Discrepancy Minimization via Regularization

Lucas Pesenti*   Adrian Vladu†

Abstract

We introduce a new algorithmic framework for discrepancy minimization based on *regularization*. We demonstrate how varying the regularizer allows us to re-interpret several breakthrough works in algorithmic discrepancy, ranging from Spencer's theorem [Spe85, Ban10] to Banaszczyk's bounds [Ban98, BDG19]. Using our techniques, we also show that the Beck-Fiala and Komlós conjectures are true in a new regime of *pseudorandom* instances.

1 Introduction

Discrepancy theory is a subfield of combinatorics which has branched in computer science due to its several connections to geometric problems, randomized algorithms, and complexity theory [Mat09, Cha00].

A landmark result in the field is Spencer's celebrated "Six standard deviations suffice" [Spe85]. In its simplest form, Spencer's paper considers a set system S of cardinality n over a ground set of n elements. The problem is to color each element of the ground set in red or blue, in such a way that all the sets are balanced, namely do not contain many more red than blue elements or vice versa. The maximum imbalance of a set induced by a coloring is called the *discrepancy* of the coloring.

Spencer's result offers an important insight into the limitations of the tools we generally employ to prove the existence of mathematical objects. A standard method to show that there always exists a low discrepancy coloring is to prove that a random coloring produces one with nonzero probability. To this end, one shows that for each set in S, a random coloring will have small discrepancy with high probability. Applying a union bound turns this into a simultaneous guarantee for all sets in S. This standard idea shows that one can always produce a coloring of discrepancy $O\left(\sqrt{n \log n}\right)$. Spencer [Spe85] shows that this approach misses even better colorings. Using a difficult nonconstructive argument, he proves that in fact, colorings of discrepancy $6\sqrt{n}$ exist. This result is tight up to constant factors, and exhibits an example where correlations between different sets in S can be exploited in order to overcome the limitations of the union bound technique, which essentially tries to ignore them.

Although in [Spe85] it is conjectured that there are inherent limitations to finding constructively a low discrepancy coloring, Bansal [Ban10] proved the contrary by exhibiting a polynomial-time algorithm whose output matches Spencer's bound up to constant factors. This provided a new direction for attacking open problems in the field, and was followed by a deluge of new algorithmic

---

*Bocconi University, `lucas.pesenti@phd.unibocconi.it`

†CNRS & IRIF, Université Paris Cité, `vladu@irif.fr`

results [LM15, ES18, Rot17, LRR17, BDG19, BDGL19, BM20, BLV22]. Notably, some of these results gave efficient algorithms to construct colorings for structured set systems, aiming towards breaking longstanding upper bounds in discrepancy theory, which are captured by the well-known Beck-Fiala and Komlós conjectures.

Interestingly, the recent works on discrepancy have increasingly relied on tools from continuous optimization, which have already shown great results in the area of fast algorithms. Hence, Bansal [Ban10] constructs a coloring by maintaining a fractional solution which gets updated by iteratively solving a semidefinite program, Lovett and Meka [LM15] perform a random walk through a polytope, Eldan and Singh [ES18] solve a linear program, and Levy, Ramadas and Rothvoss [LRR17] use an algorithm inspired by the multiplicative weights update method. Given the excellent results already delivered by continuous methods in other areas, using them to improve longstanding bounds in discrepancy theory remains an intriguing research direction. To this extent, it is tempting to ask whether casting these questions in an appropriate optimization framework will provide a natural approach that will lead to the sought answers.

1.1 Our contribution

In this paper, we further extend the connections between discrepancy and continuous optimization by providing a simple framework for discrepancy minimization based on a series of invocations of Newton's method on a regularized objective. Using this basic framework, we provide a twofold contribution.

- We show that using a version of the regularizer from [AZLO15], we obtain a simple and elegant proof of Spencer's result via a constructive algorithm. Using slightly different parameters, we also argue that in fact, "four standard deviations suffice".

  As previously noted in related works [AZLO15], the choice of the regularizer may prove critical to achieving the correct bound. Indeed, while [LRR17] regularize their problem with a negative entropy term, so that their updates reflect those corresponding to the multiplicative weights update method, we use instead the regularizer from [AZLO15]. This allows for a simpler and tighter way to control the increase of the (regularized) discrepancy over the course of the algorithm. This method parallels previous developments [LRR17, BLV22], while simultaneously providing a clean regularization framework for the problem.

- We extend the ideas of Potukuchi [Pot20], who gave an improved bound for the Beck-Fiala problem in the case where the input matrix is "pseudorandom". More precisely, we show that for Komlós instances, given an $n \times n$ matrix A with columns of at most unit ℓ_2-norm, we can achieve a discrepancy of $O(1 + \sqrt{\lambda \log n})$, where $\lambda = \max_{v \perp \mathbf{1}} \|A^{\odot 2}v\|_2/\|v\|_2$ (and $\odot$ denotes entry-wise product).

  This automatically improves Potukuchi's result from $O(\sqrt{s} + \lambda)$ to $O(\sqrt{s} + \sqrt{\lambda \log n})$ on Beck-Fiala instances with column-sparsity s, in the regime where $\lambda = \Omega(\log n)$. In addition, it implies that the Komlós conjecture is true for random rotation matrices and random Gaussian matrices.

Finally, we believe that this framework is powerful enough to provide new paths to attacking the major conjectures in the area.

1.2 Related work

Following Spencer's original paper [Spe85], an exciting series of algorithmic results has emerged. The first to provide a constructive proof was Bansal [Ban10], following which renewed efforts have attempted to make progress on the Beck-Fiala and Komlós conjectures.

In particular, Lovett and Meka [LM15] provide another algorithm recovering Spencer's theorem, based on a random walk in a polytope. Rothvoss [Rot17] shows another approach based on projections, which crucially relies on convex geometric arguments involving Gaussian width. Eldan and Singh [ES18] recover the same result by solving a linear program with a random linear objective. Levy, Ramadas and Rothvoss [LRR17] give another algorithm inspired by the multiplicative weights update method. Very recently, Bansal, Laddha and Vempala [BLV22] provided an algorithm using ℓ_p-norms as potential functions, which is very similar to our approach.

For the Beck-Fiala and Komlós problems, we note the algorithms from Bansal, Dadush and Garg [BDG19] and Bansal, Dadush, Garg and Lovett [BDGL19], that match Banaszczyk's bound of $O(\sqrt{\log n})$ for the Komlós problem, which in turn automatically implies a $O(\sqrt{s \log n})$ discrepancy bound for Beck-Fiala. In addition, a lot of work has focused on studying random instances of these conjectures [EL19, BM20, HR19, Pot20, AN22]. In particular, Potukuchi [Pot20] extracted a pseudorandom property of instances that is sufficient to obtain a low-discrepancy coloring.

We also acknowledge a series of very exciting recent developments on extended versions of Spencer's problem [RR20] and generalizations of it based on mirror descent and communication complexity [HRS22, DJR22].

Concerning regularization theory, while regularization techniques related to our ℓ_q-regularizer have long existed in the bandit literature [BC12], the first time such an idea was successfully employed in a combinatorial context was for computing optimal spectral sparsifiers [BSS14], where the authors used a barrier argument to track the evolution of the eigenvalues of a matrix. Later, Allen-Zhu, Liao and Orecchia [AZLO15] further developed this technique and made the connection to regularization explicit. This allowed them to improve the running time of the spectral sparsification algorithms from [BSS14].

We further comment on the link of some of these previous works with our approach in Section 2.1.

1.3 Organization of the paper

We will formally introduce our algorithmic framework in Section 3. Section 3.1 is devoted to reviewing the underlying iterative algorithm that is shared with several previous works. In Section 3.2, we will introduce different ways to regularize the discrepancy objective, and we will prove the corresponding technical bounds in Section 3.3.

Our first application of these tools is to the setting of Spencer's theorem. We will give three different proofs of Spencer's theorem, all based on the same idea but with different goals in mind:

- In Section 2, we sketch the main ideas of our framework on square matrices using Newton steps in the continuous limit.

- In Section 4.1, we discretize the previous approach and give a full algorithmic proof of Spencer's theorem in the general case.

- In Section 4.2, we come back to the square setting and give a more careful analysis of the

algorithm in order to optimize the leading constant. We improve Spencer's original constant from 5.32 to 3.67.

Section 5 is dedicated to the proof of our new bounds for pseudorandom instances of Komlós and Beck-Fiala conjectures. Our main results are Theorem 5.2 and Theorem 5.3, which we prove in Section 5.1 to Section 5.4. We present the consequences for random instances in Section 5.5.

Finally, we discuss some future research directions in Section 6.

1.4 Notations

Let n and m be some positive integers.

We will denote by log the natural logarithm. We set $[n] := \{1, \ldots, n\}$. We define the $(n-1)$-dimensional simplex by $\Delta_n := \{r \in \mathbb{R}^n_+ : \sum_{i \leqslant n} r_i = 1\}$, which is the set of probability distributions supported on $[n]$. We will use $\mathbf{1}$ as a shortcut for the vector $(1, \ldots, 1)$.

We consider the usual Euclidean inner product $\langle x, y \rangle := \sum_{i \leqslant n} x_i y_i$ on vectors $x, y \in \mathbb{R}^n$, and $\langle A, B \rangle := \sum_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n} A_{ij} B_{ij}$ on matrices $A, B \in \mathbb{R}^{m \times n}$. The ℓ_2-norm of a vector will by denoted by $\|x\|_2 := \langle x, x \rangle^{1/2}$ and its ℓ_∞-norm by $\|x\|_\infty := \max_{i \leqslant n} |x_i|$. The spectral norm (or $\ell_2 \to \ell_2$ operator norm) of a matrix $A \in \mathbb{R}^{m \times n}$ will be denoted by $\|A\|_{\mathrm{op}} := \max_{\|x\|_2 = 1} \|Ax\|_2$ and its Frobenius norm by $\|A\|_F := \langle A, A \rangle^{1/2}$. If $A \in \mathbb{R}^{m \times n}$ and $i \in [m]$, we write A_i for the i-th column of the matrix A^T (namely, the i-th row of A treated as a column vector). Similarly, if $j \in [n]$, A^j will denote the j-th column of A. We use the Hadamard notation $\odot$ to denote the entrywise product of vectors: for any $x, y \in \mathbb{R}^n$, we let $(x \odot y)_i := x_i y_i$ for all $i \in [n]$. We also set $x^{\odot 2} := x \odot x$. Given $A, B \in \mathbb{R}^{m \times n}$, we similarly define $(A \odot B)_{ij} := A_{ij} B_{ij}$ for all $i \in [m], j \in [n]$. Given a vector $x \in \mathbb{R}^n$, we define $\mathrm{diag}(x)$ to be the $n \times n$ diagonal matrix with the elements of x on the diagonal. Given $F \subseteq [n]$ and $x \in \mathbb{R}^F$, we will view interchangeably x as a vector on $\mathbb{R}^F$ and on $\mathbb{R}^n$, where the coordinates in $[n] \setminus F$ are filled with zeros.

Unless specified otherwise, the $\lesssim$, O and Ω notations will hide only universal constants.

2 Technical overview: Spencer's theorem via Newton steps in the continuous limit

In this section, we give an informal proof of a result of Spencer that illustrates how regularization comes into play in discrepancy minimization. For now, we aim at keeping the discussion simple and defer most details to Section 3 and Section 4.

Theorem 2.1 ([Spe85]). *There exists a universal constant $K > 0$ such that for any matrix $A \in [-1, 1]^{n \times n}$, there exists $x \in \{\pm 1\}^n$ such that*

$$\|Ax\|_\infty \leqslant K\sqrt{n}.$$

Many different proofs of Theorem 2.1 are already known: purely combinatorial [Spe85], with insights from convex geometry [Glu89, Gia97, Rot17, ES18], via random walks [LM15], with the multiplicative weights update method [LRR17], or via barrier potential functions [BLV22]. We will compare our approach with these last three techniques in Section 2.1.

4

Sticky walk. We build a deterministic sequence $x(t) := (x_1(t), \ldots, x_n(t))$ for times $t \in [0, T]$. At any time step t, $x(t)$ will be an element of the solid hypercube $[-1, 1]^n$ that represents a partial coloring. We start from $x(0) := (0, \ldots, 0)$ and the dynamic ends when $x(t)$ hits a corner of the hypercube. We define the set of active coordinates of a fractional coloring $x \in [-1, 1]^n$ as

$$F := \{j \in [n] : x_j \notin \{-1, 1\}\}.$$

The final algorithm described in Section 3 and Section 4 will essentially be a discretization of this continuous dynamic.

Regularized maximum. In order to control the quantity $\|Ax\|_\infty$ over the duration of the walk, we now define a smooth proxy for the ∞-norm. Since we can always add the negation of all the rows to the matrix A, without loss of generality it suffices to track $\max_{i \in [n]}(Ax)_i$. Naturally, for any $y \in \mathbb{R}^n$,

$$\max_{i \in [n]} y_i = \max_{r \in \Delta_n} \langle r, y \rangle, \quad \text{where } \Delta_n := \left\{ r \in \mathbb{R}_+^n : \sum_{i \leqslant n} r_i = 1 \right\}.$$

Instead, we consider the following regularized version of the right-hand side, which is the maximization problem where we added an $\ell_{1/2}$-type penalty for each element of the simplex:

$$\omega^*(y) := \max_{r \in \Delta_n} \langle r, y \rangle + \sum_{i \leqslant n} r_i^{\frac{1}{2}}.$$

In what follows, $\omega^*(Ax)$ will play a role of proxy for the ∞-norm. It is not hard to see (Lemma 3.8) that we only lose a $\sqrt{n}$ additive factor through this approximation. Therefore, for proving Theorem 2.1, it suffices to bound the total increase of the regularized maximum. We will discuss in Section 4.1 the choice of this particular $\ell_{1/2}$-regularizer.

Algorithm 1 Continuous dynamic for discrepancy minimization

1: **while** $F \neq \varnothing$ **do**
2: $\quad \dot{x} = \underset{\delta : \langle \delta, x \rangle = 0 \text{ and } \text{supp}(\delta) \subseteq F}{\arg \min} \langle A^\top \nabla \omega^*(Ax), \delta \rangle + \frac{1}{2} \cdot \delta^\top A^\top \nabla^2 \omega^*(Ax) A \delta$
3: $\quad F = \{i \in [n] : x_i \notin \{-1, 1\}\}$
4: **end while**
5: **return** x.

