An algorithm for komlós conjecture matching Banaszczyk's bound


DOI: 10.1137/17M1126795

Document status and date: Published: 30/04/2019

Document Version: Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 03. Oct. 2020
AN ALGORITHM FOR KOMLÓS CONJECTURE MATCHING BANASZCZYK’S BOUND∗

NIKHIL BANSAL†, DANIEL DADUSH‡, AND SHASHWAT GARG†

Abstract. We consider the problem of finding a low discrepancy coloring for sparse set systems where each element lies in at most \(t\) sets. We give an efficient algorithm that finds a coloring with discrepancy \(O((t \log n)^{1/2})\), matching the best known nonconstructive bound for the problem due to Banaszczyk. The previous algorithms only achieved an \(O(t^{1/2} \log n)\) bound. The result also extends to the more general Komlós setting and gives an algorithmic \(O(\log^{1/2} n)\) bound.

Key words. discrepancy, semidefinite programming, random walk

AMS subject classifications. 68W20, 11K38, 68Q25

DOI. 10.1137/17M1126795

1. Introduction. Let \((V, S)\) be a finite set system with \(V = \{1, \ldots, n\}\) and \(S = \{S_1, \ldots, S_m\}\) a collection of subsets of \(V\). For a two coloring \(\chi : V \to \{-1, 1\}\), the discrepancy of \(\chi\) for a set \(S\) is defined as \(\chi(S) = |\sum_{i \in S} \chi(i)|\) and measures the imbalance from an even split for \(S\). The discrepancy of the system \((V, S)\) is defined as

\[
\text{disc}(S) = \min_{\chi : V \to \{-1, 1\}} \max_{S \in S} \chi(S).
\]

That is, it is the minimum imbalance for all sets in \(S\), over all possible two colorings \(\chi\).

Discrepancy is a widely studied topic and has applications to many areas in mathematics and computer science. For more background we refer the reader to the books [Cha00, Mat09, CST+14]. In particular, discrepancy is closely related to the problem of rounding fractional solutions of a linear system of equations to integral ones [LSV86, Rot12], and is widely studied in approximation algorithms and optimization.

Until recently, most of the results in discrepancy were based on nonalgorithmic approaches and hence were not directly useful for algorithmic applications. However, in the last few years there has been remarkable progress in our understanding of the algorithmic aspects of discrepancy [Ban10, CNN11, LM12, Rot14, HSS14, ES14, NT15]. In particular, we can now match or even improve upon all known applications of the widely used partial-coloring method [Spe85, Mat09] in discrepancy. This has, for example, led to several other new results in approximation algorithms [Rot13, BCKL14, BN15, NTZ13].

Sparse set systems. Despite the algorithmic progress, one prominent question that has remained open is to match the known nonconstructive bounds on discrepancy for low degree or sparse set systems. These systems are parametrized by \(t\), that
denotes the maximum number of sets that contain any element. Beck and Fiala [BF81] proved, using an algorithmic iterated rounding approach, that any such set system has discrepancy at most 2t − 1. They also conjectured that the discrepancy in this case is $O(t^{1/2})$, and settling this has been an elusive open problem.

The best known result in this direction is due to Banaszczyk [Ban98], which implies an $O(\sqrt{\log n})$ discrepancy bound for the problem.¹ Unlike most results in discrepancy that are based on the partial-coloring method, Banaszczyk’s proof is based on a very different and elegant convex geometric argument, and it is not at all clear how to make it algorithmic. Prior to Banaszczyk’s result, the best known nonalgorithmic bound was $O(t^{1/2} \log n)$ [Sri97], based on the partial-coloring method. This bound was first made algorithmic in [Ban10], and by now there are several different ways known to obtain this result [LM12, Nik13, Rot14, HSS14, ES14]. However, the question of matching Banaszczyk’s bound algorithmically for the problem and its variants has been open despite a lot of attention in recent years [Nik13, Buk13, ES14, EL15, DGLN16]. In particular, as we discuss in section 1.2 there is a natural algorithmic barrier to improving the $O(t^{1/2} \log n)$ bound.

A substantial generalization of the Beck–Fiala conjecture is the following.

Komlós conjecture. Given any collection of vectors $v_1, \ldots, v_n \in \mathbb{R}^m$ such that $\|v_i\|_2 \leq 1$ for each $i \in [n]$,² there exist signs $x_1, \ldots, x_n \in \{-1, 1\}$ such that $\|\sum_{i=1}^n x_i v_i\|_\infty = O(1)$.

This implies the Beck–Fiala conjecture by choosing each $v_i$ as the column corresponding to element $i$ in the incidence matrix of the set system scaled by $t^{-1/2}$. Again, the best known nonconstructive bound here is $O(\sqrt{\log n})$ due to Banaszczyk and the previous algorithmic techniques can also be adapted to achieve $O(\log n)$ constructively for the Komlós setting.

1.1. Our results. In this paper we give the following algorithmic result for the Beck–Fiala problem, which matches the nonconstructive bound due to Banaszczyk.

**Theorem 1.** Given a set system $(V, S)$ with $|V| = n$ such that each element $i \in V$ lies in at most $t$ sets in $S$, there is an efficient randomized algorithm that finds an $O(\sqrt{\log n})$ discrepancy coloring with high probability.

Our result also extends to the Komlós setting with some minor modifications.

**Theorem 2.** Given an $m \times n$ matrix $A$ with all columns of $\ell_2$-norm at most 1, there is an efficient randomized algorithm that finds $x \in \{-1, 1\}^n$ such that $\|Ax\|_\infty = O(\sqrt{\log n})$ with high probability.

Our algorithm gives a new constructive proof of Banaszczyk’s result for the Beck–Fiala and Komlós setting. While Theorem 2 implies Theorem 1, for better clarity we first present the algorithm for the Beck–Fiala problem in sections 2 and 3 and then discuss the extension to Theorem 2 in section 4.

After this paper was published, an alternate algorithm using multiplicative weight updates giving the same result was given in [LRR16]. Also, the techniques in this paper were used to get improved bounds on other problems in discrepancy [BG16a].

1.2. High-level overview. The algorithm has a similar structure to the previous random walk based approaches [Ban10, LM12, HSS14]. It starts with the coloring $x_0 = 0^n$ at time 0, and at each time step $k$, updates the color of element $i$ by adding a small increment to its coloring at time $k−1$, i.e., $x_k(i) = x_{k−1}(i) + \Delta x_k(i)$. If a

---

¹ We assume here that $t \geq \log n$, otherwise the $O(t)$ bound is better.
² We use $[n]$ to denote the set {1, 2, ..., n}.
variable reaches $-1$ or $1$ it is frozen, and its value is not updated any more. The increment is determined by solving an appropriate SDP and projecting the resulting vectors in a random direction.

However, all the previous approaches get stuck at the $O(t^{1/2} \log n)$ barrier, and it is instructive to understand why this happens before we present our algorithm.

The $O(t^{1/2} \log n)$ barrier. Roughly speaking, the execution of the previous algorithms can be divided into $O(\log n)$ phases (either implicitly or explicitly), where in each phase about half the variables get frozen and each set incurs an expected discrepancy of $O(t^{1/2})$. This gives an overall discrepancy bound of $O(t^{1/2} \log n)$.

Intuitively, however, for a fixed set $S$, one should expect an $O(t^{1/2})$ discrepancy over all the phases for the following reason. Assume that all sets are of size $O(t)$. This can be ensured using a standard linear algebraic argument to ensure that sets incur zero discrepancy as long as they have at least $2t$ uncolored elements. After $i$ phases of partial coloring, one would expect that $S$ has about $2^{-i}$ fraction of its elements left uncolored and, hence, it should incur about $O((2^{-i}t)^{1/2})$ discrepancy in the next phase, giving a total discrepancy of $O(\sum_i 2^{-i/2} t^{1/2}) = O(t^{1/2})$.

However, the problem is that the size of sets may not evolve in this ideal manner, as the partial coloring phase does not give us a fine-grained control over how the elements of each set get colored. For example, even though half of the variables (globally) get colored during a phase, it is possible that half the sets get almost completely colored, while the other half only get $t/\log n$ of their elements colored (while still incurring an $\Omega(t^{1/2})$ discrepancy). This imbalance between the discrepancy incurred and “progress” made for each set is the fundamental barrier in overcoming the $O(t^{1/2} \log n)$ bound.

