have that $\Phi \le |V|^2 \cdot \max_{u,v\in V} |A(u,v)| < \infty$.

Fix $\rho > 0$ such that for every $f : V \to \mathbb{R}$

$$\sum_{\{u,v\}\in E} A(u,v) \cdot f(u)f(v) \le \rho \cdot \max_{\{u,v\}\in E} \left(|f(u)| \cdot |f(v)|\right).$$

We will conclude once we show that $\Phi \le C \log k \cdot \rho$.

Fix $f : V \to H$ for which

$$\sum_{\{u,v\}\in E} A(u,v) \cdot \langle f(u), f(v)\rangle \ge \frac{\Phi}{2} \cdot \max_{\{u,v\}\in E} \left(\|f(u)\|_{L_2(\Omega)} \cdot \|f(v)\|_{L_2(\Omega)}\right).$$

The existence of f follows from the fact that H and $L_2(\Omega)$ are isometric.

Observe that the following identity holds true:

$$\begin{aligned}
\frac{\Phi}{2} \cdot \max_{\{u,v\}\in E} \left(\|f(u)\|_{L_2(\Omega)} \cdot \|f(v)\|_{L_2(\Omega)}\right) &\le \sum_{\{u,v\}\in E} A(u,v) \cdot \langle f(u), f(v)\rangle) \\
= &\sum_{\{u,v\}\in E} A(u,v) \cdot \left\langle f(u)^{M\|f(u)\|_{L_2(\Omega)}}, f(v)^{M\|f(v)\|_{L_2(\Omega)}} \right\rangle + \\
&\frac{1}{2} \sum_{\{u,v\}\in E} A(u,v) \cdot \left\langle f(u) + f(u)^{M\|f(v)\|_{L_2(\Omega)}}, f(v) - f(v)^{M\|f(v)\|_{L_2(\Omega)}} \right\rangle + \\
&\frac{1}{2} \sum_{\{u,v\}\in E} A(u,v) \cdot \left\langle f(u) - f(u)^{M\|f(v)\|_{L_2(\Omega)}}, f(v) + f(v)^{M\|f(v)\|_{L_2(\Omega)}} \right\rangle.
\end{aligned} \tag{17}$$

Now,

$$\sum_{\{u,v\}\in E} A(u,v) \cdot \left\langle f(u)^{M\|f(u)\|_{L_2(\Omega)}}, f(v)^{M\|f(v)\|_{L_2(\Omega)}} \right\rangle \le M^2 \cdot \max_{\{u,v\}\in E} \left(\|f(u)\|_{L_2(\Omega)} \cdot \|f(v)\|_{L_2(\Omega)}\right) \cdot \rho.$$

Additionally, by the definition of Φ, the sum of additional terms in (17) is bounded by (as in Theorem 3.3)

$$\left(\frac{1}{2} + 8kMe^{-M^2/4}\right)^2 \cdot \Phi \cdot \max_{\{u,v\}\in E} \left(\|f(u)\|_{L_2(\Omega)} \cdot \|f(v)\|_{L_2(\Omega)}\right),$$

Plugging these estimates into (17) we get that:

$$\frac{\Phi}{2} \le M^2 \cdot \rho + \left(\frac{1}{2} + 8kMe^{-M^2/4}\right)^2 \cdot \Phi.$$

Choosing $M = 4\sqrt{\log k}$, and simplifying, yields the required result. $\square$

7 Open problems

The investigation of the parameter $K(G)$ and the optimization problem (3) leads to several interesting problems that remain open. In particular, we mention the following.

- Find an explicit family of matrices showing that $K(K_n) = \Omega(\log n)$.
- Is there a lower bound for $K(G)$ in terms of $\vartheta(\overline{G})$?
- Can the optimal value of (3) be approximated efficiently within a constant factor ? It is easy to see that this problem is MAX-SNP hard even for bipartite graphs, as observed in [3], but it will be interesting to show that for general graphs it cannot be approximated efficiently to within any constant factor.
- Let $G = G(n, 1/2)$ be the random graph on n labelled vertices obtained by picking each pair of distinct vertices to form an edge, randomly and independently, with probability 1/2. Is it true that almost surely (that is, with probability that tends to 1 as n tends to infinity), $K(G) = \Theta(\log n)$?

  Note that it is known that the clique number of G is, almost surely, $\Theta(\log n)$, and the theta function of its complement is, almost surely, $\Theta(\sqrt{n})$, (see [16]), implying that the Grothendieck constant $K(G)$ lies between $\Theta(\log\log n)$ and $\Theta(\log n)$, and that for this example it differs considerably either from the logarithm of the clique number or from that of the theta function of $\overline{G}$. Another example of a family of graphs with an even larger difference between these logarithms is given in [1], where graphs on n vertices with clique number 2 for which $\vartheta(\overline{G}) = \Theta(n^{1/3})$ are constructed.

Acknowledgement. We are grateful to Itai Benjamini for pointing out the connection to the spin glass model, and we thank Moses Charikar for useful comments and suggestions. We also thank the anonymous referees for several comments that improved the presentation of our results.

