Discrepancy Without Partial Colorings

Nicholas J. A. Harvey[*] Roy Schwartz[†] Mohit Singh[†]

I cannot pretend to feel impartial about colours. I rejoice with the brilliant ones and am genuinely sorry for the poor browns.

Sir Winston Churchill

Abstract

Spencer's theorem asserts that, for any family of n subsets of ground set of size n, the elements of the ground set can be "colored" by the values ± 1 such that the sum of every set is $O(\sqrt{n})$ in absolute value. All existing proofs of this result recursively construct "partial colorings", which assign ± 1 values to half of the ground set. We devise the first algorithm for Spencer's theorem that directly computes a coloring, without recursively computing partial colorings.

1 Introduction

In combinatorics, the *discrepancy* problem can be stated as follows. Given a universe $U = \{1, 2, \ldots, n\}$ and a family of subsets $\mathcal{S} = \{S_1, S_2, \ldots, S_m\}$ the goal is to find a function $\chi : U \to \{\pm 1\}$ that minimizes

$$\max_{1 \leq j \leq m} \left\{ \left| \sum_{i \in S_j} \chi(i) \right| \right\} . \tag{1}$$

The function χ is called a *coloring*. The *discrepancy* of $\mathcal{S}$ is the minimum of (1) over all colorings. Determining the discrepancy of a set system is a fundamental problem in combinatorics [1, 7, 14] that has a wide range of applications in computer-science [9, 5, 10, 15]. One of the most celebrated results in this area is Spencer's theorem [17] stating that any family $\mathcal{S}$ with $m = n$ has discrepancy at most $6\sqrt{n}$. More generally, if $m \geq n$, the upper bound becomes $O\left(\sqrt{n \cdot \log\left({}^{(2m)}/_{n}\right)}\right)$. This bound is tight up to constant factors for all $m \geq n$. Recently, efficient algorithms were developed [4, 3, 13] to construct colorings that match Spencer's bounds up to constant factors.

Discrepancy is also a topic of major interest in convex geometry, and many combinatorial discrepancy results have a more general geometric statement. The geometric form of Spencer's theorem is: for all $\{x_1, \ldots, x_n\} \subset [-1, 1]^n$, there exists $\chi : U \to \{\pm 1\}$ with $\left\| \sum_{i \in U} x_i \chi(i) \right\|_\infty \leq 6\sqrt{n}$. This geometric form follows from Spencer's original proof, and it is also a special case of a geometric result that was independently proven by Gluskin [12]. Gluskin's proof was simplified by Giannopoulos [11], and an algorithmic form of Giannopoulos' theorem was recently given by Rothvoss [16].

All of the previous work on Spencer's theorem, including the geometric results and the algorithmic results, are based on the idea of producing a *partial coloring*. In this approach, one first obtains a coloring of half the elements of U, then recurses on the residual family of subsets obtained by deleting all colored

[*] Department of Computer Science, University of British Columbia. Email: nickhar@cs.ubc.ca. Supported by an NSERC Discovery Grant and a Sloan Foundation Fellowship.

[†] Microsoft Research, Redmond, WA. Email: {roysch,mohits}@microsoft.com.

elements. Although the partial coloring approach suffices to obtain tight results for Spencer's theorem, there are other important discrepancy problems for which this approach does not currently (and perhaps cannot) yield tight results. A notable example is the Beck-Fiala conjecture [6], which asserts that: for every set system $\mathcal{S}$ for which every element of U is contained in at most t sets, the discrepancy of $\mathcal{S}$ is $O(\sqrt{t})$. The geometric form of the Beck-Fiala conjecture is the Komlós conjecture: for all $\{x_1, \ldots, x_n\} \subset \mathbb{R}^n$ with $\|x_i\|_2 \leq 1$, there exists $\chi : U \to \{\pm 1\}$ with $\left\|\sum_{i \in U} x_i \chi(i)\right\|_\infty \leq O(1)$.