Our approach. The key idea of our algorithm is to ensure that during the coloring updates the squared discrepancy we add to a set is proportional to the progress elements of that set make towards getting colored. More formally, the updates $\Delta x_k(i)$ that we choose at time $k$ satisfy the following properties:

1. Zero discrepancy for large sets: If a set $S$ has more than $a$ unfrozen (alive) elements at time $k$, for some constant $a$, we ensure that $\sum_{i \in S} \Delta x_k(i) = 0$. This is similar to the previous approaches and allows us to not worry about the discrepancy of a set until its size falls below $a$.

2. Proportional discrepancy property: This is the key new property and (roughly speaking) ensures that the squared discrepancy added to a set is proportional to the “energy” injected into the set. That is,

$$\left(\sum_{i \in S} \Delta x_k(i)\right)^2 \leq 2 \left(\sum_{i \in S} \Delta x_k(i)^2\right).$$

Note that the left-hand side is the square of the discrepancy increment for set $S$, and the right-hand side is the sum of squares of the increments of the elements of $S$.

Given a coloring $x_k$, let us define the energy of set $S$ at time $k$ as $\sum_{i \in S} x_k(i)^2$. Clearly, the energy of a set can never exceed its size $|S|$. As we can assume that $|S| = O(t)$ (by the zero discrepancy property above), this property suggests that if the total energy injected ($\sum_k (\sum_{i \in S} \Delta x_k(i)^2)$) into $S$ was
comparable to its final energy (which is \( O(t) \)) and the increments were mean 0 random variables, the squared discrepancy should be \( O(t) \).

3. **Approximate orthogonality constraints** to relate the injected energy to actual energy: One big problem with the above idea is that the total injected energy into a constraint may be unrelated to its final energy. For example, even for a single variable \( i \) if the coloring \( x_k(i) \) "fluctuates" a lot around 0 over time, the injected energy \( \sum_k \Delta x_k(i)^2 \) could be arbitrarily large, while the final energy for \( i \) is at most 1. For general sets \( S \), other problems can arise beyond just fluctuations due to correlations between the updates of different elements of \( S \).

To fix this we use the following idea. Suppose we could ensure that for each set \( S \) the coloring update at time \( k \) was orthogonal to the coloring at time \( k - 1 \), i.e., \( \sum_{i \in S} x_{k-1}(i) \Delta x_k(i) = 0 \). Then, by the Pythagoras theorem, the increase in energy of \( S \) would satisfy

\[
\sum_{i \in S} x_k(i)^2 - \sum_{i \in S} x_{k-1}(i)^2 = \sum_{i \in S} (x_{k-1}(i) + \Delta x_k(i))^2 - x_{k-1}(i)^2 \\
= 2 \sum_{i \in S} x_{k-1}(i) \Delta x_k(i) + \sum_{i \in S} \Delta x_k(i)^2 \\
= \sum_{i \in S} \Delta x_k(i)^2,
\]

where the last equality follows from the orthogonality constraint. As the expression in (1) is the injected energy at time \( k \), this would precisely make the total injected equal to the final energy as desired.

However, we cannot add such constraints directly for each small set as there might be too many of them. So the idea is to add a weaker version of these orthogonality constraints, where we only require that

\[
\left( \sum_{i \in S} x_{k-1}(i) \Delta x_k(i) \right)^2 \leq 2 \left( \sum_{i \in S} \Delta x_k(i)^2 \right)
\]

and show that these suffice for our purpose.

4. **Sufficient progress property**: Of course, all the properties above can be trivially satisfied by setting \( \Delta x_k(i) = 0 \) for each \( i \). So the final step is to ensure that a nontrivial update exists. To this end, we show that there exist updates with the (unnormalized) sum \( \sum_i \Delta x_k(i)^2 = \Omega(A_k) \), where \( A_k \) is the number of alive variables at time \( k \).

For this purpose, we write a semidefinite program (SDP) that captures the above constraints and use duality to show the existence of a large feasible solution.

A weaker version of these properties was used in the arXiv preprint [BG16b] to get a more size-sensitive discrepancy bound for each set, but it still only achieved an \( O(t^{1/2} \log n) \) discrepancy in the worst case.

We now describe the algorithm and the SDP we use in section 2. The analysis consists of two main parts. In section 3.1 we show the sufficient progress property mentioned above, and in section 3.2 we show how this gives an overall discrepancy bound of \( O((t \log n)^{1/2}) \).
2. Algorithm for the Beck–Fiala problem. We will index time by $k$. Let $x_k \in [-1,1]^n$ denote the coloring at the end of time step $k$. During the algorithm, variables which get set to at least $(1 - 1/n)$ in absolute value are called frozen and their values are not changed anymore. The remaining variables are called alive. We denote by $A_k$ the set of alive variables at the beginning of time step $k$. Initially all variables are alive. Let $\gamma = 1/(n^2 \log n)$, $T = (12/\gamma^2) \log n$, and $a = 6$.

We will call a set $S \in \mathcal{S}$ big at time $k$ if it has at least $at$ variables alive at time $k$, i.e., $|S \cap A_k| \geq at$ and small otherwise. We will use $\mathcal{B}_k$ to denote the collection of big sets at time $k$ and $\mathcal{L}_k$ to denote the collection of small (little) sets.

Algorithm:
1. Initialize $x_0(i) = 0$ for all $i \in [n]$ and $A_1 = \{1, 2, \ldots, n\}$.
2. For each time step $k = 1, 2, \ldots, T$ repeat the following:
   (a) Find a solution to the following semidefinite optimization problem:
   \[
   \text{Maximize } \sum_{i \in A_k} \|u_i\|_2^2 \\
   \text{s.t. } \left\| \sum_{i \in S \cap A_k} u_i \right\|_2^2 = 0 \quad \text{for each } S \in \mathcal{B}_k, \\
   \left\| \sum_{i \in S \cap A_k} u_i \right\|_2^2 \leq 2 \sum_{i \in S \cap A_k} \|u_i\|_2^2 \quad \text{for each } S \in \mathcal{L}_k, \\
   \left\| \sum_{i \in S \cap A_k} x_{k-1}(i)u_i \right\|_2^2 \leq 2 \sum_{i \in S \cap A_k} \|u_i\|_2^2 \quad \text{for each } S \in \mathcal{L}_k, \\
   \|u_i\|_2^2 \leq 1 \quad \forall i \in A_k.
   \]
   (b) Let $r_k \in \mathbb{R}^n$ be a random $\pm 1$ vector, obtained by setting each coordinate $r_k(i)$ independently to $-1$ or $1$ with probability $1/2$.
   For each $i \in A_k$, update $x_k(i) = x_{k-1}(i) + \gamma \langle r_k, u_i \rangle$. For each $i \notin A_k$, set $x_k(i) = x_{k-1}(i)$.
   (c) Initialize $A_{k+1} = A_k$.
   For each $i$, if $|x_k(i)| \geq 1 - 1/n$, update $A_{k+1} = A_{k+1} \setminus \{i\}$.
3. Generate the final coloring as follows. For the frozen elements $i \notin A_{T+1}$, set $x_T(i) = 1$ if $x_T(i) \geq 1 - 1/n$ and $x_T(i) = -1$ otherwise. For the alive elements $i \in A_{T+1}$, set them arbitrarily to $\pm 1$.

Note that the SDP at time $k$ uses the vectors $u_i$ to generate the update $\Delta x_k(i)$ by projecting $u_i$ to the random vector $r_k$ and scaling this by $\gamma$. If we think of $u_i$ as one dimensional vectors (so $\Delta x_k(i) = \gamma r u_i$, where $r$ is randomly $\pm 1$), constraints (2) will ensure that a set incurs zero discrepancy as long as it is big. Constraints (3) require the updates to satisfy the proportional discrepancy property mentioned earlier. Constraints (4) require the updates to satisfy the approximate orthogonality property mentioned earlier.

3. Analysis. We begin with some simple observations.

**Lemma 3.** For any vector $u \in \mathbb{R}^n$ and a random vector $r \in \{\pm 1\}^n$, $\mathbb{E}[\langle r, u \rangle^2] = \|u\|_2^2$ and $|\langle r, u \rangle| \leq \sqrt{n}\|u\|_2$. 