Continuous dynamic We sketch our dynamic in pseudo-code in Algorithm 1. Essentially, we impose two conditions on the update direction δ: $\text{supp}(\delta) \subseteq F$ ensures that the walk stays in the solid hypercube by fixing the coordinates of x when they reach ± 1, while $\langle \delta, x \rangle = 0$ ensures that the dynamic will eventually converge to a corner of the hypercube. Under these constraints, we select the direction that minimizes the best quadratic approximation of our potential function $\omega^*(Ax)$. In this sense, this is essentially a Newton step.

Local analysis. We would like to bound the increase in potential,

$$\frac{d\omega^*(Ax)}{d\|x\|_2^2} \approx \langle \nabla \omega^*(Ax), A\delta \rangle + \frac{1}{2} \langle A\delta, \nabla^2 \omega^*(Ax) A\delta \rangle,$$

where δ is the minimizer on line 2 of Algorithm 1.

Since the quadratic term is invariant by sign changes $\pm\delta$, we can always upper bound the term that is linear in δ by 0. Thus, it suffices to prove that the matrix $A^T \nabla^2 \omega^*(Ax) A$ has a small eigenvalue on the subspace $S := \{\delta \in \mathbb{R}^n : \langle \delta, x \rangle = 0 \text{ and } \mathrm{supp}(\delta) \subseteq F\}$. For this, we need to understand better the regularization construction – as we will see in Lemma 3.9, it follows from standard convex analysis arguments that

$$\nabla^2 \omega^*(Ax) \preccurlyeq \mathrm{diag}(\nabla)^{\frac{3}{2}} \quad \text{for some vector } \nabla \in \Delta_n \,.$$

By further use of the orthogonality trick, we can select a slightly smaller subspace than S whose elements "don't see" the rows i for which $\nabla_i \gtrsim 1/|F|$. A random element δ in this subspace achieves quadratic form at most $\|\delta\|_2^2/\sqrt{|F|}$ in expectation. The details can be found in Lemma 4.2. This ultimately implies

$$\mathrm{d}\omega^*(Ax) \lesssim \frac{\mathrm{d}\|x\|_2^2}{\sqrt{|F|}} \,. \tag{1}$$

Analysis of the whole dynamic. For $k \in [n]$, denote by $t_k := \min\{t \geqslant 0 : |F(t)| \leqslant k\}$ the first time for which the number of active coordinates reaches k. From the constraint that the update direction is always orthogonal to the current partial coloring, we get after integrating (1) over $t \in [0, T]$ that

$$\omega^*(Ax(T)) - \omega^*(0) \lesssim \sum_{k=1}^{n} \frac{\|x(t_{k-1})\|_2^2 - \|x(t_k)\|_2^2}{\sqrt{k}} \,.$$

Finally, we apply summation by parts and use the fact that $\sum_{i \leqslant k} \|x(t_{i-1})\|_2^2 - \|x(t_i)\|_2^2 \leqslant k$:

$$\omega^*(Ax(T)) - \omega^*(0) \lesssim \sqrt{n} + \sum_{k=1}^{n-1} \frac{1}{k^{\frac{3}{2}}} \sum_{i \leqslant k} \|x(t_{i-1})\|_2^2 - \|x(t_i)\|_2^2 \lesssim \sqrt{n} \,.$$

Combining this with our previous observation that $\omega^*(0) \lesssim \sqrt{n}$ concludes our proof sketch of Theorem 2.1.

2.1 Comparison with existing approaches

The general framework of tracking the discrepancy of a continuously evolving partial coloring through the means of a smooth approximation appears in the literature in various similar forms. Several other works on discrepancy use techniques that are similar to ours, the most important being those due to Lovett-Meka [LM15], Levy-Ramadas-Rothvoss [LRR17], and Bansal-Laddha-Vempala [BLV22]. In what follows we provide a brief overview of these.

Comparison with [LM15]. Here, the authors give an algorithmic proof of Spencer's theorem, where they evolve a partial coloring using a *sticky* Brownian motion on the hypercube, and similarly to our case, they freeze coordinates once they reach ± 1. Their algorithm, however, operates inside a convex set corresponding to low discrepancy colorings, and additionally requires maintaining several technical conditions. At the end of a phase, they only obtain a partial coloring, and need to repeat this procedure several times.

Comparison with [LRR17]. In this subsequent work, the authors propose an elegant deterministic algorithm, inspired by Lovett-Meka as well as by the multiplicative weights update method. Their algorithm is very close in spirit to ours, as it iteratively updates a fractional coloring while controlling the exponential weights that are assigned to the set constraints. The exponential weights can be seen as a proxy for a smooth regularization of the maximum function. In fact, this approach is related to Spencer's *hyperbolic cosine* algorithm [Spe85], that essentially consists in tracking the evolution of $\sum_i \cosh(Ax)_i$, which is up to a reparametrization equivalent to our setting when using entropic regularization (see Definition 3.7).

While the hyperbolic cosine algorithm is unable to obtain discrepancy below $O\left(\sqrt{n \log n}\right)$, the authors of [LRR17] do so by combining it with the continuous approach from [LM15]. This approach alone does not directly manage to recover Spencer's bound. The reason can be easily understood when interpreting their algorithm through the regularization perspective: the *error* introduced by entropic regularization contains a logarithmic term which carries over to the final discrepancy bound. They observe, however, that they can force the approximation provided by their regularized maximum to be far from tight, in the sense that the true discrepancy is not as large as what the potential function "sees". This property is enforced by taking sufficiently small steps, to ensure that at all times, all the rows attaining the largest discrepancy contribute equally. This forces the regularizer to spread a large part of its mass uniformly over these, and thus maximize the error it pays for beyond the true value of the maximum discrepancy. Just like in [LM15], this approach only obtains a partial coloring, and needs to be repeated with a different setting of parameters.

Comparison with [BLV22]. In this parallel work, the authors propose a unified approach for constructive discrepancy minimization using a barrier-based potential function. Their algorithm can be viewed as almost equivalent to ours, although the reason why the potential function works may appear quite magical. Compared to [LRR17], they replace exponential weights with a sum of inverse p powers of slacks, where the slacks measure for each row the distance between a desired discrepancy upper bound and the current discrepancy. In our regularization framework, a similar barrier emerges directly from choosing an appropriate regularizer. As we will soon see, we can derive it from first principles, and rather than having to guess a potential function and do tedious calculations to understand its evolution, we simply need to focus our attention on the trade-off between the error it introduces and the eigenvalues of its Hessian (see, e.g., our analysis in Section 4).

Relation to regret minimization frameworks [BSS14, AZLO15]. While both [LRR17] and [BLV22] rely on tracking a potential function, we attempt to make this approach more principled. The barrier potential present in [BLV22] appears to be related to the one employed by Batson-Spielman-Srivastava [BSS14] in the context of spectral sparsification. Interestingly, the reason a barrier was used in the case of sparsification was exactly to remove an extra logarithmic factor that would have otherwise occurred when using standard entropic regularization/multiplicative weights. Allen-Zhu, Liao and Orecchia [AZLO15] made the connection between the barrier potential and the multiplicative weights method explicit by noticing that both follow from using different regularizers on top of the maximum function (although in their case they more generally regularize matrix norms). Note that in [AZLO15], the authors provide bounds on the second-order term of their regularized spectral norms in the form of multiplicative error on the gradient term. Here, we

directly relate the second-order terms to the gradient, which allows us to obtain tighter bounds on the change in our potential functions. We believe this to be of independent interest.

3 The regularization framework

3.1 An iterative meta-algorithm

We first describe a generic iterative algorithm for discrepancy minimization that will serve as a basis for incorporating the potential functions based on regularization. Similarly to Section 2, we will construct a sequence of partial colorings $x(t) \in [-1, 1]^n$ for integer times $t = 0, 1, \ldots$. Each step consists in picking an update vector δ and adding it to $x(t)$. Whenever some coordinate of $x(t)$ becomes ± 1, we say that the coordinate is *frozen*. We will also say of an unfrozen coordinate that it is *active*.

Algorithm 2 Generic iterative algorithm for discrepancy minimization

---

1: **Input:** $A \in \mathbb{R}^{m \times n}$, $L \in (0, 1)$
2: **Output:** $x \in \{\pm 1\}^n$ (a low-discrepancy coloring of A)
3: Let $x(0) := (0, \ldots, 0)$ and $t := 0$.
4: **while** oracle$(A, x(t))$ is not **undefined do**
5:    Choose any unit vector δ in oracle$(A, x(t)) \cap \{\delta \in \mathbb{R}^n : \langle \delta, x(t) \rangle = 0\}$.
6:    Let $\varepsilon(t) := \min\{\varepsilon > 0 : \exists j \in [n], x_j(t) \notin \{-1, 1\}$ and $x_j(t) + \varepsilon \delta_j \in \{-1, 1\}\}$.
7:    Set $x(t+1) := x(t) + \min(L, \varepsilon(t))\delta$.
8:    Update $t := t + 1$.
9: **end while**
10: Let $T := t$ and $x_j^* := \text{sign}(x_j(T))$ for all $j \in [n]$.
11: **return** x^*.

The oracle. Suppose that we are given some blackbox algorithm oracle that encapsulates all the possible choices of directions of the update vector. In the sequel, oracle(A, x) will correspond to a subset of vectors that do not increase too much the value of the regularized potential function when x is the current partial coloring.

Assumption 3.1. Let $C > 0$ be some universal constant. Given a matrix $A \in \mathbb{R}^{m \times n}$ and a partial coloring $x \in [-1, 1]^n$, oracle(A, x) satisfies (with $F := \{j \in [n] : x_j \notin \{-1, 1\}\}$):

- If $|F| \geqslant C$, oracle(A, x) is a subset of $\mathbb{R}^F$ such that the intersection of oracle(A, x) with any halfspace of $\mathbb{R}^F$ contains a half-line.

- If $|F| < C$, it returns the value **undefined**.

With oracle being given, the meta-algorithm for discrepancy minimization is described as Algorithm 2. The following three immediate observations on Algorithm 2 will be central to our framework.

Observation 3.2. For any $t = 0, \ldots, T - 1$, $\|x(t+1) - x(t)\|_\infty \leqslant L$.

Observation 3.3. The final step on line 9 adds at most $C \max_{i \in [m], j \in [n]} |A_{ij}|$ to the discrepancy of the coloring, where C is the constant from Assumption 3.1.

Observation 3.4. There can be at most n/L^2 iterations of the main loop of Algorithm 2. Therefore, Algorithm 2 runs in polynomial time as long as oracle runs in polynomial time and $L \geqslant n^{-O(1)}$.

Proof. When $\varepsilon(t) \leqslant L$, at least one additional coordinate will reach ± 1 and will be frozen at the end of the iteration. This can happen at most n times. When $\varepsilon(t) > L$, since we pick our update vector orthogonal to x, we have $\|x(t+1)\|_2^2 = \|x(t)\|_2^2 + L^2$. This can happen at most n/L^2 times. $\square$

For our purposes, we will always set $L = n^{-O(1)}$ and computing oracle will only require elementary linear algebraic operations in $\mathbb{R}^n$ (intersection, orthogonal complements, direct sums, computation of eigenspaces, etc.).

3.2 Regularized maximum

Our main tool for building proxies for discrepancy is the following regularized version of the maximal entry of a vector.

Definition 3.5. For any convex function $\phi : \Delta_m \to \mathbb{R}$, we define $\phi^* : \mathbb{R}^m \to \mathbb{R}$ by

$$\phi^*(y) := \max_{r \in \Delta_m} \langle r, y \rangle - \phi(r).$$

We will call ϕ the *regularizer* – it maps elements of the simplex to some penalty in a convex way. By symmetry, it makes sense to focus on regularizers of the form $\phi(r) = \sum_{i \in [m]} \varphi(r_i)$ for some convex $\varphi : \mathbb{R} \to \mathbb{R}$. The following two special cases will play an important role in our theory.

Definition 3.6. For any $0 < q < 1$, the ℓ_q-*regularization* of the maximum, parametrized by $\eta > 0$, is the function $\omega_{q,\eta}^* : \mathbb{R}^m \to \mathbb{R}$ such that

$$\omega_{q,\eta}^*(y) := \max_{r \in \Delta_m} \langle r, y \rangle + \frac{1}{\eta q} \sum_{i=1}^m r_i^q, \quad \text{for any } y \in \mathbb{R}^m.$$

Definition 3.7. The *(negative) entropy regularization* of the maximum, parametrized by $\eta > 0$, is the function smax $: \mathbb{R}^m \to \mathbb{R}$ such that

$$\mathsf{smax}_\eta(y) := \max_{r \in \Delta_m} \langle r, y \rangle - \frac{1}{\eta} \sum_{i=1}^m r_i \log r_i, \quad \text{for any } y \in \mathbb{R}^m.$$

It is not hard to see that in this case, the solution of the maximization problem can be written in closed form: $\mathsf{smax}_\eta(y) = \frac{1}{\eta} \log \left(\sum_{i=1}^m \exp(\eta y_i) \right)$, thereby recovering the usual formulation of the softmax function.

Regret minimization interpretation. These regularization ideas have originated in the online learning community. To see how this is related to discrepancy, let us make the following thought experiment. We play an online game against an adversary, where we select at each step some $r_t \in \Delta_n$, and after that some δ_t is revealed. Our goal is to minimize the regret, which is the difference between the best static cost in hindsight, $\|A(\delta_1 + \ldots + \delta_T)\|_\infty$ and our cost, $\sum_{t \leqslant T} \langle r_t, A\delta_t \rangle$.

We can play different strategies that are robust to the future choices of δ_t. For each of these, the optimal strategy for the adversary is to follow a particular potential function that is obtained by adding a regularizer to the optimization problem.

We note that this is analogous to what Allen-Zhu, Liao, and Orecchia have proposed for graph sparsification [AZLO15]. They identified a similar online game on density matrices and used it to interpret the construction of [BSS14] as a follow-the-regularized-leader strategy. While they use the same $\ell_{1/2}$-regularizer as we did for Spencer's theorem in Section 2, the connection is more subtle, as their setting is crafted specifically for matrices with positive updates, which involves deriving a set of different bounds that charge the entire change in potential function to the first-order term.