All known results [18, 4, 13] towards these conjectures that are based on partial coloring have the drawback that they incur an extra factor of $O(\log n)$ in the discrepancy, due to the $O(\log n)$ recursive steps. The only known approach for these conjectures that avoids the extra $O(\log n)$ factor is Banaszczyk's geometric technique [2], which is not based on partial coloring, and incurs only an $O(\sqrt{\log n})$ factor, but has the drawback that it is not algorithmic. Due to the drawbacks of these previous results, it has been an open question to find new techniques that avoid partial colorings for these discrepancy problems, particularly algorithmic techniques. Such new techniques would hopefully lead to progress on the Beck-Fiala/Komlós conjectures.

1.1 Our Contribution

In this work we devise the first algorithm for Spencer's theorem that directly computes a (full) coloring, without recursively computing partial colorings. Our algorithm builds upon the techniques of Bansal [4] and Lovett and Meka [13], which we now review.

Let $c_1, \ldots, c_m$ be suitable parameters, and define

$$\mathcal{P}^{\text{disc}} \;=\; \left\{ \mathbf{x} \in \mathbb{R}^n : \left|\mathbf{1}_{S_j} \cdot \mathbf{x}\right| \leq c_j \;\; \forall 1 \leq j \leq m \right\}.$$

Bansal's breakthrough result [4] performs a random walk with Gaussian increments (i.e., discretized Brownian motion) starting at the origin. The covariance matrix of each Gaussian step comes from a feasible solution to a semidefinite program (SDP) that describes a "vector relaxation" of the discrepancy problem. If at any time the random walk approaches a face of $[-1, 1]^n$, it sticks to that face and continues walking within that face. If at any time the random walk gets very close to a discrepancy constraint, i.e., a face of $\mathcal{P}^{\text{disc}}$, that discrepancy constraint is pushed away from the origin to very carefully chosen distances, and the SDP is modified accordingly. Spencer's non-constructive theorem is used to ensure feasibility of each SDP.

Lovett and Meka [13] perform a similar random walk, except that every Gaussian step has covariance matrix equal to the identity. If at any time the random walk gets very close to a discrepancy constraint, it sticks to that face and continues walking within that face. They prove that, when the random walk stops, a constant fraction of the elements are colored, thus obtaining a partial coloring.

Both of these algorithms necessarily result in a partial coloring, not a full coloring. For Bansal's algorithm, this is because feasibility of the SDP is proven using Spencer's theorem, which only ensures existence of a "partial vector coloring", not a "full vector coloring". For the Lovett-Meka algorithm, this is because their random walk will likely stick to many discrepancy constraints before terminating. The intersection of these discrepancy constraints need not contain any point in $\{-1, 1\}^n$, and hence the walk cannot directly produce a full coloring.

Our algorithm borrows many ideas from Bansal and from Lovett-Meka, but has two key differences.

- The first difference is the way in which we distort the random walk. Like Bansal, our Gaussian steps may use different covariance matrices. Whereas Bansal's covariance matrices change only when the SDP changes (i.e., when the walk gets very close to a discrepancy constraint), our walk's covariance matrices change in every step. Thus, our walk should be viewed as a discretized diffusion process. Our covariance matrices do not directly come from vector colorings, but instead from a more geometric viewpoint. We prove the following geometric result: for any polytope $\mathcal{P}$ and point $\theta \in \mathcal{P}$ there exists an ellipsoid centered at θ and contained inside $\mathcal{P}$ such that the trace of the semi-definite matrix defining the ellipsoid is large compared to the distances of the closest faces of $\mathcal{P}$ to θ. Our proof of this geometric claim uses SDP duality and might be of independent interest. Unfortunately this geometric approach by itself is not sufficient, as one might end up close to a vertex of $\mathcal{P}^{\text{disc}} \cap [-1,1]^n$, thus getting stuck without the ability to fully color all the elements.