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
Proof. Writing \( u \) in terms of its coordinates \( u = (u(1), \ldots, u(n)) \),

\[
\mathbb{E}[(r, u)^2] = \mathbb{E}\left[\left(\sum_i r(i)u(i)\right)^2\right] = \sum_{i,j} \mathbb{E}[r(i)r(j)]u(i)u(j) = \|u\|_2^2,
\]

where the last equality uses that \( \mathbb{E}[r(i)r(j)] = 0 \) for \( i \neq j \) and \( \mathbb{E}[r(i)^2] = 1 \).

The second part follows by the Cauchy–Schwarz inequality, as \( |\langle r, u \rangle| \leq \|r\|_2\|u\|_2 = \sqrt{n}\|u\|_2 \).

This implies the following.

Observation 4. The rounding of frozen elements in step 3 of the algorithm affects the discrepancy of any set by at most \( n \cdot (1/n) = 1 \). So we can ignore this rounding error. Moreover, as \( \|u_i\|_2 \leq 1 \), \( |\langle r, u_i \rangle| \leq \gamma\sqrt{n}\|u_i\|_2 \leq 1/n \), which implies that no \( x_k(i) \) goes out of the range \([-1, 1]\) during any step of the algorithm.

The rest of the analysis is divided into three parts. In section 3.1, we show that the SDP is feasible and has value at least \( |A_k|/3 \) at each time step \( k \). In section 3.2, we use the properties of the SDP to show that each set in \( S \) has discrepancy \( O((t \log n)^{1/2}) \) after \( T \) steps with high probability. Finally, in section 3.3 we show that there are no alive elements after \( T \) steps with high probability. Together these will imply Theorem 1.

3.1. SDP is feasible and has value \( \Omega(|A_k|) \). To show that the SDP has value at least \( |A_k|/3 \) at any time step \( k \), we will consider the dual and show that no solution with objective value less than \( |A_k|/3 \) can be feasible. By strong duality, this suffices as if the optimum (primal) SDP solution was less than \( |A_k|/3 \), there would also be some feasible dual solution with that value (provided Slater’s conditions are satisfied).

It might be useful to point out here that the feasibility of our SDP is incomparable to the main result in [Nik13]; we can ensure a zero discrepancy to a few rows, which was not possible in the approach used in [Nik13] but we can only ensure a partial coloring \( \left(\sum_i \|u_i\|_2^2\right) \geq |A_k|/3 \), whereas the SDP in [Nik13] was feasible with the stronger constraint \( \|u_i\|_2 = 1 \) for all \( i \).

To make it easier to write the dual, we rewrite the SDP in the following matrix notation by setting \( X \) to be the Gram matrix of vectors corresponding to alive elements, i.e., \( X_{ij} = \langle u_i, u_j \rangle \) for \( i,j \in A_k \):

Maximize \( I \cdot X \) subject to

\[
\begin{align*}
&v_Sv_S^T \cdot X = 0 \quad \text{for each } S \in B_k, \\
&(v_Sv_S^T - 2I_S) \cdot X \leq 0 \quad \text{for each } S \in L_k, \\
&(x_Sx_S^T - 2I_S) \cdot X \leq 0 \quad \text{for each } S \in L_k, \\
&(e_i e_i^T) \cdot X \leq 1 \quad \forall i \in A_k, \\
&X \succeq 0.
\end{align*}
\]

Here \( v_S \) is the indicator vector of set \( S \cap A_k \), \( x_S \) is the vector with its \( i \)th entry equal to \( x_{k-1}(i) \) if \( i \in S \cap A_k \) and 0 otherwise, and \( I_S \) is the identity matrix restricted to set \( S \cap A_k \), i.e., \( (I_S)_{ij} = 1 \) if \( i \in S \cap A_k \) and 0 otherwise. \( \cdot \) denotes the usual inner product on matrices, \( A \cdot B = \text{Tr}(A^T B) = \sum_{ij} A_{ij} B_{ij} \).
We can write the dual of the above SDP (for reference, see [GM12]), which is given by

\[
\begin{align*}
\text{Minimize} & \quad \sum_{i \in A_k} b_i \\
\text{s.t.} & \quad \sum_{i \in A_k} b_i e_i^T + \sum_{S \in B_k} \alpha_S v_S v_S^T + \sum_{S \in L_k} (\beta_S (v_S v_S^T - 2 I_S) + \beta_S^* (x_S x_S^T - 2 I_S)) \geq I,
\end{align*}
\]

(5) \hspace{1cm} \beta_S, \beta_S^* \geq 0 \quad \forall S \in L_k,

(6) \hspace{1cm} b_i \geq 0 \quad \forall i \in A_k,

(7) \hspace{1cm} \alpha_S \in \mathbb{R} \quad \forall S \in B_k,

(8) \hspace{1cm} x_S, x_S^T \geq 0 \quad \forall S \in L_k,

Here $A \succeq B$ denotes that the matrix $A - B$ is positive semidefinite. To show strong duality we use the following result.

**Theorem 5** (Theorem 4.7.1 in [GM12]). If the primal program $(P)$ is feasible, has a finite optimum value $\eta$, and has an interior point $\tilde{x}$, then the dual program $(D)$ is also feasible and has the same finite optimum value $\eta$.

**Lemma 6.** The SDP described above is feasible and has value equal to its dual program.

**Proof.** We apply Theorem 5, with $P$ equal to the dual of the SDP. This would suffice as the dual of $P$ is our SDP.

We claim that $b_i = 1 + \epsilon$ for $\epsilon > 0$ for all $i \in A_k$, $\alpha_S = 0$ for all $S \in B_k$, and $\beta_S = \beta_S^* = \epsilon/(8n^2)$ for all $S \in L_k$ is a feasible interior point for $P$. Clearly, this solution is strictly in the interior of the constraints (6)–(8). That (5) is satisfied and has slack in every direction follows as the the number of sets $S$ can be at most $t|A_k| \leq tn \leq n^2$, and that for any vector $v$, $vv^T$ is a rank one positive semi-definite matrix with eigenvalue $\|v\|_2^2 \leq n$, and thus all eigenvalues of $v_S v_S^T - 2 I_S$ and $x_S x_S^T - 2 I_S$ lie in the range $[-2, n]$.

As this point has objective value at most $(1+\epsilon)n$ and since the $b_i$ are nonnegative, $P$ has a finite optimum value.

We wish to show that any feasible solution to the dual must satisfy $\sum_i b_i \geq |A_k|/3$. To do this, we will show that there is a large subspace $W$ of dimension at least $|A_k|/3$ where the operator

\[
\sum_{S \in B_k} \alpha_S v_S v_S^T + \sum_{S \in L_k} (\beta_S (v_S v_S^T - 2 I_S) + \beta_S^* (x_S x_S^T - 2 I_S))
\]

is negative semidefinite. This would imply that to satisfy (5), the $b_i$’s have to be quite large on average. We first give two general lemmas.

**Lemma 7.** Given an $h \times n$ matrix $M$ with columns $z_1, z_2, \ldots, z_n$. If $\|z_i\|_2 \leq 1 \forall i \in [n]$, then there exists a subspace $W$ of $\mathbb{R}^n$ satisfying

(i) $\dim(W) \geq \frac{n}{2}$, and

(ii) $\forall y \in W^\perp$, $\|M y\|_2^2 \leq 2\|y\|_2^2$.

**Proof.** Let the singular value decomposition of $M$ be given by $M = \sum_{i=1}^n \sigma_i p_i q_i^T$, where $0 \leq \sigma_1 \leq \cdots \leq \sigma_n$ are the singular values of $M$ and $\{p_i : i \in [n]\}, \{q_i : i \in [n]\}$ are two sets of orthonormal vectors (if $h < n$, some $p_i$’s and the corresponding $\sigma_i$’s will be zero). Then,

\[
\sum_{i=1}^n \sigma_i^2 = \text{Tr} \left[ \sum_{i=1}^n \sigma_i^2 q_i q_i^T \right] = \text{Tr} [M^T M] = \sum_{i=1}^n \|z_i\|_2^2 \leq n.
\]
So at least \(\left\lceil \frac{n}{\sqrt{2}} \right\rceil\) of the squared singular values \(\sigma^2_i\)'s have value at most 2, and thus \(\sigma_1 \leq \cdots \leq \sigma_{\left\lceil \frac{n}{\sqrt{2}} \right\rceil} \leq \sqrt{2}\). Let \(W = \text{span}\{q_1, \ldots, q_{\left\lceil \frac{n}{\sqrt{2}} \right\rceil}\}\). For \(y \in W\),

\[
\|My\|_2^2 = \left\| \sum_{i=1}^n \sigma_i p_i q_i^T y \right\|_2^2 = \left\| \sum_{i=1}^\left\lceil \frac{n}{\sqrt{2}} \right\rceil \sigma_i p_i q_i^T y \right\|_2^2
\]

\[
\leq \sum_{i=1}^\left\lceil \frac{n}{\sqrt{2}} \right\rceil \sigma_i^2 (q_i^T y)^2 \quad \text{(since the } p_i \text{ are orthonormal)}
\]

\[
\leq 2 \sum_{i=1}^\left\lceil \frac{n}{\sqrt{2}} \right\rceil (q_i^T y)^2
\]

\[
= 2 \|y\|_2^2 \quad \text{(since the } q_i \text{ are orthonormal).} \quad \Box
\]

This implies the following result.