3.3 Regularization bounds

We now present our two main technical lemmas that give an analytic justification for the ℓ_q and negative entropy regularization. The first one (Lemma 3.8) estimates the additive error incurred when tracking the regularized version of the maximum instead of the true maximum. For constant η and q, the approximation is worse for ℓ_q-regularization than for negative entropy regularization (polynomial vs logarithmic in the size of the vector).

Lemma 3.8. *Let $y \in \mathbb{R}^m$ and $q \in (0,1)$. If $M(y) := \max_{1 \leqslant i \leqslant m} y_i$,*

$$M(y) \leqslant \omega_{q,\eta}^*(y) \leqslant M(y) + \frac{m^{1-q}}{\eta q} \ \text{ and } \ M(y) \leqslant \mathsf{smax}_\eta(y) \leqslant M(y) + \frac{\log m}{\eta} \ .$$

Proof. The lower bounds follow from picking r to be the Dirac mass function centered on the maximum coordinate. For the upper bounds, note that on the one hand, for all $r \in \Delta_m$, $\langle r, y \rangle \leqslant M(y)$, and on the other hand, $\sum_i r_i^q \leqslant m^{1-q}$ (resp. $-\sum_i r_i \log r_i \leqslant \log m$) by Jensen's inequality. $\square$

The second one (Lemma 3.9) bounds the first two terms in the Taylor expansion of the potential function. In the sequel, this will allow us to control the increase in ℓ_∞-norm when making a small update in our iterative algorithm. As we demonstrated in Section 2, what matters in this expansion is the second-order term. Indeed, in applications to discrepancy, we will always trivially upper bound the first-order term by simply picking an update that is positively correlated with the gradient (which will be an easy additional condition to impose).

Lemma 3.9. *Fix $y \in \mathbb{R}^m$ and $q \in (0,1)$. Let $\nabla := \nabla \omega_{q,\eta}^*(y)$. Then $\nabla \in \Delta_m$ and for all $\delta \in \mathbb{R}^m$ with $\|\delta\|_\infty \leqslant \frac{1-q}{8\eta}$,*

$$\omega_{q,\eta}^*(y + \delta) \leqslant \omega_{q,\eta}^*(y) + \langle \nabla, \delta \rangle + \frac{\eta}{1-q} \sum_{i=1}^m \nabla_i^{2-q} \delta_i^2 \ .$$

Similarly, if $\nabla := \nabla \mathsf{smax}_\eta(y)$, then $\nabla \in \Delta_m$ and for all $\delta \in \mathbb{R}^m$ with $\|\delta\|_\infty \leqslant \frac{1}{3\eta}$,

$$\mathsf{smax}_\eta(y + \delta) \leqslant \mathsf{smax}_\eta(y) + \langle \nabla, \delta \rangle + \eta \sum_{i=1}^m \nabla_i \delta_i^2 \ .$$

Proof. Consider first the ℓ_q-regularizer with $\eta = 1$. To lighten notations we write ω^* for $\omega_{q,\eta}^*$. Recall that

$$\omega^*(y) = \max_{r \in \Delta_m} \langle r, y \rangle + \frac{1}{q} \sum_{i=1}^m r_i^q \ . \tag{2}$$

By Danskin's theorem (see e.g. [Ber99, Proposition B.25]), we have $\nabla\omega^*(y) = r^* \in \Delta_m$, where r^* is the optimum in (2). For the KKT conditions to hold, we must have for some $\lambda : \mathbb{R}^m \to \mathbb{R}$ (the Lagrange multiplier associated to the equality constraint of the simplex): $y_i + (r_i^*)^{q-1} = \lambda(y)$ for all $i \in [m]$. Note that the Lagrange multipliers associated to the inequality constraints disappear by complementary slackness since necessarily $r_i^* \neq 0$. Also we must have $\lambda(y) > \max_{i \in [m]} y_i$ by the previous equality. In fact, $\lambda(y)$ is the unique solution to $\sum_{i \in [m]} (\lambda(y) - y_i)^{1/(q-1)} = 1$.

In summary, $\nabla\omega^*(y) = (\lambda(y)\mathbf{1} - y)^{\odot \frac{1}{q-1}} \in \Delta_m$. Differentiating once more, we see that

$$\nabla^2\omega^*(y) = \frac{1}{1-q}\left(\text{diag}(\nabla\omega^*(y)^{\odot 2-q}) - (\nabla\lambda(y))(\nabla\omega^*(y)^{\odot 2-q})^T\right).$$

Let $M := (\nabla\lambda(y))(\nabla\omega^*(y)^{\odot 2-q})^T$. Observe that M has rank 1 and must be symmetric as the Hessian itself is symmetric. Further, $\lambda(y)$ is a nondecreasing function of y_i for all $i \in [m]$, so that every entry of M is nonnegative. It follows that M is positive semidefinite, and thus

$$\nabla^2\omega^*(y) \preccurlyeq \frac{1}{1-q}\,\text{diag}(\nabla\omega^*(y)^{\odot 2-q}). \tag{3}$$

Now fix $\delta \in \mathbb{R}^m$. The function $s \mapsto \sum_i (s - y_i)^{\frac{1}{q-1}}$ defined for $s > \max_i y_i$ is nonincreasing, so for all i,

$$|\lambda(y + \delta) - \lambda(y)| \leqslant \|\delta\|_\infty \text{ and } \lambda(y) \geqslant 1 + y_i. \tag{4}$$

Now fix $i \in [m]$ and suppose that δ satisfies $\|\delta\|_\infty \leqslant \frac{1-q}{8}$. We write

$$(\nabla\omega^*(y + \delta))_i^{2-q} = (\nabla\omega^*(y))_i^{2-q}\left(1 + \frac{\lambda(y + \delta) - \lambda(y) - \delta_i}{\lambda(y) - y_i}\right)^{\frac{2-q}{q-1}}$$
$$\leqslant (\nabla\omega^*(y))_i^{2-q}\exp\left(\frac{2-q}{1-q}\frac{\lambda(y) + \delta_i - \lambda(y + \delta)}{\lambda(y + \delta) - y_i - \delta_i}\right),$$

where we used the inequality $\log(1 + y) \geqslant \frac{y}{1+y}$. Now we plug in the inequalities (4):

$$(\nabla\omega^*(y + \delta))_i^{2-q} \leqslant (\nabla\omega^*(y))_i^{2-q}\exp\left(\frac{2-q}{1-q}\frac{\lambda(y) + \delta_i - \lambda(y + \delta)}{1 + \lambda(y + \delta) - \lambda(y) - \delta_i}\right)$$
$$\leqslant (\nabla\omega^*(y))_i^{2-q}\exp\left(\frac{2}{1-q}\frac{2\|\delta\|_\infty}{1 - 2\|\delta\|_\infty}\right)$$
$$\leqslant 2(\nabla\omega^*(y))_i^{2-q}. \tag{5}$$

Finally, from Taylor's inequality, under the same assumption $\|\delta\|_\infty \leqslant \frac{1-q}{8}$,

$$|\omega^*(y + \delta) - \omega^*(y) - \langle\nabla\omega^*(y), \delta\rangle| \leqslant \frac{1}{2}\sup_{u \in [0,1]}\left|\delta^T\nabla^2\omega^*(y + u\delta)\delta\right| \leqslant \frac{1}{1-q}\sum_{i=1}^m (\nabla\omega^*(y))_i^{2-q}\delta_i^2,$$

where the last inequality follows from (3) and (5).

For the entropy regularizer and $\eta = 1$, it holds that

$$\nabla\text{smax}(y) = \frac{\exp(y)}{\sum_{i=1}^m \exp(y_i)} \text{ and } \nabla^2\text{smax}(y) \preccurlyeq \text{diag}(\nabla\text{smax}(y)).$$

Therefore for all $i \in [m]$, $(\nabla\text{smax}(y + \delta))_i \leqslant (\nabla\text{smax}(y))_i\exp(2\|\delta\|_\infty)$, and we conclude in the same way as for the ℓ_q-regularizers.

For general η, observe that $\nabla \omega_{q,\eta}^*(y) = \nabla \omega_{q,1}^*(\eta y)$ and $\nabla^2 \omega_{q,\eta}^*(y) = \eta \nabla^2 \omega_{q,1}^*(\eta y)$ (and similarly for smax_η). Therefore, the same argument based on Taylor's inequality gives the desired result as long as $\|\delta\|_\infty \leqslant \frac{1-q}{8\eta}$ for $\omega_{q,\eta}^*$ and $\|\delta\|_\infty \leqslant \frac{1}{3\eta}$ for smax_η. $\qquad\square$

Remark 3.10. This gives an analytic explanation for why we might prefer ℓ_q-regularization to negative entropy regularization in certain situations, although the approximation error from Lemma 3.8 is worse (for the same value of η). Observe that a typical entry ∇_i of the gradient is much smaller that 1. Hence ℓ_q-regularization can be advantageous whenever we can leverage the fact that ∇_i^{2-q} is typically much smaller than ∇_i. As we will now see, this is the case in Spencer's setting.

4 Spencer's setting

We now focus on the setting of Spencer's theorem (Theorem 2.1), namely the discrepancy of matrices with bounded entries. Our goal in this section is twofold: first, we give a rigorous version of the proof of Spencer's theorem that we sketched in Section 2. Then, we show how to improve the constant with a slightly more careful analysis.

4.1 Full proof of Spencer's theorem

We give a complete proof of Spencer's theorem in the general case where the matrix has m rows and n columns. Our choice of $q \in (0,1)$ in the ℓ_q-regularization is going to depend on the ratio m/n.

Theorem 4.1. *Let $n \leqslant m$. There is a deterministic algorithm running in polynomial time that for each $A \in [-1,1]^{m \times n}$, finds $x \in \{\pm 1\}^n$ such that*

$$\|Ax\|_\infty = O\left(\sqrt{n \log\left(\frac{2m}{n}\right)}\right).$$

We start by proving the following lemma, which will allow us to find an update vector that does not increase too much the ℓ_q-regularization of the maximal coordinate when there are k active coordinates remaining.

Lemma 4.2. *Let k, m be such that $4 \leqslant k \leqslant 2m - 2$. Let $u_1, \ldots, u_m$ be unit vectors in $\mathbb{R}^k$ and $\nabla \in \Delta_m$. Consider*

$$M := \sum_{i=1}^m \nabla_i^{2-q} u_i u_i^T.$$

There is a subspace S of dimension at least 2 such that for all $v \in S$,

$$v^T M v \leqslant 8k^{q-2} \|v\|_2^2.$$

Moreover, this subspace can be found efficiently.

Proof. Without loss of generality, suppose that $\nabla_1 \geqslant \ldots \geqslant \nabla_m$. Let

$$S_1 := \left\{ v \in \mathbb{R}^k : \langle v, u_i \rangle = 0 \text{ for all } i = 1, \ldots, \left\lceil \frac{k}{2} \right\rceil - 1 \right\}.$$

Observe that $\nabla_{\lceil \frac{k}{2} \rceil} \leqslant \frac{2}{k}$, so for all $v \in S_1$,

$$v^T M v \leqslant \left(\frac{2}{k} \right)^{1-q} \sum_{i=1}^{m} \nabla_i \langle u_i, v \rangle^2. \tag{6}$$

Let $R := \sum_{i \leqslant m} \nabla_i u_i u_i^T$ and consider an orthonormal basis $w_1, \ldots, w_l$ of S_1 such that $w_1^T R w_1 \leqslant \ldots \leqslant w_l^T R w_l$. Select S to be the span of $\{w_1, w_2\}$. Observe that

$$\sum_{j=1}^{l} w_j^T R w_j = \sum_{i=1}^{m} \nabla_i \sum_{j=1}^{l} \langle w_j, u_i \rangle^2 \leqslant 1.$$

Thus, by an averaging argument, it holds that

$$v^T R v \leqslant \frac{1}{l-1} \|v\|_2^2 \leqslant \frac{2}{k-2} \|v\|_2^2$$

for all $v \in S$. We conclude by combining this with (6) and using the assumption on k. $\qquad\square$

Proof of Theorem 4.1. We first double all the rows of A and consider the matrix $\begin{bmatrix} A \\ -A \end{bmatrix}$. Thus, we assume without loss of generality that we are given a $2m \times n$ matrix A such that for all $x \in \{\pm 1\}^n$, $\|Ax\|_\infty = \max_i (Ax)_i$.

We set the parameter L of Algorithm 2 to be $L := \frac{1-q}{8\eta n}$, where q and η are the parameters of the ℓ_q-regularizer to be fixed later.

We now describe our construction of $\mathsf{oracle}(A, x(t))$ with $F(t)$ being the set of active coordinates of $x(t)$ and $k = k(t) := |F(t)|$. To simplify notations, we write $x = x(t)$ and $F = F(t)$. Observe that $\|A\delta\|_\infty \leqslant n\|\delta\|_\infty$. Hence, by Lemma 3.9, there exists $\nabla \in \Delta_n$ such that for all update $\delta \in \mathbb{R}^n$ with $\|\delta\|_\infty \leqslant L$,

$$\omega_{q,\eta}^*(A(x+\delta)) - \omega_{q,\eta}^*(Ax) \leqslant \langle \nabla, A\delta \rangle + \frac{\eta}{1-q} \sum_{i=1}^{n} \nabla_i^{2-q} \langle A_i, \delta \rangle^2.$$

By assumption, $\sum_{j \in F} A_{i,j}^2 \leqslant k$, so we can apply Lemma 4.2 to get a 2-dimensional subspace S such that for all $\delta \in S$,

$$\sum_{i=1}^{n} \nabla_i^{2-q} \langle A_i, \delta \rangle^2 \leqslant \frac{4\|\delta\|_2^2}{k^{1-q}}. \tag{7}$$

The second-order term is invariant if we change δ to $-\delta$, but the first-order term changes sign. We return from $\mathsf{oracle}(A, x(t))$ the subspace S intersected with the halfspace $\{\delta \in \mathbb{R}^n : \langle A^T \nabla, \delta \rangle \leqslant 0\}$.