- The second difference is that we slowly move all discrepancy constraints away from the origin in every step. This allows the random walk to escape potential areas of $\mathcal{P}^{\text{disc}}$ in which it might get stuck. Furthermore, the distance that every constraint is moved is a deterministic function of the the time step, whereas in Bansal's approach the movement of the constraints depends on the random walk. The movement of our constraints must also be carefully chosen. On one hand the rate in which the discrepancy constraints are pushed needs to be slow enough such that one still obtains optimal discrepancy bounds. On the other hand this rate needs to be fast enough so the random walk does not get stuck too close to some discrepancy constraints. At the heart of this approach is the following simple observation: the variance of the distance between the location of the random walk and any fixed discrepancy constraint is upper bounded by the number of elements which are still uncolored. Hence, a delicate balance is needed to ensure that the rate in which the discrepancy constraints are moved is larger than the number of uncolored elements.

Let us make one final remark concerning the difference between our algorithm and previous ones. As stated above, our algorithm directly computes a full coloring without recursively computing partial colorings. On the other hand, our *analysis* partitions the algorithm's execution into several phases, which are somewhat analogous to the partial coloring steps of the Bansal and Lovett-Meka algorithms. Nevertheless, it is still accurate to say that our algorithm does not use partial colorings as the algorithm's behavior is oblivious to these phases of the analysis.

2 Preliminaries

Ellipsoids: Given a positive semi-definite $n \times n$ matrix Σ and its symmetric $n \times n$ square root matrix B (i.e. $\Sigma = B^2$) we denote the ellipsoid it defines centered at point $\theta \in \mathbb{R}^n$ by:

$$E(\Sigma, \theta) \triangleq \{\mathbf{y} = B \cdot \mathbf{u} + \theta : \mathbf{u} \in \mathbb{R}^n, ||\mathbf{u}||_2 \leq 1\} \ .$$

Additionally, we denote the Euclidean sphere centered at θ with radius r by: Ball(θ, r).

The Discrepancy Polytope: We denote by $\mathbf{1}_S \in \mathbb{R}^n$ the characteristic vector of the subset of elements $S \subseteq U$. Our algorithm conducts iterations which we index by t. For any subset $S_j \in \mathcal{S}$ and iteration t we define:

$$c_j(t) \triangleq C \cdot \sqrt{n \cdot \ln\left(\frac{2m}{n}\right)} \cdot \left(1 - 2^{-\frac{\gamma^2 \cdot t}{a}}\right) \ .$$

Here C and a are absolute constants and γ a parameter depending on n, all to be chosen later. We define the following polytopes:

$$\mathcal{P}^{\text{disc}}(t) \triangleq \left\{\mathbf{x} \in \mathbb{R}^n : \forall 1 \leq j \leq m \ \ \left|\mathbf{1}_{S_j} \cdot \mathbf{x}\right| \leq c_j(t)\right\}$$
$$\mathcal{P}(t) \triangleq \mathcal{P}^{disc}(t) \cap \{\mathbf{x} \in \mathbb{R}^n : \forall 1 \leq i \leq n \ \ |x_i| \leq 1\} \ .$$

Distances: In order to define our notion of *effective* distance of a point $\theta \in \mathcal{P}(t)$ from the jth discrepancy constraint, we need the following definition of the set of variables which are active, i.e., all variables which are not colored:

$$\mathcal{C}^{\text{act}}(\theta) \triangleq \{1 \leq i \leq n : |\theta_i| < 1 - 1/n\} \ .$$

We also denote the set of active variables in $S \subseteq U$ by: $S^{\text{act}}(\theta) \triangleq S \cap \mathcal{C}^{\text{act}}(\theta)$. For every point $\theta \in \mathcal{P}$ we denote its *distance* with respect to the jth discrepancy constraint by:

$$d_j(\theta, t) \triangleq \frac{c_j(t) - \left|\mathbf{1}_{S_j} \cdot \theta\right|}{||\mathbf{1}_{S_j^{\text{act}}(\theta)}||_2} .$$

Gaussian Distribution and Concentration: If $X \sim N(0,1)$ we denote the cumulative distribution function of the normalized gaussian by: $\Phi(t) = \Pr\left[X \leq t\right]$. We also require the following concentration result by [4].