**Theorem 8.** Let \(V\) be any finite collection of vectors \(v_1, \ldots, v_h\) in \(\mathbb{R}^n\), and for each \(v \in V\), there is some nonnegative multiplier \(\beta_v \geq 0\). Consider the operator

\[
B = \sum_{v \in V} \beta_v \left( vv^T - 2 \sum_{i=1}^n \langle v, e_i \rangle^2 e_i e_i^T \right),
\]

where the \(e_i\) are the standard basis of \(\mathbb{R}^n\). Then there exists a subspace \(W\) of dimension at least \(n/2\) such that \(\langle y, By \rangle \leq 0\) for every \(y \in W\) or, equivalently, \(y^T By \leq 0\) for every \(y \in W\).

**Proof.** Let \(v_i\) denote \(\langle v, e_i \rangle\). We can express \(y^T By\) as

\[
B \cdot yy^T = \sum_v \beta_v \left( vv^T - 2 \sum_{i=1}^n \langle v, e_i \rangle^2 e_i e_i^T \right) \cdot yy^T
\]

\[
= \sum_v \beta_v \left( \sum_i v_i y_i \right)^2 - 2 \sum_i v_i^2 y_i^2
\]

Construct a matrix \(M\) with rows indexed by \(v\) for each \(v \in V\) and columns indexed by \(i \in [n]\). The entries of \(M\) are given by \(M_{e,i} = \beta_v^{1/2} v_i\). Then, we can write

\[
\sum_v \beta_v \left( \sum_i v_i y_i \right)^2 = \|My\|_2^2.
\]

For each \(i \in [n]\), define \(\beta_i^2 = \sum_v \beta_v v_i^2\) as the squared \(\ell_2\)-norm of column \(i\) of \(M\), and let \(D\) be an \(n \times n\) diagonal matrix with entries \(D_{ii} = \beta_i\). Then,

\[
\sum_v \beta_v \left( \sum_i v_i y_i \right)^2 - 2 \sum_i v_i^2 y_i^2 = \|My\|_2^2 - 2\|Dy\|_2^2.
\]

Let \(N \subseteq [n]\) be the set of coordinates with \(\beta_i > 0\). We claim that it suffices to focus on the coordinates in \(N\). Let us first observe that if \(i \notin N\), i.e., \(\beta_i^2 = 0\), then we can set \(y_i\) arbitrarily as (9) is unaffected. As the directions \(e_i\) for \(i \in N\) are orthogonal
to the directions in \([n] \setminus N\), it suffices to show that there is an \(|N|/2\) dimensional subspace \(W\) in \(\text{span}\{e_i : i \in N\}\) such that \(|M y|_2^2 - 2|D y|_2^2 \leq 0\) for each \(y \in W\). The overall subspace we desire is simply \(W \oplus \text{span}\{e_i : i \in [n] \setminus N\}\) which has dimension \(|N|/2 + (n - |N|) \geq n/2\).

So, let us assume that \(N = [n]\) (or, equivalently, restrict \(M\) and \(D\) to columns in \(N\)), which gives us that \(\beta_i > 0\) for all \(i \in N\) and hence that \(D\) is invertible.

Let \(M' = MD^{-1}\). The squared \(\ell_2\)-norm of each column in \(M'\) is \(\sum_v \beta_v v^2 / D_v^2\) which equals 1, and by Lemma 7, there is a subspace \(W'\) of dimension at least \(|N|/2\) such that \(|M' y'|_2^2 \leq 2|y'|_2^2\) for each \(y' \in W'\). Setting \(y = D^{-1} y'\) gives

\[
|M y|_2^2 = |M' y'|_2^2 \leq 2|y'|_2^2 = 2|D y|_2^2,
\]
and thus \(W = \{D^{-1} y' : y' \in W'\}\) gives the desired subspace since \(\text{dim}(W) = \text{dim}(W')\).

Going back to the dual SDP, this gives the following.

**Lemma 9.** Let \(B_k = \sum_{S \in \mathcal{L}_k} (\beta_S (v_S v_S^T - 2I_S) + \beta_S^2 (x_S x_S^T - 2I_S))\). Then, there exists a subspace \(W \subseteq \mathbb{R}^{|A_k|}\) of dimension at least \(|A_k|/2\) such that for all \(y \in W\), \(y^T B_k y \leq 0\).

**Proof.** We apply Theorem 8 with vectors \(v_S\) and \(x_S\) for each small set \(S \in \mathcal{L}_k\) with the multipliers \(\beta_S\) and \(\beta_S^2\). Then,

\[
B = \sum_{S \in \mathcal{L}_k} \left[ \beta_S \left( v_S v_S^T - 2 \sum_{i \in A_k} \langle v_S, e_i \rangle^2 e_i e_i^T \right) + \beta_S^2 \left( x_S x_S^T - 2 \sum_{i \in A_k} \langle x_S, e_i \rangle^2 e_i e_i^T \right) \right]
= \sum_{S \in \mathcal{L}_k} \left[ \beta_S (v_S v_S^T - 2I_S) + \beta_S^2 (x_S x_S^T - 2I_S) \right]
\geq \sum_{S \in \mathcal{L}_k} (\beta_S (v_S v_S^T - 2I_S) + \beta_S^2 (x_S x_S^T - 2I_S))
= B_k.
\]

Here we use that \(v_S\) is the indicator vector for set \(S \cap A_k\) with entries \(\langle v_S, e_i \rangle = 1\) iff \(i \in S \cap A_k\) and, thus, \(\sum_{i \in A_k} \langle v_S, e_i \rangle^2 e_i e_i^T = I_S\). Similarly for the vectors \(x_S\), \(\langle x_S, e_i \rangle = x_{k-1}(i)\) for \(i \in S \cap A_k\) and 0 otherwise. The last step uses that \(x_{k-1}(i)^2 \leq 1\) and thus

\[
-2 \sum_{i \in S \cap A_k} x_{k-1}(i)^2 e_i e_i^T \geq -2I_S.
\]

By Theorem 8, there is a subspace \(W\) with \(\text{dim}(W) \geq |A_k|/2\) such that \(y^T B_k y \leq 0\) for each \(y \in W\). As \(B \geq B_k\), it also holds that \(y^T B_k y \leq 0\) for each \(y \in W\).

We now come to the main theorem of this subsection.

**Theorem 10.** At time step \(k\), the dual program has value at least \(|A_k|/3\).

**Proof.** As element \(i\) in \(A_k\) appears in at most \(t\) sets, the number of big sets \(|B_k|\) at time step \(k\) is at most \(|A_k|/t\). Let \(W_1\) be the subspace orthogonal to \(C = \text{span}\{v_S : S \in B_k\}\). Clearly, \(\text{dim}(C) \leq |B_k| \leq |A_k| / a\).