Now we switch to the global analysis of Algorithm 2 and estimate what is the total discrepancy incurred over the whole walk. Assumption 3.1 is here satisfied for $C = 3$, so since the entries of A are bounded, the last step of Algorithm 2 only changes the discrepancy of the final coloring by an additive constant.

Denote by β_k the sum of the ℓ_2-squared norm of the update vectors starting from the point where there are at most k unfrozen coordinates remaining. Recall that we always choose our update

vector orthogonal to the current position, so that $\beta_k \leqslant k$. We now sum by parts the main term of the increases (7) over the execution of the algorithm,

$$\sum_{k=4}^{n} \frac{\beta_k - \beta_{k-1}}{k^{1-q}} = \frac{\beta_n}{n^{1-q}} + \sum_{k=4}^{n-1} \beta_k \left(\frac{1}{k^{1-q}} - \frac{1}{(k+1)^{1-q}} \right) \leqslant n^q + \sum_{k=4}^{n-1} \frac{1}{k^{1-q}} \leqslant \frac{2n^q}{q}. \tag{8}$$

Thus, by (8) and Lemma 3.8, the final coloring $x(T)$ satisfies

$$\|Ax(T)\|_\infty \leqslant \omega_{q,\eta}^*(Ax(T)) \leqslant \frac{(2m)^{1-q}}{\eta q} + \frac{8\eta}{q(1-q)} n^q.$$

The result follows by setting

$$\eta = \sqrt{\frac{(1-q)m^{1-q}}{n^q}} \text{ and } q = 1 - \frac{1}{\log\left(\frac{2m}{n}\right)}. \qquad \square$$

Remark 4.3. At this point, it is worth looking at what happens in this proof if we replace the ℓ_q-regularizer with the entropic regularizer. For simplicity, consider the case where $m = O(n)$. While the constant cost is only $\log n/\eta$, we are not able to win anything in the local update as in Lemma 4.2 and we would get an ηn loss in the potential during the walk. Optimizing over η would give discrepancy $\sqrt{n \log n}$. Negative entropy regularization in this context corresponds merely to a derandomization of the Chernoff and union bound argument.

In fact, one could repeat the same analysis by replacing the regularizer by a general function of the form $\phi(r) = \sum_{i \leqslant n} \varphi(r_i)$ for some convex, non-positive function $\varphi : \mathbb{R} \to \mathbb{R}$. Under additional conditions on φ (for example the fact that $x \mapsto x\varphi''(x)$ is non-increasing) one would obtain a discrepancy of

$$O\left(\sqrt{-n\varphi\left(\frac{1}{n}\right) \sum_{k \leqslant n} \frac{k}{\varphi''\left(\frac{1}{k}\right)}} \right). \tag{9}$$

With this bound established, we can quickly verify that setting φ to be the negative entropy, we obtain $\varphi(1/n) = -\log n/n$ and $\varphi''(1/k) = k$, which immediately recovers a discrepancy of $O(\sqrt{n \log n})$.

Given this expression in (9), it appears that we can derive the best possible regularizer by solving a differential equation. Since there is no silver bullet for such problems, one can simply test various elementary functions. Setting $\varphi(x) = -x^q$ for $0 < q < 1$ we verify the required condition and obtain $\varphi(1/n) = -1/n^q$, and $\varphi''(1/k) = q(1-q)/k^{2-q}$, which removes the logarithmic factor for constant q.

Remark 4.4 (Spherical discrepancy). A slight variation of the same algorithm, which does not freeze variables, automatically achieves optimal bounds for *spherical discrepancy*. This setting is a relaxation of the Komlós problem, where the columns of the input matrix are vectors with at most unit ℓ_2-norm, but the sought coloring only has an ℓ_2-norm constraint i.e. $\|x\|_2 = \sqrt{n}$, rather than $x \in \{\pm 1\}^n$ [JM20]. The key difference between this setting and that of Komlós is that we are not forced to lose degrees of freedom by freezing variables, so throughout the entire execution of the algorithm we have $\Theta(n)$ degrees of freedom to update the partial coloring.

To show this, we simply observe that at all times there is an update that does not increase the discrepancy of rows with large *global* ℓ_2 norm, which represent only at most a constant fraction of

the entire set of rows, by Markov's inequality. The rate of increase in discrepancy entirely depends on the ℓ_2-norm of the rows of the underlying matrix (restricted to the unfrozen variables, which in this case are all the variables). Following through with the same argument we used for Spencer, we obtain discrepancy $O(1)$.

4.2   Towards a better constant for Spencer's theorem

In his seminal paper, Spencer proved that in the $m = n$ setting, Theorem 2.1 holds with constant $K = 5.32$. In this section, we improve it to $K = 3\sqrt{3/2} \approx 3.675$. Furthermore, we can find a coloring achieving the constant $K = 3\sqrt{3/2} + \varepsilon$ in polynomial time for every fixed $\varepsilon > 0$. In [Bel13], the result with $K = 3.65$ is claimed but the proof does not immediately correspond to an efficient algorithm and some computations rely on personal communication.

Our strategy for obtaining the theorem with $K = 3.675$ is to repeat the argument of Section 4.1 by tracking more carefully all the constants. We start by giving an analog of Lemma 4.2 with a tighter leading constant.

Lemma 4.5. *Let $q \in (0,1)$, $u_1, \ldots, u_m$ be unit vectors in $\mathbb{R}^k$, and $\nabla \in \Delta_m$. Consider*

$$M := \sum_{i=1}^m \nabla_i^{2-q} u_i u_i^T.$$

There is a subspace S of dimension 2 such that for all $v \in S$ with $\|v\|_2 = 1$,

$$v^T M v \leqslant k^{q-2} + O(k^{q-3}).$$

Moreover, for all fixed $\varepsilon > 0$, there is a randomized algorithm running in time polynomial in n that with high probability returns a 2-dimensional subspace S such that for all $v \in S$ with $\|v\|_2 = 1$,

$$v^T M v \leqslant (1 + \varepsilon) k^{q-2} + O(k^{q-3}).$$

Proof. Assume without loss of generality that $\nabla_1 \geqslant \ldots \geqslant \nabla_m$. Sample α uniformly in the interval $(\frac{1}{2}, 1)$. We will prove that

$$f(\nabla) := \mathbf{E}_\alpha \left[\sum_{i \geqslant \lfloor \alpha k \rfloor} \nabla_i^{2-q} \right] \leqslant \frac{k^{q-1}}{4} + O(k^{q-2}) = \mathbf{E}_\alpha[(1-\alpha)k^{q-1}] + O(k^{q-2}). \tag{10}$$

Let us first see why it implies the desired result. Let $\alpha \in (\frac{1}{2}, 1)$ be such that $\frac{1}{(1-\alpha)k} \sum_{i \geqslant \lfloor \alpha k \rfloor} \nabla_i^{2-q} \leqslant k^{q-2} + O(k^{q-3})$. Then we can repeat the proof of Lemma 4.2 to get a 2-dimensional subspace S such that for all $v \in S$,
$$v^T M v \leqslant k^{q-2} + O(k^{q-3}).$$

Furthermore, if ε is fixed, the corresponding α can be found with high probability by repeating the experiment and using Markov's inequality.

Now we prove (10). We can compute

$$f(\nabla) = 2 \int_{\frac{1}{2}}^1 \sum_{i \geqslant \lfloor \alpha k \rfloor} \nabla_i^{2-q} \, d\alpha = \sum_{i \geqslant \lfloor \frac{k}{2} \rfloor} \left(\frac{2(i+1)}{k} - 1 \right) \nabla_i^{2-q}.$$

15

Let $P = \{\nabla \in \Delta_m : \nabla_1 \geqslant \ldots \geqslant \nabla_m\}$. We can see that $f : P \to \mathbb{R}$ is a convex function so it attains its maximum at an extreme point of P. The extreme points of P are of the form $z_l := \left(\frac{1}{l} \ldots \frac{1}{l} 0 \ldots 0\right)$ (with exactly l nonzero coordinates) for some $l \in [m]$. Moreover, the maximum of f has to be attained when $l = \gamma k$, with $\gamma \in [\frac{1}{2}, 1]$. However, in that case,

$$f(z_l) = l^{q-2}\left(\frac{l(l+1) - \frac{k}{2}(\frac{k}{2}+1)}{k} - \left(l - \frac{k}{2} + 1\right) + O(1)\right) = k^{q-1}\left(\sqrt{\gamma} - \frac{1}{\sqrt{\gamma}} + \frac{1}{4\gamma^{3/2}}\right) + O(k^{q-2}).$$

Finally, $\gamma \mapsto \sqrt{\gamma} - \frac{1}{\sqrt{\gamma}} + \frac{1}{4\gamma^{3/2}}$ is increasing on $[\frac{1}{2}, 1]$, with maximum equal to $\frac{1}{4}$ for $\gamma = 1$. $\qquad\square$

We also remark that we can improve the constant in front of the second-order term in Lemma 3.9.

Lemma 4.6. *There exists universal constants $0 < C_1, C_2 \leqslant 1$ such that if $\nabla := \nabla \omega_{q,\eta}^*(y)$, then for all $\delta \in \mathbb{R}^n$ with $\|\delta\|_\infty \leqslant C_1 \frac{1-q}{n\eta}$,*

$$\omega_{q,\eta}^*(y+\delta) \leqslant \omega_{q,\eta}^*(y) + \langle \nabla, \delta \rangle + \frac{\eta}{2(1-q)}\left(1 + \frac{C_2}{n}\right)\sum_{i=1}^{n} \nabla_i^{2-q}\delta_i^2.$$

Proof. The proof is identical to the proof of Lemma 3.9. We simply replace (5) by the stronger inequality following from the stronger assumption on $\|\delta\|_\infty$. $\qquad\square$

Theorem 4.7. *For every $A \in [-1,1]^{n \times n}$, there exists $x \in \{\pm 1\}^n$ such that $\|Ax\|_\infty \leqslant 3\sqrt{\frac{3n}{2}} + O(1)$. Moreover, for any fixed $\varepsilon > 0$, there is a randomized algorithm running in polynomial time to find $x \in \{\pm 1\}^n$ such that $\|Ax\|_\infty \leqslant (3\sqrt{\frac{3}{2}} + \varepsilon)\sqrt{n} + O(1)$.*

Proof. The proof is similar to our proof of Theorem 4.1. We take the parameter of Algorithm 2 to be $L = \frac{C_1}{4n^2}$, where C_1 is the constant from Lemma 4.6. Let $\omega^* := \omega_{q,\eta}^*$, for some parameters q and η to pick later.

We now explain how to construct $\mathsf{oracle}(A, x(t))$, with $k := |F(t)|$. Let $t_i = t_i(t) < t$ be the last step of the algorithm for which $\langle A_i, x(t_i)\rangle$ has a different sign from $\langle A_i, x(t)\rangle$. We define $\pi_t : \mathbb{R}^n \to \mathbb{R}^n$ such that for all $i \in [n]$ and $x \in \mathbb{R}^n$

$$\pi_t(x)_i := \begin{cases} \langle +A_i, x - x(t_i + 1)\rangle & \text{if } \langle A_i, x\rangle > 0 \\ \langle -A_i, x - x(t_i + 1)\rangle & \text{otherwise} \end{cases} \tag{11}$$

We apply Lemma 4.6 to get that there is some $\nabla \in \Delta_n$ such that for all $\delta \in \mathbb{R}^n$ with $\|\delta\|_\infty \leqslant L$,

$$\omega^*(\pi_t(x(t) + \delta)) - \omega^*(\pi_t(x(t))) \leqslant u(\delta) + \frac{\eta}{2(1-q)}\left(1 + \frac{C_2}{n}\right)\sum_{i=1}^{n}\nabla_i^{2-q}\langle A_i, \delta\rangle^2,$$

where $u : \mathbb{R}^n \to \mathbb{R}$ is a linear form. We then use Lemma 4.5 to find a 2-dimensional subspace S such that if $\delta \in S$ and $\|\delta\|_2 \leqslant 1$,

$$\sum_{i=1}^{n}\nabla_i^{2-q}\langle A_i, \delta\rangle^2 \leqslant \|\delta\|_2^2 k^{q-1} + O(k^{q-2}).$$

We also pick a signing of $\pm\delta$ to make $u(\pm\delta) \leqslant 0$. Moreover,

$$\omega^*(\pi_{t+1}(x(t) \pm \delta)) \leqslant \omega^*(\pi_t(x(t) \pm \delta))$$

16

since setting the discrepancy proxy of some rows to zero only decreases the value of ω^*.

Then by a similar argument to the one in Section 4.1, we can ensure that the total increase $\omega^*(\pi_T(x(T))) - \omega^*(0)$ over the duration of the walk is at most $\frac{\eta n^q}{2q(1-q)} + O(1)$. Hence,

$$\|Ax(T)\|_\infty \leqslant nL + \|\pi_T(x(T))\|_\infty \leqslant \frac{n^{1-q}}{q\eta} + \frac{\eta n^q}{2q(1-q)} + O(1).$$

Optimizing over $q \in (0,1)$ and $\eta > 0$, we get for $q = 2/3$ and $\eta = \sqrt{3/2}$ that indeed $\|Ax(T)\|_\infty \leqslant 2\sqrt{3n/2} + O(1)$. Furthermore, the corresponding statement with ε follows from using instead the constructive part of Lemma 4.5. $\square$

We do not expect this analysis to be tight – for example, we have not made use of the first-order term. There is still a substantial gap with the best known lower bound of $K \geqslant 1$, obtained for Hadamard matrices.

5 New discrepancy bounds for pseudorandom instances

An advantage of approaches for discrepancy minimization based on potential functions is that it is easy to track two (or a constant number of) potentials in parallel. Up to changing a few constants in the analysis, we can ensure that the best of both worlds happens. In this section, we illustrate this idea by proving a new discrepancy bound for instances of Beck-Fiala and Komlós conjectures that satisfy a certain pseudorandomness condition.

Following [Pot20], we define the quantity $\lambda(A)$ associated to a matrix $A \in \mathbb{R}^{m \times n}$ as follows:

$$\lambda(A) := \sup_{\|u\|_2=1, \langle u, \mathbf{1} \rangle = 0} \|Bu\|_2, \text{ where } B_{ij} := A_{ij}^2 \text{ for all } i \in [m], j \in [n].$$

In the special case where A is the adjacency matrix of a d-regular graph, $\lambda(A)$ is the second largest eigenvalue of A and is bounded by d. As observed in [Pot20], $\lambda(A)$ is typically much smaller than this worst-case bound when A is the incidence matrix of a *random* regular set system. We will also check in Section 5.5 that this still holds for natural random instances of Komlós conjecture.