Lemma 1 (Bansal [4])**.** *Let $X_1, \ldots, X_T$ be random variables and $Y_1, \ldots, Y_T$ be random variables where each Y_i is a function of X_i. Suppose that for all $1 \leq i \leq T$ and $x_1, \ldots, x_{i-1} \in \mathbb{R}$, $Y_i|_{X_1=x_1,\ldots,X_{i-1}=x_{i-1}} \sim N(0, \rho(x_1, \ldots, x_{i-1}))$ where $\rho(x_1, \ldots, x_{i-1}) \leq 1$. Then for any $\lambda \geq 0$:*

$$\Pr\left[|Y_1 + \ldots Y_T| \geq \lambda\sqrt{T}\right] \leq 2 \cdot e^{-\lambda^2/2} .$$

3 Algorithm

In this section we present an algorithm that fully colors all elements without resorting to partial coloring. Our algorithm conducts a random walk where in each step the direction that the algorithm moves to is determined by a suitable ellipsoid. Specifically, given a point θ and time t, a maximum trace ellipsoid that is contained inside $\mathcal{P}^{\text{disc}}(t)$ is found. In addition to being contained inside $\mathcal{P}^{\text{disc}}(t)$, we impose two additional requirements. The first is that the ellipsoid is also contained in the subspace of all variables that are still active: $\{\mathbf{x} \in \mathbb{R}^n : x_i = \theta_i \ \forall i \in U \setminus \mathcal{C}^{\text{act}}(\theta)\}$. The reason for that is that a variable i which is not active anymore is fully colored, i.e., $|x_i| \geq 1 - {}^1\!/\!{}_n$, and its value should not be changed. The second requirement is that the ellipsoid is not too large and is in fact contained inside the Euclidean unit sphere centered at θ. Such an ellipsoid can be found, for example, by solving the following semi-definite program:

$$\begin{array}{lll} SDP(\theta,t) & \max & Tr\left(\Sigma\right) \\ & s.t. & E\left(\Sigma,\theta\right) \subseteq \mathcal{P}^{\text{disc}}(t) \cap \left\{\mathbf{x} \in \mathbb{R}^n : x_i = \theta_i \ \forall i \in U \setminus \mathcal{C}^{\text{act}}(\theta)\right\} \cap \text{Ball}(\theta, 1) \end{array}$$

Note that the above semi-definite program is parameterized by a point θ and time t. Let us now provide a precise description of our algorithm and our main theorem.

Algorithm 1: $(n, \mathcal{S}, \gamma)$

```
Initialize: $\mathbf{x}(0) \leftarrow \mathbf{0}$, $t \leftarrow 0$.
while $\mathcal{C}^{act}(\mathbf{x}(t)) \neq \emptyset$ do
    Let $B(\mathbf{x}(t), t)$ be the square root of the solution for $SDP(\mathbf{x}(t), t)$.
    Choose $\mathbf{g}(t) \in \mathbb{R}^n$ s.t. $g_i(t) \sim N(0,1)$ i.i.d $\forall 1 \leq i \leq n$.
    $\mathbf{x}(t+1) \leftarrow \mathbf{x}(t) + \gamma \cdot B(\mathbf{x}(t), t) \cdot \mathbf{g}(t)$.
    if $\mathbf{x}(t+1) \notin \mathcal{P}(t+1)$ then
        Abort.
    $t \leftarrow t + 1$.
for $i = 1$ to $n$ do
    Round $x_i(t)$ to the closest integer.
Output $\mathbf{x}(t)$.
```

Theorem 1. *With a probability of at least $1/poly(n)$ Algorithm 1 terminates in polynomial time without aborting and outputs a coloring with discrepancy of* $O\left(\sqrt{n \cdot \ln\left((2m)/n\right)}\right)$.