Let \(W_0\) be the subspace guaranteed by Lemma 9 for matrix \(B_k\) such that \(\text{dim}(W_0) \geq |A_k|/2\) and for all \(y \in W_0\), \(y^T B_k y \leq 0\). Define the subspace \(W = W_1 \cap W_0\). Then,

\[
\text{dim}(W) \geq \text{dim}(W_0) - \text{dim}(C) \geq |A_k|/2 - |A_k|/a = |A_k|/3.
\]
Let $P_W$ be the projection operator on the subspace $W$. Projecting the dual constraint (5) onto $W$, we get

$$P_W \left( \sum_{i \in A_k} b_i e_i e_i^T + \sum_{S \in B_k} \alpha_S v_S v_S^T + B_k \right) \succeq P_W.$$  

By linearity of $P_W$ and as $P_W(v_S v_S^T) = 0$ for each $S \in B_k$, this implies

$$P_W \left( \sum_{i \in A_k} b_i e_i e_i^T \right) + P_W(B_k) \succeq P_W.$$  

Taking the trace on both sides and noting that $\text{Tr}[P_W(B_k)] \leq 0$ since $y^T B_k y \leq 0$ for all $y \in W$, we get

$$\text{Tr} \left[ P_W \left( \sum_{i \in A_k} b_i e_i e_i^T \right) \right] \geq \text{Tr}[P_W] = \dim(W).$$  

As $b_i \geq 0$ for all $i \in A_k$, the operator $\sum_{i \in A_k} b_i e_i e_i^T$ is positive semidefinite. As taking the projection of a positive semidefinite operator can only decrease its trace, we can lower bound the dual objective as

$$\sum_{i \in A_k} b_i = \text{Tr} \left( \sum_{i \in A_k} b_i e_i e_i^T \right) \geq \text{Tr} \left[ P_W \left( \sum_{i \in A_k} b_i e_i e_i^T \right) \right] \geq \text{Tr}[P_W] = \dim(W) \geq |A_k|/3$$

which completes the proof. \hfill \qed

### 3.2. Bounding the discrepancy

Let $D_S(k)$ denote the signed discrepancy of set $S \in \mathcal{S}$ at the end of time step $k$, i.e., $D_S(k) = \sum_{i \in S} x_k(i)$.

We now show the following key result.

**Theorem 11.** Fix a set $S \in \mathcal{S}$. Then, for any $\lambda \geq 0$, the discrepancy of $S$ at time step $T$ satisfies

$$\Pr \left[ |D_S(T)| \geq \lambda \sqrt{T} \right] \leq 8 \exp \left( -\lambda^2 / (100a) \right).$$

Setting $\lambda = O(\log^{1/2} n)$ would imply that with high probability every set has discrepancy $O((t \log n)^{1/2})$ at time $T$.

Among other things, the proof of Theorem 11 will use a powerful concentration inequality for martingales due to Freedman that we describe below.

**Martingales and Freedman’s inequality.** Let $X_1, X_2, \ldots, X_n$ be a sequence of independent random variables on some probability space, and let $Y_k$ be a function of $X_1, \ldots, X_k$. The sequence $Y_0, Y_1, Y_2, \ldots, Y_n$ is called a martingale with respect to the sequence $X_1, \ldots, X_n$ if for all $k \in [n]$, $\mathbb{E}[Y_k | Y_{k-1}]$ is finite and $\mathbb{E}[Y_k | X_1, X_2, \ldots, X_{k-1}] = Y_{k-1}$. We will use $\mathbb{E}_{k-1}[Z]$ to denote $\mathbb{E}[Z | X_1, X_2, \ldots, X_{k-1}]$, where $Z$ is any random variable.

**Theorem 12** (Freedman [Fre75]). Let $Y_0, \ldots, Y_n$ be a martingale with respect to $X_1, \ldots, X_n$ such that $|Y_k - Y_{k-1}| \leq M$ for all $k$, and let

$$W_k = \sum_{j=1}^{k} \mathbb{E}_{j-1} [(Y_j - Y_{j-1})^2] = \sum_{j=1}^{k} \text{Var}[Y_j | X_1, \ldots, X_{j-1}].$$
Then for all \( \lambda \geq 0 \) and \( \sigma^2 \geq 0 \), we have

\[
\Pr[|Y_n - Y_0| \geq \lambda \text{ and } W_n \leq \sigma^2] \leq 2 \exp \left( -\frac{\lambda^2}{2(\sigma^2 + M\lambda/3)} \right).
\]

Observe crucially that the above inequality is much more powerful than the related Azuma–Hoeffding or Bernstein’s inequality. In particular, the term \( W_n \) is the variance encountered by the martingale on the particular sample path it took, as opposed to a worst case bound on the variance over all possible paths.

**Simple observations.** We now get back to the proof of Theorem 11 and begin with a few simple observations.

Fix a set \( S \in \mathcal{S} \). Let the vector solution returned by the SDP at time \( k \) be given by vectors \( u^k_i \) for \( i \in [n] \), where we take \( u^k_i = 0 \) if \( i \notin A_k \). We say that \( S \) becomes active at time \( k \) if \( k \) is the first time step when \( |S \cap A_k| \leq \lambda t \).

**Observation 13.** Before a set \( S \in \mathcal{S} \) is active, it incurs zero discrepancy.

**Proof.** Suppose \( S \) becomes active at time \( k_S \). Then,

\[
D_S(k_S - 1) = \sum_{k=1}^{k_S-1} \gamma(r_k, \sum_{i \in S \cap A_k} u^k_i) = 0,
\]

since by SDP constraint (2), \( \sum_{i \in S \cap A_k} u^k_i = 0 \) for \( k < k_S \).

**Observation 14.** As a set has no more than \( at \) alive variables when it becomes active, Observation 13 implies that the maximum discrepancy any set can have is \( 2at \), which Theorem 11 for \( \lambda > 2at^{1/2} \). So henceforth we can assume that \( \lambda \leq 2at^{1/2} \).

Define the energy of set \( S \) at the end of time step \( k \) as \( E_S(k) = \sum_{i \in S} x_k(i)^2 \) and the change in energy of \( S \) at time step \( k \) as \( \Delta_k E_S = E_S(k) - E_S(k - 1) \). Then,

\[
\Delta_k E_S = \sum_{i \in S} x_k(i)^2 - \sum_{i \in S} x_{k-1}(i)^2 = \sum_{i \in S} \left( (x_{k-1}(i) + \gamma \langle r_k, u^k_i \rangle)^2 - x_{k-1}(i)^2 \right)
\]

\[
= \gamma^2 \sum_{i \in S} \langle r_k, u^k_i \rangle^2 + 2\gamma \langle r_k, \sum_{i \in S} x_{k-1}(i)u^k_i \rangle.
\]

The following is a simple but crucial observation.

**Observation 15.** Once a set \( S \in \mathcal{S} \) becomes active, its energy can increase overall by at most \( at \).

**Proof.** When \( S \) becomes active, it has at most \( at \) alive variables. Moreover, a frozen variable is never updated by the algorithm and can never become alive again. As the energy of a single variable is bounded by 1, the energy of \( S \) can increase by at most \( at \) after it becomes active.

**Remark.** Note that the energy of a set \( S \) does not necessarily increase monotonically over time. It evolves randomly and can also decrease. So, even though the overall increase is at most \( at \), the total energy “injected” \( \sum_{k \geq k_S} |\Delta_k E_S| \) can be arbitrarily larger than \( at \). Here \( k_S \) denotes the time when \( S \) becomes active.

By Observation 13, we only need to bound the discrepancy of \( S \) after it becomes active. For notational convenience, let us call the time \( S \) becomes active as time 0. So, \( D_S(k) \) and \( E_S(k) \) will be the signed discrepancy and energy of \( S \), respectively, \( k \) time steps after it becomes active.
Observation 16. After $S$ becomes active, $D_S(k)$ behaves like a martingale with variance of the increment at time step $k$ bounded by
\[
\mathbb{E}_{k-1} \left[ (D_S(k) - D_S(k-1))^2 \right] \leq 2\gamma^2 \sum_{i \in S' \setminus A_k} \|u_i^k\|^2.
\]

Proof. The discrepancy update of $S$ at time $k$ is $\gamma (r_k, \sum_{i \in S' \setminus A_k} u_i^k)$. This has expectation 0 (averaging of $r_k$) and by Lemma 3 this variance is exactly $\gamma^2 \sum_{i \in S' \setminus A_k} \|u_i^k\|^2$, which by SDP constraint (3) is upper bounded by $2\gamma^2 \sum_{i \in S' \setminus A_k} \|u_i^k\|^2$.

Proof of Theorem 11. The plan of the proof is the following. Freedman’s inequality allows us to bound the discrepancy at time $T$ as a function of the variance $\sum_{k=1}^T \mathbb{E}_{k-1} [(D_S(k) - D_S(k-1))^2]$ which is at most $2\gamma^2 \sum_{k=1}^T \sum_{i \in S' \setminus A_k} \|u_i^k\|^2$ by Observation 16. As we will see, this term is expected total energy injected into $S$.