We now recall Potukuchi's result for pseudorandom Beck-Fiala instances.

Theorem 5.1 (Theorem 1.1 in [Pot20])**.** *Let $A \in \{0,1\}^{m \times n}$ be such that each column has at most s nonzero entries. Then there is a randomized algorithm running in polynomial time to find $x \in \{\pm 1\}^n$ such that*
$$\|Ax\|_\infty = O(\sqrt{s} + \lambda(A)).$$

The algorithm behind Theorem 5.1 relies on iteratively running the random walk of Lovett and Meka [LM15] to obtain a good partial coloring on the rows that behave "randomly" on the set of current active coordinates. When $\lambda(A)$ is small, one can make sure that the rows behave "randomly" as long as their ℓ_2-mass on the active coordinates is large enough. However, the bound on the discrepancy of the rows after their ℓ_2-mass has become small crucially uses the fact that the instance is $\{0,1\}$-valued.

Our contributions here are new discrepancy bounds that generalize both Theorem 5.1 and Banaszczyk's bound [Ban98, BDG19]. Moreover, they hold both in the Beck-Fiala setting *and* in the Komlós setting. In Section 5.5, we will apply these results to deduce random versions of Komlós

conjecture. Next, we state the two theorems that we will prove in this section. For simplicity, we focus on the case of square matrices ($m = n$), although we naturally expect the techniques to generalize to the $m \geqslant n$ case.

Theorem 5.2 (Bound for pseudorandom Komlós instances). *Let $A \in \mathbb{R}^{n \times n}$ be such that each column has ℓ_2-norm at most 1. Then there is a deterministic, polynomial-time algorithm to find $x \in \{\pm 1\}^n$ such that*

$$\|Ax\|_\infty = O(1 + \sqrt{\lambda(A) \log n}).$$

Theorem 5.3 (Bound for pseudorandom Beck-Fiala instances). *Let $A \in \{0, \pm 1\}^{n \times n}$ be such that each column has at most s nonzero entries. Then there is a deterministic, polynomial-time algorithm to find $x \in \{\pm 1\}^n$ such that*

$$\|Ax\|_\infty = O(\sqrt{s} + \min(\sqrt{\lambda(A) \log n}, \lambda(A))).$$

As immediate corollaries, Theorem 5.2 implies Komlós conjecture when $\lambda(A) = O(\frac{1}{\log n})$ and Theorem 5.3 implies Beck-Fiala conjecture when $\lambda(A) = O(\sqrt{s} + \frac{s}{\log n})$. This strictly improves Theorem 5.1 for Beck-Fiala instances in the regime $\lambda(A) \in [O(\log n), O(s)]$. Furthermore, to the best of the authors' knowledge, Theorem 5.2 is the first result for pseudorandom instances of Komlós conjecture.

Remark 5.4. The mere column-sparsity assumption in Theorem 5.3 does not suffice to ensure that $\lambda(A) \leqslant s$.[1] However, as we will see (Remark 5.16), we can essentially replace $\lambda(A)$ by $\min(\lambda(A), s)$ in the analysis. In this sense, our algorithm also matches Banaszczyk's bound.

5.1 Proof strategy and notations

We start by giving some idea of how we will use the fact that $\lambda(A)$ is small in our discrepancy framework. The main insight of [Pot20] is that at any point in time in a discrepancy walk, we can control the ℓ_2-mass restricted to active coordinates of all rows simultaneously.

Lemma 5.5 (Lemma 2.3 in [Pot20]). *Let $A \in \mathbb{R}^{n \times n}$ and $F \subseteq [n]$ be of size k. Then, for any constant $D > 0$, there exists a subset $S \subseteq [n]$ such that $|S| \leqslant k/D^2$ and for any $i \notin S$,*

$$\sum_{j \in F} A_{ij}^2 \leqslant \frac{k}{n} \sum_{j=1}^n A_{ij}^2 + D\lambda(A).$$

Intuitively, the term $\frac{k}{n} \sum_{j \leqslant n} A_{ij}^2$ corresponds to the ℓ_2-squared-mass we would expect the row A_i to have if the set of active coordinates F were picked at random. The parameter $\lambda(A)$ gives a bound on the deviation from this random behavior. In particular, the ℓ_2-mass of a row essentially decreases as in the average case as long as it is $\Omega(\sqrt{\lambda(A)})$. We recall the proof of Lemma 5.5 for completeness.

---

[1] For example, consider a vector v with half $+1$ and half -1 entries, take the first row of A to be v and fill the other rows with zeros. A has one nonzero entry per column but $\|Av\|_2 = \sqrt{n}\|v\|_2$.

Proof. Let us denote $B_{ij} := A_{ij}^2$ for all $i, j \in [n]$. If $u \in \mathbb{R}^n$ is a vector orthogonal to $\mathbf{1}$, then by definition we have $\sum_i \langle B_i, u \rangle^2 \leqslant \lambda(A)^2 \|u\|_2^2$. Consider $u \in \mathbb{R}^n$ such that $u_j := 1 - \frac{k}{n}$ if $j \in F$ and $u_j := -\frac{k}{n}$ if $j \notin F$. Then u is orthogonal to $\mathbf{1}$ and $\|u\|_2^2 \leqslant k$. Hence,

$$\sum_{i=1}^n \left(\sum_{j \in F} A_{ij}^2 - \frac{k}{n} \sum_{j=1}^n A_{ij}^2 \right)^2 \leqslant \lambda(A)^2 k \,.$$

The result follows from a simple counting argument. $\qquad\square$

Roadmap of the proof. In order to prove Theorem 5.2 and Theorem 5.3, we will track two different types of potential functions, depending on which of the main term or the error term in Lemma 5.5 dominates. Section 5.2 will be devoted to bounding the discrepancy incurred in the regime where the row mass decreases as if the input were random. The analysis here will mirror our proof of Spencer's theorem. In Section 5.3, we will consider the case where the error term dominates. There we leverage the fact that the row mass has become small. In this setting, we will use a potential function that was introduced in [LRR17, Appendix B] to recover Banaszczyk's bound with the multiplicative weights update method.

Before starting the proof, we introduce some useful concepts and notations. From now on, we fix a matrix $A \in \mathbb{R}^{n \times n}$ with column ℓ_2-norm bounded by 1. With the context being clear, we will write $\lambda := \lambda(A)$. Our algorithm will follow the structure of the meta-algorithm Algorithm 2, therefore we will use our usual notations: $x(t)$ for the coloring at time t, $F(t)$ for the active coordinates at time t, etc.

Definition 5.6. For each row $i \in [n]$, we define

$$t_i := \min \left\{ t \geqslant 0 : \sum_{j \in F(t)} A_{ij}^2 \leqslant 8\lambda \right\} \,.$$

In words, t_i represents the time at which the i-th row stops behaving as if it were random. We will write $P(t) := \{i \in [n] : t \leqslant t_i\}$ for the set of "pseudorandom" rows at time t. The following observation explains what we mean by "pseudorandom" – we can pretend as if the freezing process decreases linearly the ℓ_2-mass of the rows. It is an easy consequence of Lemma 5.5.

Claim 5.7. Fix any time step $t \geqslant 0$. There exists a subset of rows $I = I(t) \subseteq [n]$ such that $|I| \leqslant \frac{|F(t)|}{16}$ and for any $i \in P(t), i \notin I$:

$$\sum_{j \in F(t)} A_{ij}^2 \leqslant \frac{2|F(t)|}{n} \sum_{j=1}^n A_{ij}^2 \,.$$

For any row $i \in [n]$, we will track separately the contributions to its discrepancy for $t \leqslant t_i$ and $t > t_i$.

Random regime. Similarly to [Pot20], we group together the rows that have similar *total ℓ_2-mass*. For any $r \in \{1, \ldots, \lceil \log_2 n \rceil\}$, let

$$R_r := \left\{ i \in [n] : \sum_{j=1}^n A_{ij}^2 \in (2^{r-1}, 2^r] \right\} \,.$$

19

We also consider $R_0 := \{i \in [n] : \sum_{j=1}^n A_{ij}^2 \leq 1\}$. An easy double counting argument bounds the size of each R_r: $|R_r| \leq n2^{1-r}$ for any $r \leq \lceil \log_2 n \rceil$. Our strategy will be to play several Spencer's games in parallel (restricted to the rows in R_r, for each value of r) and carefully allocate some "effective dimension" to each of them at any step of the walk.

We now define $\pi_{r,t} : \mathbb{R}^n \to \mathbb{R}^n$ to be a projection to the coordinates of R_r of the discrepancy of the rows that behave pseudorandomly. Once $t > t_i$, we keep tracking in the i-th row the same value $\langle A_i, x(t_i) \rangle$. In short, for any $x \in \mathbb{R}^n$ and $i \in [n]$,

$$(\pi_{r,t}(x))_i = \begin{cases} 0 & \text{if } i \notin R_r \\ \langle A_i, x(t_i) \rangle & \text{if } i \in R_r \text{ but } i \notin P(t) \\ \langle A_i, x \rangle & \text{if } i \in R_r \cap P(t) \end{cases}$$

Now we are able to define our potential functions in the random regime. For any $r = 0, \ldots, \lceil \log_2 n \rceil$, let

$$\Phi_{r,t}(x) := \omega^*_{\frac{1}{2}, \sqrt{n2^{1-r}}}(\pi_{r,t}(x)) .$$

The choice of $\eta_r = \sqrt{n2^{1-r}}$ as regularization parameter can be justified by the fact that there are at most $n2^{1-r}$ rows in the r-th group – so this is essentially the smallest value of η_r that makes the additive approximation error of the regularized maximum $O(1)$ (which is our target discrepancy in this regime).

Our main lemma states that it is possible to design oracle with a bit of slack in the dimension requirements, in such a way that all the potential functions Φ_r only pay a constant amortized increase over the duration of the walk.

Lemma 5.8. *There is a construction of* oracle$(A, x(t))$ *that always returns a subspace of codimension at most $\frac{k}{4} + O(1)$ (with k being the number of active coordinates of $x(t)$), such that for any $r = 0, \ldots, \lceil \log_2 n \rceil$, $\Phi_{r,T}(x(T)) \leq O(1)$.*

Small row regime. Set $\eta := \sqrt{\frac{\log n}{\lambda}}$ (it will be the parameter of the regularizer in this regime). We define $B \in \mathbb{R}^{n \times n}$ to be the following thresholded version of A: for any $i, j \in [n]$, let $B_{ij} := A_{ij}$ if $A_{ij}^2 \leq \frac{1}{16K^2\eta^2}$ and $B_{ij} := 0$ otherwise (for some constant K that we will fix later). We will later see in Section 5.4 that it is sufficient to monitor the discrepancy of B instead of A.

Recall that at this point of the walk, each row will have effective ℓ_2-mass $O(\lambda)$. For algorithms based on orthogonality constraints, there is not much difference between having bounded rows and bounded columns, so this justifies patching an algorithm for Banaszczyk's setting at this point. We implement the potential function from [LRR17] in our regularization framework.

Let $\pi'_t : \mathbb{R}^n \to \mathbb{R}^n$ to be such that for any $x \in \mathbb{R}^n$ and $i \in [n]$,

$$(\pi'_t(x))_i = \begin{cases} \langle B_i, x - x(t_i) \rangle - K\eta \sum_{j=1}^n B_{ij}^2 (x_j - x_j(t_i))^2 & \text{if } i \notin P(t) \\ 0 & \text{if } i \in P(t) \end{cases}$$

for some constant $K > 0$ that we will fix in the proof. Finally, we define our potential function for this regime:

$$\Psi_t(x) := \mathsf{smax}_\eta(\pi'_t(x)) .$$

Our main lemma states we can make this potential non-increasing (and moreover the subspace dimension necessary for that allows some slack).

Lemma 5.9. *There is a construction of* oracle$(A, x(t))$ *that always returns a subspace of codimension at most $\frac{k}{4} + O(1)$ (where k is the number of active coordinates of $x(t)$), such that $\Psi_T(x(T)) \leqslant \Psi_0(x(0))$.*

Remark 5.10. Some intuition for π'_t comes from the fact that to get a coloring of discrepancy $\sqrt{\lambda \log n}$ for an input matrix with *rows* of ℓ_2-norm bounded by λ, the Chernoff and union bound argument suffices. If $v_1, \ldots, v_n$ are the rows of the matrix, it is essentially consisting in arguing that when $x \sim \{\pm 1\}^n$, it holds for any $\eta \geqslant 0$ that:

$$\log \mathbf{E} \exp\left(\eta \max_{i \in [n]} |\langle v_i, x \rangle|\right) \leqslant \log \mathbf{E} \sum_{i=1}^{n} \exp(\eta |\langle v_i, x \rangle|) \leqslant \log \sum_{i=1}^{n} \exp\left(\frac{\eta^2 \|v_i\|_2^2}{2}\right).$$

If we interpret $\|v_i\|_2^2$ as $\sum_{j \in [n]} v_{i,j}^2 x_j^2$, this might motivate us to look at the softmax of $\{\langle v_i, x \rangle - \eta \langle v_i^{\odot 2}, x^{\odot 2} \rangle / 2\}$.

The plan for the rest of this section is as follows. First, we prove Lemma 5.8 in Section 5.2 and Lemma 5.9 in Section 5.3. Then in Section 5.4 we show how to deduce Theorem 5.2 and Theorem 5.3. Finally, we study the consequences of Theorem 5.2 and Theorem 5.3 for random instances in Section 5.5.

5.2 Discrepancy in the random regime

Our main goal in this section is to prove Lemma 5.8. Throughout this discussion, we fix a small constant $\varepsilon \in (0, 1/5)$. Our first lemma describes a construction of oracle that bounds locally the increase of the potential. This part is very similar in spirit to our proof of Spencer's theorem.