Added in proof. Some of the problems posed above have been addressed in the recent manuscript "On non-approximability for quadratic programs", by S. Arora, E. Berger, E. Hazan, G. Kindler and S. Safra. It is shown there that there exists a constant $\gamma > 0$ such that if $NP \not\subseteq DTIME\left(n^{(\log n)^3}\right)$, then the optimal value of (3), in the case $G = K_n$, cannot be approximated in polynomial time to a factor smaller than $O\left((\log n)^{\gamma}\right)$. Additionally, an *explicit* family of matrices is constructed, showing that $K(K_n) = \Omega\left(\frac{\log n}{\log\log n}\right)$.

References

[1] N. Alon, Explicit Ramsey graphs and orthonormal labelings, The Electronic J. Combinatorics 1 (1994), R12, 8pp.

[2] N. Alon, Covering a hypergraph of subgraphs, Discrete Mathematics 257 (2002), 249-254.

[3] N. Alon and A. Naor, Approximating the Cut-Norm via Grothendieck's Inequality, Proc. of the 36^{th} ACM STOC, 72-80, 2004.

[4] N. Alon, G. Gutin and M. Krivelevich, Algorithms with large domination ratio, J. Algorithms 50 (2004), 118-131.

[5] N. Alon and A. Orlitsky, Repeated communication and Ramsey graphs, IEEE Transactions on Information Theory 41 (1995), 1276-1289.

[6] N. Bansal, A. Blum and S. Chowla, Correlation Clustering, Proc. of the 43 IEEE FOCS, 238-247, 2002.

[7] F. Barahona, On the computational complexity of Ising spin glass models, J. Phys. A: Math. Gen. 15 (1982), 3241-3253.

[8] A. Bonamie, Etude de coefficients Fourier des fonctiones de $L^p(G)$. Ann. Inst. Fourier 20, 335-402, 1970.

[9] M. Charikar, V. Guruswami and A. Wirth, Clustering with qualitative information. Proc. of the 44 IEEE FOCS, 524–533, 2003.

[10] M. Charikar and A. Wirth, Maximizing quadratic programs: extending Grothendieck's Inequality, FOCS 2004, 54-60.

[11] G. Ding, P. D. Seymour and P. Winkler, Bounding the vertex cover number of a hypergraph, *Combinatorica* **14** (1994), 23–34.

[12] A. M. Frieze and R. Kannan, Quick Approximation to matrices and applications, Combinatorica **19** (1999), 175–200.

[13] M. Grötschel, L. Lovász and A. Schrijver, The ellipsoid method and its consequences in combinatorial optimization, Combinatorica 1 (1981), 169–197.

[14] A. Grothendieck, Résumé de la théorie métrique des produits tensoriels topologiques. Bol. Soc. Mat. Sao paolo, 8:1-79, 1953.

[15] W. B. Johnson and J. Lindenstrauss, Basic concepts in the geometry of Banach spaces, Handbook of the geometry of Banach spaces, Vol. I, 1–84, North-Holland, Amsterdam, 2001.

[16] F. Juhász, The asymptotic behaviour of Lovász' θ function for random graphs, Combinatorica 2 (1982), 153–155.

[17] S.-J. Kim, A. Kostochka and K. Nakprasit, On the Chromatic Number of Intersection Graphs of Convex Sets in the Plane, The Electronic J. Combinatorics 11 (2004), R52, 12pp.

[18] D. Karger, R. Motwani and M. Sudan, Approximate graph coloring by semidefinite programming, Journal of the ACM, Vol. 45(2), 1998, pp. 246-265.

[19] B. S. Kashin and S. J. Szarek, On the Gram Matrices of Systems of Uniformly Bounded Functions. Proceedings of the Steklov Institute of Mathematics, Vol. 243, 2003, pp. 227-233.

[20] J. Krivine, Sur la constante de Grothendieck. C. R. Acad. Sci. Paris Ser. A-B, 284:445-446, 1977.

[21] J. Lindenstrauss and A. Pełczyński, Absolutely summing operators in L_p spaces and their applications. Studia Math. 29, 275-326, 1968.

[22] L. Lovász. Kneser's conjecture, chromatic number and homotopy, *Journal of Combinatorial Theory*, 25:319–324, 1978.

[23] L. Lovász, On the Shannon capacity of a graph, IEEE Transactions on Information Theory 25(1) (1979), pp. 1-7.

[24] L. Lovász and M. D. Plummer, *Matching Theory*, North Holland, Amsterdam (1986).

[25] A. Megretski, Relaxation of Quadratic Programs in Operator Theory and System Analysis. In Systems, Approximation, Singular Integral Operators, and Related Topics (Bordeaux, 2000), Basel: Birkhäuser, 2001, pp. 365-392.

[26] A. Nemirovski, C. Roos and T. Terlaky, On Maximization of Quadratic Form over Intersection of Ellipsoids with Common Center. Mathematical Programming, 1999, Vol. 86 Issue 3, pp. 463-473.

[27] M. Talagrand, Spin glasses: a challenge for mathematicians. Cavity and mean field models, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge, 46. Springer-Verlag, Berlin, 2003.