As the overall energy increase of $S$ can be at most at (Observation 15), it would suffice to show that the injected energy into $S$ is comparable to $at$. To do this, we will use the approximate orthogonality constraints (4) and apply Freedman’s inequality again to show that the injected energy is tightly concentrated around the energy increase. We now give the details.

Recall that by (10), the energy change at time $k$ is a random variable given by
\[
\Delta_k E_S = \gamma \sum_{i \in S} \langle r_k, u_i^k \rangle^2 + 2\gamma \sum_{i \in S} x_{k-1}(i) u_i^k.
\]
Denote the first term above as
\[
\Delta_k Q_S = \gamma^2 \sum_{i \in S} \langle r_k, u_i^k \rangle^2
\]
which we will call the change in quadratic energy of $S$ at time step $k$ and let $Q_S(k) = \sum_{j=1}^k \Delta_j Q_S$, the total quadratic energy of $S$ until time $k$.

Similarly, denote the second term as
\[
\Delta_k L_S = 2\gamma \sum_{i \in S} x_{k-1}(i) u_i^k
\]
which we will call the change in linear energy of $S$ at time step $k$, and let $L_S(k) = \sum_{j=1}^k \Delta_j L_S$, the total linear energy of $S$ until time $k$. The energy of $S$ at time $k$ is given by $E_S(k) = Q_S(k) + L_S(k)$.

Define $Q'_S(k)$ as
\[
Q'_S(k) = \sum_{j=1}^k \mathbb{E}_{j-1} [\Delta_j E_S] = \sum_{j=1}^k \mathbb{E}_{j-1} [\Delta_j Q_S].
\]
By Lemma 3,
\[
Q'_S(k) = \sum_{j=1}^k \gamma^2 \sum_{i \in S} \|u_i^j\|^2.
\]

We are now ready to prove the tail bound on discrepancy. The probability that the discrepancy of $S$ at time $T$ exceeds $\lambda \sqrt{T}$ can be written as
\[
(11) \quad \Pr \left[ |D_S(T)| \geq \lambda \sqrt{T} \right] \leq \Pr \left[ |D_S(T)| \geq \lambda \sqrt{T}, Q'_S(T) \leq 16at \right] + \Pr \left[ Q'_S(T) > 16at \right].
\]
We now bound each of the terms in (11) separately.
Bounding the first term. Recall that $D_S(k)$ is a martingale. To apply Freedman’s inequality (Theorem 12) we bound $M$ and $W_k$ as follows. By Lemma 3,

$$M \leq |D_S(k) - D_S(k - 1)| = \left| \gamma \left\langle r_k, \sum_{i \in S} u_i^k \right\rangle \right| \leq \gamma \sqrt{n} \left\| \sum_{i \in S} u_i^k \right\|_2 \leq \gamma n^{3/2}.$$  

Similarly from Lemma 3 and the SDP constraint (3),

$$W_k = \sum_{j=1}^k \mathbb{E}_{j-1} \left( (D_S(j) - D_S(j - 1))^2 \right) = \sum_{j=1}^k \mathbb{E}_{j-1} \left[ \left( \gamma^2 \left\langle r_j, \sum_{i \in S} u_i^j \right\rangle \right)^2 \right]$$

$$= \sum_{j=1}^k \gamma^2 \left\| \sum_{i \in S} u_i^j \right\|_2^2 \leq \sum_{j=1}^k 2\gamma^2 \sum_{i \in S} \left\| u_i^j \right\|_2^2 = 2Q'_S(k).$$

Freedman’s inequality now gives

$$\Pr \left[ |D_S(T)| \geq \lambda \sqrt{T} \text{ and } Q'_S(T) \leq 16at \right] \leq \Pr \left[ |D_S(T)| \geq \lambda \sqrt{T} \text{ and } W_T \leq 32at \right]$$

$$\leq 2 \exp \left( \frac{-\lambda^2 t}{2[32at + \gamma n^{3/2} \lambda \sqrt{T}/3]} \right) \leq 2 \exp \left( \frac{-\lambda^2}{100at} \right) \text{ (using } \lambda \leq 2\sqrt{t}). \quad (12)$$

Bounding the second term. We can write

$$\Pr[Q'_S(T) > 16at] = \sum_{j=0}^{\infty} \Pr[2^{j+4}at < Q'_S(T) \leq 2^{j+5}at]$$

$$\leq \Pr[Q'_S(T) \leq Q'_S(T) - 8at]$$

$$+ \sum_{j=0}^{\infty} \Pr[Q'_S(T) \leq 2^{j+5}at, Q'_S(T) \geq 2^{j+3}at]. \quad (13)$$

The inequality above holds as the event $\{Q'_S(T) > 16at\}$ is contained in the union of the two events in (13).

As the energy $E_S(T)$ of $S$ cannot exceed $at$, we have $E_S(T) = L_S(T) + Q_S(T) \leq at$. Thus, $Q_S(T) \geq 2^{j+3}at$ implies $L_S(T) \leq at - 2^{j+3}at \leq -7 \cdot 2^jat$, giving

$$\Pr[Q'_S(T) > 16at] \leq \Pr[Q'_S(T) \leq Q'_S(T) - 8at]$$

$$+ \sum_{j=0}^{\infty} \Pr[L_S(T) \leq -7 \cdot 2^jat, Q'_S(T) \leq 2^{j+5}at]. \quad (14)$$

To bound the second term on the right-hand side of (14), we will crucially use the approximate orthogonality constraints in the SDP (4) and use Freedman’s inequality.
To this end, note that $L_S(k)$ is a martingale whose difference sequence can be bounded by

$$M \leq |L_S(k) - L_S(k - 1)| = \left| 2\gamma \left< r_k, \sum_{i \in S} x_{k-1}(i) u_i^k \right> \right|$$

$$\leq 2\gamma \sqrt{n} \left\| \sum_{i \in S} x_{k-1}(i) u_i^k \right\|_2 \leq 2\gamma n^{3/2},$$

where we used Lemma 3 in the first inequality and the fact that $|x_{k-1}(i)| \leq 1$.

By Lemma 3 and SDP constraint (4),

$$W_k = \sum_{j=1}^k E_{j-1} \left[ |L_S(j) - L_S(j - 1)|^2 \right] = \sum_{j=1}^k E_{j-1} \left[ 4\gamma^2 \left< r_j, \sum_{i \in S} x_{j-1}(i) u_i^j \right> \right]$$

$$= \sum_{j=1}^k 4\gamma^2 \left\| \sum_{i \in S} x_{j-1}(i) u_i^j \right\|_2 \leq \sum_{j=1}^k 8\gamma^2 \sum_{i \in S} \| u_i^j \|_2^2 = 8Q_S'(k).$$

Applying Freedman’s inequality now with these bounds on $M$ and $W_k$, we obtain

$$\Pr \left[ |L_S(T)| \geq 7 \cdot 2^j at \text{ and } Q_S'(T) \leq 2^j + 5 at \right]$$

$$\leq \Pr \left[ |L_S(T)| \geq 7 \cdot 2^j at \text{ and } W_T \leq 2^j + 8 at \right]$$

$$\leq 2 \exp \left( \frac{-49 \cdot 2^{j+2}a^2}{2^{j+8}at + 2\gamma n^{3/2} \cdot 7 \cdot 2^j at / 3} \right)$$

$$\leq 2 \exp \left( \frac{-2^j at}{20} \right).$$

Together with $\lambda \leq 2a\sqrt{t}$ (by our assumption), this gives

$$\sum_{j=0}^{\infty} \Pr \left[ L_S(T) \leq -7 \cdot 2^j at, Q_S'(T) \leq 2^j + 5 at \right] \leq 4 \exp \left( \frac{-at}{20} \right) \leq 4 \exp \left( \frac{-\lambda^2}{100a} \right).$$

It remains to bound $\Pr[Q_S(T) \leq Q_S'(T) - 8at]$, the first term in (14). We use Freedman’s inequality in a simple way (even the Azuma–Hoeffding inequality would suffice here).