Lemma 5.11. *Let $x := x(t)$ and $k = k(t) := F(t)$. Let $R_0 = \lceil \log_2(32n/k) \rceil$. There exists a subspace $S = S(t) \subseteq F(t)$ such that S has codimension at most $\frac{k}{4}$ and if $\delta \in S$ satisfies $\|\delta\|_\infty \leqslant 1/poly(n)$,*

$$\Phi_{r,t}(x + \delta) - \Phi_{r,t}(x) \lesssim u_r(\delta) + \frac{1}{k}\left(\frac{k2^r}{n}\right)^{\frac{1-3\varepsilon}{2}} \|\delta\|_2^2 \quad \text{for any } r \leqslant R_0$$

and

$$\Phi_{r,t}(x + \delta) = \Phi_{r,t}(x) \quad \text{for any } r > R_0$$

where $\{u_r : r \leqslant R_0\}$ are linear forms.

Proof. For any $r = 0, \ldots, R_0$, let $k_r = k_r(t) := C_1 \left(\frac{k2^r}{n}\right)^\varepsilon k$ be the effective subspace dimension devoted to rows in the r-th group, where $C_1 = C_1(\varepsilon)$ is chosen so that $\sum_{r \leqslant R_0} k_r \leqslant k/8$.

Let S' be the orthogonal complement of the span of the large rows and the row in I, namely $(\bigcup_{r > R_0} \{A_i : i \in R_r\})^\perp$. These rows all have total ℓ_2-squared mass larger than $2^{R_0 - 1} \geqslant 16n/k$, so there are at most $k/16$ of them and thereby S_1 has codimension at most $k/16$.

Applying Lemma 3.9 to $r \leqslant R_0$, for some $\nabla_r \in \Delta_n$, it holds for any δ with $\|\delta\|_\infty \leqslant 1/poly(n)$ that

$$\Phi_{r,t}(x + \delta) - \Phi_{r,t}(x) \leqslant u_r(\delta) + 2\sqrt{n 2^{1-r}} \sum_{i \in R_r \cap P(t)} \nabla_{r,i}^{\frac{3}{2}} \langle A_i, \delta \rangle^2 \quad \text{for some linear form } u_r : \mathbb{R}^n \to \mathbb{R}.$$

Let I_r be the set of coordinates that are in the top $k_r/2$ entries of the gradient ∇_r. We define S_r to be the intersection of the orthogonal complement of the span of the rows in $I \cup I_r$ (where I is the

21

set of rows from Claim 5.7 that satisfies $|I| \leqslant k/16$), and of the top $k_r/2$-dimensional eigenspace of $\sum_{i \in R_r \cap P(t), i \notin I \cup I_r} \nabla_{r,i}^{\frac{3}{2}} A_i A_i^T$ over $\mathbb{R}^{F(t)}$. Then if $\delta \in S_r$,

$$\Phi_{r,t}(x + \delta) - \Phi_{r,t}(x) - u_r(\delta) \lesssim \frac{\|\delta\|_2^2 \sqrt{n}}{2^{\frac{r}{2}} k_r} \sum_{i \in R_r \cap P(t), i \notin I \cup I_r} \nabla_{r,i}^{\frac{3}{2}} \sum_{j \in F(t)} A_{ij}^2 \qquad \text{(since } \delta \in S_r\text{)}$$

$$\lesssim \frac{\|\delta\|_2^2 k}{k_r \sqrt{n} 2^{\frac{r}{2}}} \sum_{i \in R_r \cap P(t), i \notin I \cup I_r} \nabla_{r,i}^{\frac{3}{2}} \sum_{j=1}^{n} A_{ij}^2 \qquad \text{(by Claim 5.7 and } i \in P(t), i \notin I\text{)}$$

$$\leqslant \frac{\|\delta\|_2^2 k 2^{\frac{r}{2}}}{k_r \sqrt{n}} \sum_{i \in R_r \cap P(t), i \notin I \cup I_r} \nabla_{r,i}^{\frac{3}{2}} \qquad \text{(since } i \in R_r\text{)}$$

$$\lesssim \frac{\|\delta\|_2^2 k 2^{\frac{r}{2}}}{k_r^{\frac{3}{2}} \sqrt{n}} \qquad \text{(since } i \notin I_r\text{)}$$

$$\lesssim \frac{1}{k} \left(\frac{k 2^r}{n} \right)^{\frac{1-3\varepsilon}{2}} \|\delta\|_2^2. \qquad \text{(by definition of } k_r\text{)}$$

Finally we set $S := S' \cap \bigcap_{r \leqslant R_0} S_r$. One can check that S has codimension at most $\frac{k}{16} + \frac{k}{16} + \sum_{r \leqslant R_0} k_r \leqslant \frac{k}{4}$. $\qquad \square$

The following step is a trick to handle the first-order terms. Indeed, a caveat is that unlike in Spencer's setting, we cannot afford to move perpendicularly to all the gradients simultaneously.

Lemma 5.12. *Fix $x \in \mathbb{R}^n$. Let S be a subspace such that for any $\delta \in S$ and $r \leqslant R_0$,*

$$\Phi_{r,t}(x + \delta) - \Phi_{r,t}(x) \lesssim u_r(\delta) + \frac{1}{k} \left(\frac{k 2^r}{n} \right)^{\frac{1-3\varepsilon}{2}} \|\delta\|_2^2 \quad \text{for some linear forms } \{u_r : r \leqslant R_0\}.$$

Then for any $\delta \in S$, at least one of $+\delta$ or $-\delta$ satisfies that for any $r \leqslant R_0$,

$$\Phi_{r,t}(x \pm \delta) - \Phi_{r,t}(x) \lesssim \frac{1}{k} \left(\frac{k 2^r}{n} \right)^{\varepsilon} \|\delta\|_2^2.$$

Proof. By picking $\varepsilon < 1/5$, we have for any $\delta \in S$

$$\sum_{r \leqslant R_0} 2^{-\varepsilon r} (\Phi_{r,t}(x + \delta) - \Phi_{r,t}(x) - u_r(\delta)) \leqslant \frac{1}{k} \left(\frac{k}{n} \right)^{\varepsilon} \|\delta\|_2^2.$$

By a trivial upper bound, this means that for any $\delta \in S, r \leqslant R_0$,

$$2^{-\varepsilon r} (\Phi_{r,t}(x + \delta) - \Phi_{r,t}(x)) \leqslant \sum_{s \leqslant R_0} 2^{-\varepsilon s} u_s(\delta) + \frac{1}{k} \left(\frac{k}{n} \right)^{\varepsilon} \|\delta\|_2^2.$$

In particular, by picking the signing $\pm \delta$ that satisfies $\sum_{s \leqslant R_0} 2^{-\varepsilon s} u_s(\pm \delta) \leqslant 0$, we get that for any $r \leqslant R_0$,

$$\Phi_{r,t}(x \pm \delta) - \Phi_{r,t}(x) \lesssim \frac{1}{k} \left(\frac{k 2^r}{n} \right)^{\varepsilon} \|\delta\|_2^2. \qquad \square$$

Proof of Lemma 5.8. By combining Lemma 5.11 and Lemma 5.12, we know that at any step where there are k active coordinates remaining, the potential $\Phi_{r,t}$ increases by at most $\frac{1}{k}\left(\frac{k2^r}{n}\right)^\varepsilon \|\delta\|_2^2$ if $r \leqslant 1 + \log_2(32n/k)$ and is unchanged for $r \geqslant 1 + \log_2(32n/k)$.

Now fix any $r \leqslant \lceil\log_2(n)\rceil$. By a similar argument to the one in the proof of Theorem 2.1, after letting β_k be the ℓ_2-squared mass injected into x starting from the first time for which there are at most k active coordinates remaining, we can upper bound

$$\Phi_{r,T}(x(T)) - \Phi_{r,0}(0) \lesssim \left(\frac{2^r}{n}\right)^\varepsilon \sum_{k=1}^{64n2^{-r}} \frac{\beta_k - \beta_{k-1}}{k^{1-\varepsilon}} = O(1)\,.$$

The claimed bound follows after noting that our choice of parameters for the regularizers also implies $\Phi_{r,0}(0) = O(1)$. $\qquad\square$

5.3 Discrepancy in the small row regime

Our next goal is to prove Lemma 5.9, which is a direct consequence of the following lemma.

Lemma 5.13. *Fix any time t and let $k := |F(t)|$. There is a subspace S of $\mathbb{R}^{F(t)}$ of codimension at most $\frac{k}{4} + O(1)$ such that*

$$\Psi_t(x + \delta) \leqslant \Psi_t(x)$$

holds for any $\delta \in S$ with $\|\delta\|_\infty \leqslant 1/poly(n)$.

Before proving it, we recall the following well-known result (see for example [BDG19, Theorem 8] or [LRR17, Lemma 21]):

Lemma 5.14. *For any $w_1, \ldots, w_m \geqslant 0$, $B \in \mathbb{R}^{m \times n}$, and $\alpha \in (0, 1)$, there exists a subspace S of $\mathbb{R}^n$ of codimension at most αn such that for all $\delta \in S$,*

$$\sum_{i \leqslant m} w_i \left(\sum_{j \leqslant n} B_{ij}\delta_j\right)^2 \leqslant \frac{1}{\alpha} \sum_{i \leqslant m} w_i \sum_{j \leqslant n} B_{ij}^2 \delta_j^2\,. \tag{12}$$

Proof. Up to considering $\sqrt{w_i}B_i$, assume without of generality that $w_i = 1$ for all $i \in [m]$. Moreover, the statement is invariant if we remove the zero columns from B and replace δ_j by $\frac{\delta_j}{\|B^j\|_2}$ for all $j \in [n]$. Therefore we can also assume that all columns of B have unit Euclidean length.

Now the right-hand side is just $\frac{\|\delta\|_2^2}{\alpha}$ and the left-hand side is $\delta^T \sum_{i \leqslant m} B_i B_i^T \delta$. We can simply choose S to be the subspace of vectors δ orthogonal to the top αn eigenspace of the linear operator $\sum_{i \leqslant m} B_i B_i^T$, which has trace n. The result follows a counting argument. $\qquad\square$

Proof of Lemma 5.13. To avoid overcharging notations we drop in this proof the dependencies on t and let $x := x(t)$, $F := F(t)$ and $P := P(t)$. When $\|\delta\|_\infty \leqslant 1/poly(n)$, we can apply Taylor expansion (Lemma 3.9) – for some $\nabla \in \Delta_n$:

$$\Psi_t(x + \delta) - \Psi_t(x) \leqslant \sum_{i \notin P} \nabla_i \sum_{j \in F} B_{ij}\delta_j + \eta \sum_{i \notin P} \nabla_i \left(\sum_{j \in F} B_{ij}\delta_j\right)^2$$

$$- 2K\eta \sum_{i \notin P} \nabla_i \sum_{j \in F} B_{ij}^2 x_j \delta_j + 2K^2 \eta^3 \sum_{i \notin P^t} \nabla_i \left(\sum_{j \in F} B_{ij}^2 x_j \delta_j \right)^2$$

$$- K\eta \sum_{i \notin P} \nabla_i \sum_{j \in F} B_{ij}^2 \delta_j^2 + K^2 \eta^3 \sum_{i \notin P} \nabla_i \left(\sum_{j \in F} B_{ij}^2 \delta_j^2 \right)^2.$$

First, note that the last term scales as δ_j^4 so we can make it negligible by picking $\|\delta\|_\infty \leqslant 1/\text{poly}(n)$.

We consider the subspace S_1 of codimension at most 2 that is the orthogonal complement of the span of the 2 vectors from the linear terms in δ in the previous right-hand side. Also, by applying Lemma 5.14 to two different matrices with $\alpha = \frac{1}{8}$, we can find S_2 of codimension $k/4$ such that any $\delta \in S_3$ satisfies the following two conditions:

$$\eta \sum_{i \notin P} \nabla_i \left(\sum_{j \in F} B_{ij} \delta_j \right)^2 \leqslant 8\eta \sum_{i \notin P} \nabla_i \sum_{j \in F} B_{ij}^2 \delta_j^2. \tag{13}$$

$$2K^2 \eta^3 \sum_{i \notin P} \nabla_i \left(\sum_{j \in F} B_{ij}^2 x_j \delta_j \right)^2 \leqslant 16K^2 \eta^3 \sum_{i \notin P} \nabla_i \sum_{j \in F} B_{ij}^4 x_j^2 \delta_j^2. \tag{14}$$

Note that whenever (14) is satisfied, it also follows from $|x_i| \leqslant 1$ and the assumption $B_{ij}^2 \leqslant \frac{1}{16K^2\eta^2}$ that

$$2K^2 \eta^3 \sum_{i \notin P} \nabla_i \left(\sum_{j \in F} B_{ij}^2 x_j \delta_j \right)^2 \leqslant \eta \sum_{i \notin P} \nabla_i \sum_{j \in F} B_{ij}^2 \delta_j^2.$$

Let $S := S_1 \cap S_2$. S has codimension at most $k/4 + O(1)$ by construction. Picking $K := 9$, we get

$$\Psi_t(x + \delta) - \Psi_t(x) \leqslant 0,$$

for any $\delta \in S$ satisfying $\|\delta\|_\infty \leqslant 1/\text{poly}(n)$. $\qquad\square$

Finally we bound the error of replacing A by B.

Lemma 5.15. *For any $i \in [n]$,*

$$|\langle A_i - B_i, x(T) - x(t_i) \rangle| \lesssim \sqrt{\lambda \log n}.$$

Proof. Fix $i \in [n]$. Let $F := F(t_i)$ be the set of active coordinates when the i-th row becomes small. Since $\sum_{j \in F} A_{ij}^2 = O(\lambda)$, it must be that

$$|\{j \in F : A_{ij} - B_{ij} \neq 0\}| \lesssim \log n.$$

Therefore, by Cauchy-Schwarz,

$$|\langle A_i - B_i, x(T) - x(t_i) \rangle| \leqslant \sum_{j \in F} |A_{ij} - B_{ij}| \lesssim \sqrt{\log n} \sqrt{\sum_{j \in F} (A_{ij} - B_{ij})^2} \lesssim \sqrt{\lambda \log n}. \qquad\square$$

5.4 Putting everything together

We are now ready to prove Theorem 5.2 and Theorem 5.3.

Proof of Theorem 5.2 and Theorem 5.3. We assume without loss of generality that $\max_i (Ax)_i = \|Ax\|_\infty$ by the usual trick of doubling the rows. We define $\mathsf{oracle}(A, x(t))$ to be the intersection of the halfspace from Lemma 5.11 and Lemma 5.12, and of the subspace from Lemma 5.13.

On the one hand, Lemma 5.8 implies that for any $r \leqslant \lceil \log_2 n \rceil$ and $i \in R_r$,

$$\langle A_i, x(t_i) \rangle = \pi_{r,T}(x(T))_i \leqslant \Phi_{r,T}(x(T)) \leqslant O(1) \,.$$

On the other hand, Lemma 5.9 implies that for any $i \in [n]$,

$$\langle B_i, x(T) - x(t_i) \rangle - K\eta \sum_{j=1}^n B_{ij}^2 (x_j(T) - x_j(t_i))^2 \leqslant \Psi_T(x(T)) \leqslant \Psi_0(x(0)) \leqslant \sqrt{\lambda \log n} \,.$$

Observe that

$$\sum_{j=1}^n B_{ij}^2 (x_j(T) - x_j(t_i))^2 \leqslant 4 \sum_{j=1}^n B_{ij}^2 \lesssim \lambda \,.$$

Hence, $\langle B_i, x(T) - x(t_i) \rangle \lesssim \sqrt{\lambda \log n}$, and by Lemma 5.15, $|\langle A_i - B_i, x(T) - x(t_i) \rangle| \lesssim \sqrt{\lambda \log n}$ holds as well. It remains to use the triangle inequality:

$$\|Ax(T)\|_\infty = O(1 + \sqrt{\lambda \log n}) \,.$$

This implies Theorem 5.2 and the first part of Theorem 5.3. For the second part of Theorem 5.3, simply observe that if A is a rescaled Beck-Fiala instance, namely $A_{ij} \in \{0, 1/\sqrt{s}\}$ for all $i, j \in [n]$, then the condition $\sum_{j \in F(t_i)} A_{ij}^2 = O(\lambda)$ implies by that A_i has at most $O(s\lambda)$ nonzero entries in $F(t_i)$, and so $\sum_{j \in F(t_i)} |A_{ij}| = O(\lambda\sqrt{s})$. In particular, the time steps $t \in [t_i, T]$ affect the discrepancy of A_i by at most $O(\lambda\sqrt{s})$. This shows that the constructed coloring also has discrepancy $O(1 + \lambda\sqrt{s})$ in this case, which is equivalent to the second part of the bound in Theorem 5.3. $\qquad\square$

Remark 5.16. One may also recover Banaszszyk's bound for Komlós (or Beck-Fiala) instances by repeating the argument from Section 5.3, replacing λ by some universal constant larger than 1 and adding an additional orthogonality constraint to large rows, so that the algorithm can pretend that the rows all have bounded ℓ_2-mass. It follows from our previous observations on the link between the multiplicative weights update method and negative entropy regularization that this would be equivalent to the approach for recovering Banaszczyk's bound in [LRR17].

5.5 Application to random instances

5.5.1 Random orthogonal matrices

The next consequence of our bound for pseudorandom instances is that Komlós conjecture is true for *random* rotation matrices. An equivalent geometric way to state our result is the following: there exists a universal constant $C > 0$ such that when we randomly rotate the n-dimensional hypercube around the origin, with high probability there exists a corner at ℓ_∞-distance at most C from the origin.

Rotation matrices appear to be hard instances for proving the Komlós conjecture, as present proof techniques merely manage to match the discrepancy bounds to those for the suprema of Rademacher processes involving the transpose matrix. Improving beyond $O(\sqrt{\log n})$ discrepancy for orthogonal matrices would therefore provide new techniques for treating Rademacher/Gaussian processes for structured matrices. A first step in making progress on this front would therefore be to consider *random* orthogonal matrices.

What we mean by random rotation is a random matrix distributed according to the Haar measure on the orthogonal group $\mathcal{O}(n)$. The Haar measure is a natural generalization of the uniform distribution. We can just think of the sampling as picking the matrix columns to be i.i.d. standard Gaussians in $\mathbb{R}^n$ (which will be linearly independent almost surely), and orthonormalizing them with the Gram-Schmidt process.

We explicitly computed small moments of the entries of such a random matrix with the help of the Maple package `IntHaar` [GK21].

Claim 5.17. Suppose A is distributed according to the Haar measure on $\mathcal{O}(n)$. Then,

$$\mathbf{E}\left[A_{11}^8\right] = \frac{105}{n(n+2)(n+4)(n+6)},$$

$$\mathbf{E}\left[A_{11}^4 A_{12}^4\right] = \frac{9}{n(n+2)(n+4)(n+6)},$$

$$\mathbf{E}\left[A_{11}^2 A_{12}^2 A_{21}^2 A_{22}^2\right] = \frac{n^2+4n+15}{n(n+2)(n-1)(n+1)(n+4)(n+6)}.$$

Corollary 5.18 (Komlós conjecture for random rotations)**.** *There is a deterministic algorithm that given a Haar-distributed random matrix A on $\mathcal{O}(n)$, finds with high probability $x \in \{\pm 1\}^n$ such that*

$$\|Ax\|_\infty = O(1).$$

Proof. Our proof is inspired by the observations in the proof of Theorem 1 of [Ber01] for the Haar measure on the unitary group. Consider $B := (A^{\odot 2})^T A^{\odot 2}$. Using Claim 5.17, we see that

$$\begin{aligned}
\mathbf{E}\,\mathrm{Tr}\,B^2 &= \sum_{1 \leqslant i,j,k,l \leqslant n} \mathbf{E}\left[A_{ij}^2 A_{il}^2 A_{kj}^2 A_{kl}^2\right] \\
&= n^2 \mathbf{E}\left[A_{11}^8\right] + 2n^2(n-1)\mathbf{E}\left[A_{11}^4 A_{12}^4\right] + n^2(n-1)^2 \mathbf{E}\left[A_{11}^2 A_{12}^2 A_{21}^2 A_{22}^2\right] \\
&= 1 + O\left(\frac{1}{n}\right).
\end{aligned}$$

Note that 1 and $\lambda(A)^2$ are eigenvalues of the positive semidefinite matrix B, so $1 + \lambda(A)^4 \leqslant \mathrm{Tr}\,B^2$, and $\mathbf{E}\,\lambda(A)^4 \leqslant O\left(\frac{1}{n}\right)$ by the previous estimate. By Markov's inequality, $\lambda(A) \leqslant \frac{1}{\log n}$ with high probability, and we conclude by applying Theorem 5.2. $\qquad\square$

5.5.2 Random Gaussian matrices

Next we show that for matrices with random Gaussian entries, the corresponding λ parameter is small. Without loss of generality, we assume that the input matrix is square, as otherwise (in the regime $m \geqslant n$) we can add extra columns while only worsening λ. We assume that each entry

is sampled i.i.d. from $\mathcal{N}(0, \sigma^2)$ with $\sigma = \frac{1}{\sqrt{n}}$, so that all column norms are tightly concentrated around 1, i.e.

$$\mathbf{Pr}(1 - \varepsilon \leqslant \|A^j\|_2^2 \leqslant 1 + \varepsilon) \geqslant 1 - 2\exp\left(-\frac{n\varepsilon^2}{8}\right),$$

which follows from standard concentration bounds.

Claim 5.19. Given a random Gaussian matrix $A \in \mathbb{R}^{n \times n}$, where entries are i.i.d. Gaussians $\mathcal{N}(0, \sigma^2)$ with $\sigma = \frac{1}{\sqrt{n}}$, one has that

$$\max_{\langle u, \mathbf{1}\rangle = 0} \frac{\|A^{\odot 2}u\|_2}{\|u\|_2} \leqslant \frac{1}{\sqrt{n}}$$

with high probability.

Proof. Let $B := A^{\odot 2} - \frac{\mathbf{1}\mathbf{1}^T}{n}$. Observe that

$$\max_{\|u\|_2 = 1, \langle u, \mathbf{1}\rangle = 0} \|A^{\odot 2}u\|_2 \leqslant \max_{\|u\|_2 = 1} \|Bu\|_2.$$

Now, B is a matrix with i.i.d. entries such that $\mathbf{E}\, B_{11} = 0$, $\mathbf{E}\, B_{11}^2 = O(1/n^2)$ and $\mathbf{E}\, B_{11}^4 = O(1/n^4)$, so by a standard result from random matrix theory (see e.g. Theorem 2.3.8 of [Tao12]), it holds that $n\|B\|_{\text{op}} = O(\sqrt{n})$ with high probability. It readily follows that $\lambda(A) \leqslant 1/\sqrt{n}$ with high probability. $\square$

This shows that by applying Theorem 5.2 random Gaussian matrices have discrepancy $O(1)$.

Corollary 5.20. *Given a random Gaussian matrix $A \in \mathbb{R}^{n \times n}$, where entries are i.i.d. Gaussians $\mathcal{N}(0, \sigma^2)$, $\sigma = \frac{1}{\sqrt{n}}$, there exists a deterministic algorithm that finds a coloring $x \in \{\pm 1\}^n$ such that $\|Ax\|_\infty = O(1)$.*

6 Discussion on symmetric Beck-Fiala instances

An interesting special case of the Beck-Fiala conjecture is when the matrix A is the adjacency matrix of some s-regular graph. It turns out that in this setting, a folklore argument based on Lovász Local Lemma implies that there exists a coloring with discrepancy $O(\sqrt{s \log s})$. Although there is an algorithm to construct such a coloring in polynomial time, it is not captured by the iterative framework we introduced in this paper. It is in our opinion a great open problem to unify those two lines of work.

To formulate the problem more precisely, we provide a new streamlined and self-contained analysis of the algorithm matching the bound based on Lovász Local Lemma. In particular, it highlights the differences with the sticky walk approach. Our inspiration is an argument of [AIS19] for finding a satisfying assignment of bounded degree k-SAT instances.

Theorem 6.1 (Folklore). *There is a randomized algorithm that, given $A \in \{0, 1\}^{n \times n}$ with at most s nonzero entries per row and at most s nonzero entries per column, finds with high probability in polynomial time a coloring $x \in \{\pm 1\}^n$ such that $\|Ax\|_\infty = O(\sqrt{s \log s})$.*

Proof. We will call a row *bad* (w.r.t. an implicit full coloring) when its discrepancy is larger than $4\sqrt{s \log s}$. We consider the following algorithm. First, we generate a uniformly random coloring. Then we repeat t times the operation of picking the bad row with smallest index (unless there is none, in which case we stop) and resampling all the variables appearing in it.

Since each constraint contains at most s variables and each variable appears in at most s constraints, any constraint has nonempty intersection with at most s^2 other constraints.

Define C_t to be the set of all ordered sequences of t constraints that have nonzero probability to be picked in that order by the algorithm. The execution of the algorithm can be described as a rooted forest of t vertices, each one of these corresponding to a constraint that is picked. When a constraint is picked, it can create at most s^2 children, each of which corresponding to a constraint of lower index that intersects it and became bad after the resampling.

Therefore, we can encode an element $c \in C_t$ by giving $\{c_i : \forall j \in \{1, \ldots, i-1\}, c_j < c_i\}$, and a rooted forest on t vertices, each (except the roots) with labels between 1 and s^2. It follows from standard combinatorics that

$$|C_t| \leqslant 2^n \binom{2t}{t} s^{2t} = 2^n (2s)^{2t} .$$

Fix a sequence of resampled constraints $c \in C_t$ and a sequence of $t+1$ colorings $u_1, \ldots, u_{t+1} \in \{\pm 1\}^n$. c and u can correspond to a potential execution of the algorithm only if u_i is u_{i+1} where the c_i-th constraint of u_i is bad. Applying Chernoff bounds, we see that there can be at most $2^s/s^8$ such u_i's. It follows by induction that there are at most $2^{st}/s^{8t}$ possible sequences $u_1, \ldots, u_t$. On the other hand, for any fixed $u_1, \ldots, u_{t+1}, c_1, \ldots, c_t$, the probability that the algorithm follows exactly this sequence of constraints and colorings is at most 2^{-st}. Hence, by a union bound, the probability that the final coloring that the sequence of constraints is $c_1, \ldots, c_t$ is at most $2^n s^{-8t}$.

To conclude, we can apply another union bound to get that the probability that the final discrepancy is larger than $\sqrt{2s \log s}$ is at most $|C_t| 2^n s^{-8t}$, which is $n^{-\Omega(1)}$ for some $t = n^{O(1)}$. $\qquad\square$

Instead of working with fractional colorings, here we walk directly in the space of full colorings. While the row-sparsity assumption is not really restrictive in the sticky walk framework (as we can always pick update vectors orthogonal to large rows), it seems crucial for arguments based on Lovász Local Lemma.

A concrete family of instances. We now introduce a family of discrepancy instances for which the tools we use to analyze our iterative framework fail to provide interesting bounds. It is an interesting question whether a more refined analysis will yield an improved discrepancy bound.

We believe these examples are essential in benchmarking attempts at improving discrepancy bounds, so we will consider them to be candidate hard instances.

Definition 6.2 (Twisted Hypercubes). The graph on one vertex is the only twisted hypercube of dimension 0. A *twisted hypercube* of dimension d is then obtained by taking two copies of the same twisted hypercube of dimension $d-1$, and adding a matching between both vertex sets.[2]

Twisted hypercubes of dimension d have $n = 2^d$ vertices, each of degree $O(\log n)$. Theorem 6.1 implies that they have colorings of discrepancy $O(\sqrt{\log n \log \log n})$, but Theorem 5.3 only gives a

---

[2] A slightly different construction consists in adding a matching between two *potentially distinct* twisted hypercubes of dimension $d-1$. This is for example the convention chosen in one previous use of the term "twisted hypercube" in the litterature [DPP+18]. We believe it does not make much difference in the discrepancy setting.

bound of $O(\log n)$ (note that this bound would also follow from [BF81]). We believe it would be interesting to find a construction of colorings of twisted hypercubes of discrepancy $o(\log n)$ using the sticky walk approach.[3]

Remark 6.3. Although smoothing the twisted hypercube is not sufficient to apply our bounds on pseudorandom Beck-Fiala instances, we can at least transform it into the adjacency matrix of a (multi-)graph with constant spectral expansion, while only losing an $O(1)$-additive factor on the discrepancy of any coloring. Therefore, our question on the discrepancy of symmetric instances could be reduced to that of the discrepancy of symmetric *expanding* instances.