Define the martingale $Z_k = Q_S(k) - \sum_{j=1}^k E_{j-1}[\Delta_j Q_S] = Q_S(k) - Q_S'(k)$ (this is the standard Doob decomposition of $\Delta_k Q_S$). By Lemma 3,

$$M \leq |Z_k - Z_{k-1}| = |\Delta_k Q_S - E_{k-1}[\Delta_k Q_S]| \leq 2|\Delta_k Q_S| \leq 2\gamma^2 n \sum_{i \in S} \| u_i^k \|_2^2 \leq 2\gamma^2 n^2.$$

Using the trivial bound $E_{j-1}[|Z_j - Z_{j-1}|^2] \leq M^2$, we obtain that

$$W_T = \sum_{j=1}^T E_{j-1} \left[ |Z_j - Z_{j-1}|^2 \right] \leq 4T \gamma^4 n^4 = 48\gamma^2 n^4 \log n.$$
As $Q_S(T) \leq Q'_S(T) - 8at$ is the same as $Z_T \leq -8at$, by Freedman’s inequality we get
\[
\Pr[Q_S(T) \leq Q'_S(T) - 8at] \leq \Pr[|Z_T| \geq 8at] \\
\leq 2 \exp \left( \frac{-64a^2t^2}{2|W_T + 16\gamma^2n^2at/3|} \right) \\
\leq 2 \exp \left( \frac{-a^2t^2}{2\gamma^2n^4\log n} \right) \\
\leq 2 \exp \left( -\frac{a^2t^2}{2} \right) \leq 2 \exp \left( -\frac{\lambda^2}{100a} \right).
\]
(16)

In the last step we use that $\gamma = 1/(a^2 \log n)$, and Observation 14.

Combining (11), (12), (14), (15), and (16), we obtain the desired bound
\[
\Pr[|D_S(T)| - \lambda \sqrt{t}] \leq 8 \exp \left( -\frac{\lambda^2}{100a} \right).
\]

3.3. Termination and finishing the proof. To finish the proof, we show that the last rounding step at time $T + 1$ does not cause problems.

**Theorem 17.** After time $T$, there are no alive variables left with probability at least $1 - O(n^{-2})$.

**Proof.**

Given the coloring $x_k$ at time $k$, define $G_k = \sum_{i \in A_k} (1 - x_k(i)^2)$. Clearly $G_1 \leq n$. As $x_k(i) = x_{k-1}(i) + \gamma(r_k, u_i^k)$, we have that $E_{k-1}[x_k(i)^2] = x_{k-1}(i)^2 + \gamma^2 \|u_i^k\|^2$. It follows that
\[
E_{k-1}[G(k)] = E_{k-1} \left[ \sum_{i \in A_k} (1 - x_k(i)^2) \right] = \sum_{i \in A_k} (1 - x_{k-1}(i)^2) - \gamma^2 \sum_{i \in A_k} \|u_i^k\|^2 \\
\leq \sum_{i \in A_k} (1 - x_{k-1}(i)^2) - \gamma^2 |A_k|/3 \leq (1 - \gamma^2/3) \sum_{i \in A_k} (1 - x_{k-1}(i)^2) \\
\leq (1 - \gamma^2/3) \sum_{i \in A_{k-1}} (1 - x_{k-1}(i)^2) = (1 - \gamma^2/3) G_{k-1}.
\]

Thus by induction,
\[
E[G_{T+1}] \leq (1 - \gamma^2/3)^T G_1 \leq e^{-\gamma^2T/3} n = n^{-4} \cdot n = 1/n^3.
\]
Thus by Markov’s inequality, $\Pr[G_{T+1} \geq 1/n] \leq 1/n^2$. However, $G_{T+1} \leq 1/n$ implies that $A_{T+1} = 0$ as each alive variable contributes at least $1 - (1 - 1/n)^2 > 1/n$ to $G_{T+1}$. 

Theorem 1 now follows directly. Applying Theorem 11 with $\lambda = c \log^{1/2} n$ for $c$ a large enough constant and taking a union bound over the at most $nt \leq n^2$ sets, we get that $|D_S(T)| = O((t \log n)^{1/2})$ with probability at least $1 - 1/\text{poly}(n)$ for all sets $S$. By Theorem 17 with probability at least $1 - O(n^{-2})$, all variables are frozen by time $T$ and, hence, at most an additional discrepancy of 1 is added by rounding the frozen variables to $\pm 1$.

4. Extension to the Komlós setting. The algorithm also extends to the more general Komlós setting with some additional modifications. Recall that in the Komlós setting, we are given an $m \times n$ matrix $B$ with arbitrary real entries $b_{ji}$ such that for each column $i$, it holds that $\sum_j b_{ji}^2 \leq 1$. Let $r_j$ denote the $j$th row of $B$ and let $a$ be the constant as in the previous section. We will show the following result.
Theorem 18. Fix any row $r_j$ of matrix $B$. Then, for any $\lambda \geq 0$, the discrepancy of $r_j$ at time step $T$ (the end of the algorithm) satisfies

$$\Pr[|D_{T}(r_j)| \geq \lambda] \leq 8 \exp\left(-\lambda^{2}/(1000a)\right),$$

where $|D_{T}(r_j)|$ is the discrepancy of row $j$ after time step $T$.

The previous argument does not work directly when the entries $b_{ji}$ are arbitrary as we may not get a strong concentration if some entries $b_{ji}$ are too large. So we consider the following modified algorithm.

Algorithm. Given a matrix $B$, for any $\lambda > 0$ we denote by $r^\lambda_j$ the $\lambda$-truncation of row $j$ containing only the entries $b_{ji}$ that are at most $4a/\lambda$ in absolute value, i.e., $r^\lambda_j$ only contains those entries $i$ of row $j$ for which $|b_{ji}| \leq 4a/\lambda$ and is 0 otherwise.

As previously, let $A_k$ denote the set of alive variables at the beginning of time step $k$, and we set $\gamma = 1/n^6$ and $T = (12/\gamma^2) \log n$. A row $j$ is called big at time step $k$ if $\sum_{i \in A_k} b^2_{ji} > a$, and small otherwise. As the $\ell_2$-norm of columns of $B'$ is at most 1, there are at most $|A_k|/a$ big rows at any time step $k$.

The modified SDP. The SDP is modified as follows: Similarly to (2) we still require the discrepancy of big rows to be zero. That is,

$$\sum_{i \in A_k} b_{ji} u_i = 0 \quad \text{for each big row } j. \quad (17)$$

For an active (not big) row $r_j$ at time $k$, we add proportional discrepancy and approximate orthogonality constraints for every $\lambda$-truncation $r^\lambda_j$ of $r_j$, i.e., for every $\lambda > 0$, we add the proportional discrepancy constraint (3) (same as before, we just multiply the $u_i$’s by $b_{ji}$’s)

$$\sum_{i \in A_k, |b_{ji}| \leq 4a/\lambda} b_{ji} u_i \leq 2 \sum_{i \in A_k, |b_{ji}| \leq 4a/\lambda} b^2_{ji} \|u_i\|_2^2 \quad \text{and the approximate orthogonality constraints } (4) \quad (18)$$

and the approximate orthogonality constraints (4)

$$\sum_{i \in A_k, |b_{ji}| \leq 4a/\lambda} b^2_{ji} x_{k-1}(i) u_i \leq 2 \sum_{i \in A_k, |b_{ji}| \leq 4a/\lambda} b^4_{ji} \|u_i\|_2^2. \quad (19)$$

Notice that as stated, for each active row we add two SDP constraints for every value of $\lambda > 0$. However it suffices to add at most $2n$ constraints in total for each active row: Just sort the entries of a row in increasing order of absolute value and add the proportional discrepancy and orthogonality constraints in the SDP for every prefix of this sorted row (alternately, one could also consider geometrically increasing values of $\lambda$). Thus the SDP has a polynomial number of constraints at any time step.