Acknowledgments

References

[AIS19]   Dimitris Achlioptas, Fotis Iliopoulos, and Alistair Sinclair. Beyond the Lovász Local Lemma: Point to Set Correlations and Their Algorithmic Applications. In *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019*, pages 725–744. IEEE Computer Society, 2019. 27

[AN22]   Dylan J. Altschuler and Jonathan Niles-Weed. The Discrepancy of Random Rectangular Matrices. *Random Struct. Algorithms*, 60(4):551–593, 2022. 3

[AZLO15] Zeyuan Allen-Zhu, Zhenyu Liao, and Lorenzo Orecchia. Spectral Sparsification and Regret Minimization Beyond Matrix Multiplicative Updates. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015*, pages 237–245. ACM, 2015. 2, 3, 7, 10

[Ban90]   Wojciech Banaszczyk. A Beck-Fiala-type Theorem for Euclidean Norms. *Eur. J. Comb.*, 11(6):497–500, 1990. 33

[Ban98]   Wojciech Banaszczyk. Balancing Vectors and Gaussian Measures of n-dimensional Convex Bodies. *Random Structures & Algorithms*, 12(4):351–360, 1998. 1, 17

[Ban10]   Nikhil Bansal. Constructive Algorithms for Discrepancy Minimization. In *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010*, pages 3–10. IEEE Computer Society, 2010. 1, 2, 3

---

[3]One could also consider randomly twisted hypercubes, obtained by taking recursively uniformly random matchings. They form a family of structured random instances that are not captured by our pseudorandom bounds.

[BC12]   Sébastien Bubeck and Nicolò Cesa-Bianchi. Regret Analysis of Stochastic and Non-stochastic Multi-armed Bandit Problems. *Found. Trends Mach. Learn.*, 5(1):1–122, 2012. 3

[BDG19]   Nikhil Bansal, Daniel Dadush, and Shashwat Garg. An Algorithm for Komlós Conjecture Matching Banaszczyk's Bound. *SIAM J. Comput.*, 48(2):534–553, 2019. 1, 2, 3, 17, 23

[BDGL19]   Nikhil Bansal, Daniel Dadush, Shashwat Garg, and Shachar Lovett. The Gram-Schmidt Walk: A Cure for the Banaszczyk Blues. *Theory Comput.*, 15:1–27, 2019. 2, 3, 33

[Bel13]   Adrian William Belshaw. *Strong Normality, Modular Normality, and Flat Polynomials: Applications of Probability in Number Theory and Analysis*. PhD thesis, Simon Fraser University, Department of Mathematics, 2013. 15

[Ber99]   Dimitri P. Bertsekas. *Nonlinear Programming*. Athena Scientific, 1999. 11

[Ber01]   Gregory Berkolaiko. Spectral Gap of Doubly Stochastic Matrices Generated from Equidistributed Unitary Matrices. *Journal of Physics A: Mathematical and General*, 34(22):L319, 2001. 26

[BF81]   József Beck and Tibor Fiala. "Integer-making" Theorems. *Discrete Applied Mathematics*, 3(1):1–8, 1981. 29

[BLV22]   Nikhil Bansal, Aditi Laddha, and Santosh S. Vempala. A Unified Approach to Discrepancy Minimization. In Amit Chakrabarti and Chaitanya Swamy, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2022, September 19-21, 2022, University of Illinois, Urbana-Champaign, USA (Virtual Conference)*, volume 245 of *LIPIcs*, pages 1:1–1:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022. 2, 3, 4, 6, 7

[BM20]   Nikhil Bansal and Raghu Meka. On the Discrepancy of Random Low Degree Set Systems. *Random Struct. Algorithms*, 57(3):695–705, 2020. 2, 3

[BSS14]   Joshua D. Batson, Daniel A. Spielman, and Nikhil Srivastava. Twice-Ramanujan Sparsifiers. *SIAM Rev.*, 56(2):315–334, 2014. 3, 7, 10

[Cha00]   Bernard Chazelle. *The Discrepancy Method: Randomness and Complexity*. Cambridge University Press, 2000. 1

[DJR22]   Daniel Dadush, Haotian Jiang, and Victor Reis. A New Framework for Matrix Discrepancy: Partial Coloring Bounds via Mirror Descent. In Stefano Leonardi and Anupam Gupta, editors, *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*, pages 649–658. ACM, 2022. 3

[DPP+18]   Andrzej Dudek, Xavier Pérez-Giménez, Pawel Pralat, Hao Qi, Douglas B. West, and Xuding Zhu. Randomly Twisted Hypercubes. *Eur. J. Comb.*, 70:364–373, 2018. 28

[EL19]   Esther Ezra and Shachar Lovett. On the Beck-Fiala Conjecture for Random Set Systems. *Random Struct. Algorithms*, 54(4):665–675, 2019. 3

[ES18]   Ronen Eldan and Mohit Singh. Efficient Algorithms for Discrepancy Minimization in Convex Sets. *Random Struct. Algorithms*, 53(2):289–307, 2018. 2, 3, 4

[Gia97]   Apostolos Giannopoulos. On some Vector Balancing Problems. *Studia Mathematica*, 122(3):225–234, 1997. 4

[GK21]   Alejandro Ginory and Jongwon Kim. Weingarten Calculus and the IntHaar Package for Integrals over Compact Matrix Groups. *J. Symb. Comput.*, 103:178–200, 2021. 26

[Glu89]   Efim D. Gluskin. Extremal Properties of Orthogonal Parallelepipeds and their Applications to the Geometry of Banach Spaces. *Mathematics of the USSR-Sbornik*, 64(1):85–96, 1989. 4

[HR19]   Rebecca Hoberg and Thomas Rothvoss. A Fourier-Analytic Approach for the Discrepancy of Random Set Systems. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2019*, pages 2547–2556. SIAM, 2019. 3

[HRS22]   Samuel B. Hopkins, Prasad Raghavendra, and Abhishek Shetty. Matrix Discrepancy from Quantum Communication. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 637–648, 2022. 3

[JM20]   Chris Jones and Matt McPartlon. Spherical Discrepancy Minimization and Algorithmic Lower Bounds for Covering the Sphere. In Shuchi Chawla, editor, *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SODA 2020, Salt Lake City, UT, USA, January 5-8, 2020*, pages 874–891. SIAM, 2020. 14

[LM15]   Shachar Lovett and Raghu Meka. Constructive Discrepancy Minimization by Walking on the Edges. *SIAM J. Comput.*, 44(5):1573–1582, 2015. 2, 3, 4, 6, 7, 17

[LRR17]   Avi Levy, Harishchandra Ramadas, and Thomas Rothvoss. Deterministic Discrepancy Minimization via the Multiplicative Weight Update Method. In *Integer Programming and Combinatorial Optimization - 19th International Conference, IPCO 2017*, volume 10328 of *Lecture Notes in Computer Science*, pages 380–391. Springer, 2017. 2, 3, 4, 6, 7, 19, 20, 23, 25

[Mat09]   Jiri Matousek. *Geometric Discrepancy: An Illustrated Guide*. Algorithms and Combinatorics. Springer, 2009. 1

[Pot20]   Aditya Potukuchi. A Spectral Bound on Hypergraph Discrepancy. In *47th International Colloquium on Automata, Languages, and Programming, ICALP 2020*, volume 168 of *LIPIcs*, pages 93:1–93:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. 2, 3, 17, 18, 19

[Rot17]   Thomas Rothvoss. Constructive Discrepancy Minimization for Convex Sets. *SIAM J. Comput.*, 46(1):224–234, 2017. 2, 3, 4

[RR20]   Victor Reis and Thomas Rothvoss. Vector Balancing in Lebesgue Spaces. *arXiv preprint arXiv:2007.05634*, 2020. 3

[Spe85]   Joel Spencer. Six Standard Deviations Suffice. *Transactions of the American Mathematical Society*, 289(2):679–706, 1985. 1, 3, 4, 7

[Tao12]   Terence Tao. *Topics in Random Matrix Theory*. Graduate studies in mathematics. American Mathematical Soc., 2012. 27

A An iterative algorithm for ellipsoid discrepancy

In view of attacking harder discrepancy problems such as the Beck-Fiala and Komlós conjectures, it is worth noting that a major obstacle in achieving the conjectured bounds lies in the fact that once many variables get frozen to ± 1, we have a smaller degree of freedom in choosing our update without having seen a significant decrease in the norms of the matrix rows, when restricted to the unfrozen coordinates.

An alternate strategy would be to use an amortized analysis that bounds how much each coordinate of the diagonal matrix that we employ to upper bound the Hessian of the potential function has contributed so far. The fact that the Hessian can change quite drastically throughout the execution of the algorithm poses some difficulties in realizing this approach. However, we can show that in the case where the Hessian essentially stays constant, we can obtain interesting bounds.

Ellipsoid discrepancy. To this extent, we study a simpler discrepancy problem that we call *ellipsoid discrepancy*, and that was initially introduced by Banaszczyk [Ban90]. Given some positive semidefinite matrix $B \in \mathbb{R}^{n \times n}$, we consider the norm $\|\cdot\|_B$ on $\mathbb{R}^n$ defined by

$$\|z\|_B := \sqrt{\langle z, Bz \rangle}, \quad \text{for all } z \in \mathbb{R}^n.$$

We are interested in the following "Euclidean" version of the Komlós problem: given a matrix $Q \in \mathbb{R}^{n \times n}$ whose columns have ℓ_2-norm at most 1, find $x \in \{\pm 1\}^n$ such that $\|Qx\|_B$ is small. Banaszczyk [Ban90] proved that there always exists a coloring of Q achieving ellipsoid discrepancy $\sqrt{\operatorname{Tr} B}$ (and this is tight for any B, as can be seen by taking the columns of Q to be an orthonormal basis of eigenvectors of B). It follows implicitly from other works, including the Gram-Schmidt walk algorithm [BDGL19], that this bound can be matched algorithmically (at least up to the leading constant).[4] We give a different algorithmic proof, based on our iterative meta-algorithm Algorithm 2, that highlights how measuring the discrepancy in an amortized sense is sometimes necessary.

Theorem A.1. *There is a deterministic algorithm running in polynomial time that given a positive semidefinite matrix $B \in \mathbb{R}^{n \times n}$ and a matrix $Q \in \mathbb{R}^{n \times n}$ with column ℓ_2-norm bounded by 1, returns $x \in \{\pm 1\}^n$ such that*

$$\|Qx\|_B = O\left(\sqrt{\operatorname{Tr} B}\right).$$

Our proof of Theorem A.1 uses the iterative machinery described in Section 3.1. Since the $\|\cdot\|_B$-norm squared is already a smooth degree-2 polynomial function, we will not need any regularization. However, we stress that repeating our analysis for Spencer's theorem from Section 4 would only give an ellipsoid discrepancy bound of $\sqrt{\operatorname{Tr} B \log n}$.

Proof. Up to rotating Q (without affecting the norm of its columns), we assume without loss of generality that B is diagonal, with diagonal elements $D_1 \geqslant \ldots \geqslant D_n \geqslant 0$.

We make use of the meta-iterative algorithm Algorithm 2. Set $L := 1/2$ (since there is no regularization involved, any constant would work here). We show how to construct $\mathsf{oracle}(A, x(t))$

---

[4]We thank the anonymous reviewers for pointing this out.

with $F(t) := \{j \in [n] : x_j(t) \notin \{-1, 1\}\}$ and $k = k(t) := |F(t)|$. For any $\delta \in \mathbb{R}^n$,

$$\|Q(x(t) + \delta)\|_D^2 - \|Qx(t)\|_D^2 = 2 \sum_{i=1}^n D_i \langle Q_i, x(t) \rangle \langle Q_i, \delta \rangle + \sum_{i=1}^n D_i \langle Q_i, \delta \rangle^2 . \tag{15}$$

Let S be the subspace of $\mathbb{R}^{F(t)}$ of all the vectors that are orthogonal to $Q_1, \ldots, Q_{\lfloor k/2 \rfloor - 1}$ (restricted to $\mathbb{R}^{F(t)}$). Now we use that if $\delta \in S$,

$$\sum_{i=1}^n D_i \langle Q_i, \delta \rangle^2 \leqslant D_{\lfloor k/2 \rfloor} \left\langle \sum_{i=1}^n Q_i Q_i^T, \delta \delta^T \right\rangle .$$

Restricted to the coordinates in $F(t)$, the matrix $\sum_{i \in [n]} Q_i Q_i^T$ has trace $\sum_{i \in [n], j \in F(t)} Q_{ij}^2 \leqslant k$ by assumption on the norm of the columns of Q. Therefore, by averaging, we can find a 2-dimensional subspace of S for which the previous quadratic form is $O\left(\|\delta\|_2^2\right)$. Finally, we return the intersection of this subspace with the halfspace that makes the first-order term non-positive in (15).

Now we study the total ellipsoid discrepancy incurred over time steps $t = 0, \ldots T$. Let β_k be the ℓ_2-squared norm injected into $x(t)$ between $t_k := \min\{t \geqslant 0 : |F(t)| \leqslant k\}$ and $T := \min\{t \geqslant 0 : |F(t)| \leqslant 3\}$. In particular, we have $\beta_k \leqslant k$. We sum by parts the second-order increase of (15) over the execution of the algorithm,

$$\begin{aligned}
\sum_{i=2}^{\lfloor n/2 \rfloor} D_i(\beta_{2i+1} - \beta_{2i-1}) &= \sum_{i=3}^{\lfloor n/2 \rfloor} \beta_{2i-1}(D_{i-1} - D_i) + D_{\lfloor n/2 \rfloor} \beta_n \\
&\leqslant \sum_{i=3}^{\lfloor n/2 \rfloor} (2i - 1)(D_{i-1} - D_i) + 2n D_{\lfloor n/2 \rfloor} \\
&\lesssim \operatorname{Tr} D .
\end{aligned}$$

Since the first-order term of the increase is always non-positive and the final step of Algorithm 2 only changes the (squared) ellipsoid discrepancy by at most $O(\operatorname{Tr} D)$, the final coloring x^* satisfies $\|Ax^*\|_D^2 \lesssim \operatorname{Tr} D$. $\qquad \square$