Analysis. First, exactly as before the SDP is feasible and has a solution with value at least $|A_k|/3$. This follows from Theorem 8, which shows that there is a subspace $W$ of dimension at least $|A_k|/2$, where the corresponding operator is negative semidefinite on $W$, and then applying the argument in Theorem 10. In fact, this would be true even if (19) was replaced by the stronger constraint

$$\sum_{i \in A_k, |b_{ji}| \leq 4a/\lambda} b^2_{ji} x_{k-1}(i) u_i \leq 2 \sum_{i \in A_k, |b_{ji}| \leq 4a/\lambda} b^4_{ji} x_{k-1}(i)^2 \|u_i\|_2^2.$$
Let $D_k(r_j)$ denote the signed discrepancy of row $j$ at the end of time step $k$,

$$D_k(r_j) = \sum_{i \in [n]} b_{ji} x_k(i).$$

We also extend this definition to truncations of rows:

$$D_k(r_j^\lambda) = \sum_{i \in [n], |b_{ji}| \leq 4a/\lambda} b_{ji} x_k(i).$$

We now show Theorem 18. Fix a row $r_j$ and a $\lambda \geq 0$. Call an entry $b_{ji}$ large if $|b_{ji}| > 4a/\lambda$. We first make the following key observation.

**Observation 19.** When a row becomes active, the $\ell_1$-norm of the alive variables in that row that are large can be at most $\lambda/4$.

**Proof.** Each large entry is at least $4a/\lambda$ in absolute value. As a row $r_j$ becomes active when $\sum_{i \in A_k} b_{ji}^2 \leq a$, there can be at most $a/(4a/\lambda)^2 = \lambda^2/16a$ alive variables with $b_{ji}$ large. By the Cauchy–Schwarz inequality,

$$\left( \sum_{i \in A_k, |b_{ji}| > 4a/\lambda} |b_{ji}| \right) \leq \left( \sum_{i \in A_k} b_{ji}^2 \right)^{1/2} \left( \frac{\lambda^2}{16a} \right)^{1/2} \leq \lambda/4.$$

The above observation implies that when a row becomes active, the large entries in it can change discrepancy by at most $\lambda/2$. Thus to prove Theorem 18, it suffices to show

$$\Pr \left[ |D_T(r_j^\lambda)| \geq \lambda/2 \right] \leq 8 \exp \left( -\lambda^2/(1000a) \right).$$

This follows similarly to the analysis as before, using the proportional discrepancy (18) and approximate orthogonality constraints (19) for $r_j^\lambda$ and noting that (19) implies that

$$\left\| \sum_{i \in A_k, |b_{ji}| \leq 4a/\lambda} b_{ji}^2 x_{k-1}(i) u_i \right\|_2^2 \leq \sum_{i \in A_k, |b_{ji}| \leq 4a/\lambda} b_{ji}^4 \|u_i\|_2^2 \leq \frac{32a^2}{\lambda^2} \sum_{i \in A_k, |b_{ji}| \leq 4a/\lambda} b_{ji}^2 \|u_i\|_2^2$$

as $|b_{ji}| \leq 4a/\lambda$ for all entries in the truncated row $r_j^\lambda$. Let us define the energy of $\lambda$-truncation of row $j$ at time $k$ as

$$E_k(r_j^\lambda) = \sum_{i \in [n], |b_{ji}| \leq 4a/\lambda} b_{ji}^2 x_k(i)^2.$$

As previously, once the row becomes active, its energy can rise by at most $a$.

The analysis in section 3.2 had two main ideas:

1. First we showed that the expected squared discrepancy of a set $S$ at time $T$ was $O(1)$ times the energy injected into the set $Q_S(T)$ (using constraints (3)). This argument works exactly as before using constraints (18) and we sketch the details below.

For ease of notation we will denote the entries of the truncated row $r_j^\lambda$ as $b_{ji}$, where it is understood that we are setting $b_{ji} = 0$ if $b_{ji}$ was large in the original matrix. The change in energy at time $k$ is a random variable given by

$$\Delta_k E(r_j^\lambda) = \gamma^2 \sum_{i \in [n]} b_{ji}^2 \langle r_k, u_i^k \rangle^2 + 2\gamma \langle r_k, \sum_{i \in [n]} b_{ji}^2 x_{k-1}(i) u_i^k \rangle.$$
Denote the first term above as $\Delta_k Q(r_j^k)$, the change in quadratic energy of $r_j^k$ at time step $k$, and let $Q_k(r_j^k) = \sum_{k=1}^{k} \Delta_k Q(r_j^k)$, the total quadratic energy of $r_j^k$ until time $k$.

Similarly, denote the second term as $\Delta_k L(r_j^k)$, the change in linear energy of $r_j^k$ at time step $k$, and let $L_k(r_j^k) = \sum_{k=1}^{k} \Delta_k L(r_j^k)$, the total linear energy of $r_j^k$ until time $k$.

Define $Q'_k(r_j^k) = \sum_{k'=1}^{k} \mathbb{E}_{k'-1} [\Delta_{k'} E(r_j^{k'})] = \sum_{k'=1}^{k} \mathbb{E}_{k'-1} [\Delta_{k'} Q(r_j^{k'})]$. By Lemma 3,

$$Q'_k(r_j^k) = \sum_{k'=1}^{k} \gamma^2 \sum_{i \in [n]} \|u_i^{k'}\|_2^2.$$ 

Just as before, discrepancy $D_k(r_j^k)$ behaves as a martingale with the variance $W_k$ bounded by $2Q'_k(r_j^k)$. Freedman’s inequality then gives

$$(21) \quad \text{Pr} \left[ |D_T(r_j^T)| \geq \lambda/2 \text{ and } Q'_T(r_j^T) \leq 16a \right] \leq 2 \exp \left( \frac{-\lambda^2}{1000a} \right).$$

Next we showed that $Q'_S(T)$ was essentially the same as $Q_S(T)$ (shown in (16)). In fact this difference can be made arbitrarily small by reducing $\gamma$ and the argument here works exactly as before. In particular, we get

$$(22) \quad \text{Pr} \left[ Q_T(r_j^T) \leq Q'_T(r_j^T) - 8a \right] \leq 2 \exp \left( \frac{-\lambda^2}{1000a} \right).$$

2. The second part was to show that the linear term does not cause problems. In particular, the crucial argument was that $Q_S(T)$ cannot be much more than at as (i) the total rise in energy $L_S(T) + Q_S(T)$ cannot exceed at and (ii) $L(T)$ was a martingale with squared deviation comparable to $Q_S(T)$ and hence cannot be much larger than $Q_S^{1/2}(T)$. This step used the constraints (4).

This argument also works similarly in our setting here. For a truncated row $r_j^k$, $Q_T(r_j^k)$ cannot be much more than $a$ as (i) the total rise in energy $L_T(r_j^k) + Q_T(r_j^k)$ cannot exceed $a$ and (ii) $L_T(r_j^k)$ is a martingale with squared deviation comparable to $\frac{2\log n}{\lambda^2} Q_T(r_j^k)$ (by (19) and (20)). Proceeding exactly as before and applying Freedman’s inequality we obtain that,

$$(23) \quad \sum_{t=0}^{\infty} \text{Pr} \left[ L_T(r_j^t) \leq -7 \cdot 2^t a, \quad Q'_T(r_j^t) \leq 2^t + 5a \right] \leq 4 \exp \left( \frac{-\lambda^2}{1000a} \right).$$

Theorem 18 now follows by combining (21),(22),(23) as before and using Observation 19.

Theorem 2 now follows easily, by observing that $m$ can be assumed to be polynomially bounded in $n$ and applying a union bound. Indeed, we can discard all rows of $\ell_1$-norm less than $\sqrt{\log n}$ since they can only ever have discrepancy at most $\sqrt{\log n}$.

The remaining rows have squared $\ell_2$-norm at least $\frac{\log n}{n}$, as by the Cauchy–Schwarz inequality

$$\sqrt{\log n} \leq \sum_{i \in [n]} |b_{ji}| \leq \left( \sum_{i \in [n]} b_{ji}^2 \right)^{1/2} (n)^{1/2}.$$
As \( \sum_{i,j} b_{ij}^2 \leq n \), there can be at most \( n^2 / \log n \) such rows. We now set \( \lambda = O(\sqrt{\log n}) \) in Theorem 18 and take a union bound over all these rows.

**Acknowledgments.** We would like to thank Nick Harvey, Shachar Lovett, Aleksandar Nikolov, Thomas Rothvoss, Aravind Srinivasan, and Mohit Singh for various useful discussions about discrepancy over the years. A special thanks to Aleksandar Nikolov and Kunal Talwar for organizing a recent workshop on discrepancy at the American Institute of Mathematics which led to several fruitful discussions related to this work. This work was also done in part while the first author was visiting the Simons Institute for the Theory of Computing.

**REFERENCES**