Note: It is important to note that we can in fact compute a suitable ellipsoid without solving $SDP(\theta, t)$. It can be inferred from our proof techniques that one can directly compute a feasible solution to $SDP(\theta, t)$, whose objective value is sufficiently high that it is enough to guarantee the correctness of Theorem 1. This direct computation requires only the use of Gram-Schmidt orthogonalization, thus making Algorithm 1 considerably faster and simpler. Details are deferred to the full version of the paper.

4 Analysis

We first present the geometric core of our argument, namely that there is a suitable ellipsoid Algorithm 1 can choose in every iteration t. Then we proceed by showing that this maximum trace ellipsoid is enough to prove the correctness of the algorithm as stated in Theorem 1.

4.1 Geometric Core

At the heart of our geometric approach lies the following theorem, which proves the existence of a suitable ellipsoid. Specifically, the ellipsoid we find is centered at a given point θ and is contained inside the given polytope $\mathcal{P}$. The trace of the semi-definite matrix defining the ellipsoid is comparable to the distances of the closest faces of $\mathcal{P}$ to θ.

Theorem 2. *Let* $\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}_i \cdot \mathbf{x} = b_i \ \forall 1 \leq i \leq k,\ \mathbf{v}_j \cdot \mathbf{x} \leq c_j \ \forall 1 \leq j \leq m\}$ *such that* $\mathbf{a}_i \cdot \mathbf{v}_j = 0$ *for each* $1 \leq i \leq k$ *and* $1 \leq j \leq m$. *Let* $\theta \in \mathcal{P}$ *and* $d_j(\theta) \triangleq \frac{c_j - \mathbf{v}_j \cdot \theta}{||\mathbf{v}_j||_2}$. *Then there exists an ellipsoid* $E(\Sigma, \theta) \subseteq \mathcal{P}$ *such that* $Tr(\Sigma) \geq \min_{J:|J|=n-k}\left\{\sum_{j\in J} d_j(\theta)^2\right\}$.

Proof. Using the fact that $E(\Sigma, \theta) \subseteq \mathcal{P}$ if and only the constraints in *(Primal-SDP)* are satisfied (see Chapter 8, page 428 [8]), we obtain that it is enough to show that the objective of the following semi-definite program is more than $\min_{J:|J|=n-k}\left\{\sum_{j\in J} d_j(\theta)^2\right\}$.

$$\textit{(Primal-SDP)} \qquad \begin{array}{rll} \max & Tr(\Sigma) & \\ s.t. & \langle \mathbf{a}_i \mathbf{a}_i^T, \Sigma \rangle = 0 & \forall 1 \leq i \leq k \\ & \left(||\mathbf{v}_j||_2^2\right)^{-1} \langle \mathbf{v}_j \mathbf{v}_j^T, \Sigma \rangle \leq \left(||\mathbf{v}_j||_2^2\right)^{-1} (c_j - \mathbf{v}_j \cdot \theta)^2 = d_j(\theta)^2 & \forall 1 \leq j \leq m \\ & \Sigma \succeq 0 & \end{array}$$

Consider the dual of *(Primal-SDP)*:

$$\textit{(Dual-SDP)} \qquad \begin{array}{rll} \min & \displaystyle\sum_{j=1}^{m} \lambda_j d_j(\theta)^2 & \\ s.t. & \displaystyle\sum_{i=1}^{k} \mu_i \left(\mathbf{a}_i \mathbf{a}_i^T\right) + \sum_{j=1}^{m} \lambda_j \left(||\mathbf{v}_j||_2^2\right)^{-1} \left(\mathbf{v}_j \mathbf{v}_j^T\right) \succeq I & \\ & \lambda_j \geq 0 & \forall 1 \leq j \leq m \end{array}$$

By renaming the constraints assume, without loss of generality, that $d_1(\theta) \leq \ldots \leq d_m(\theta)$. We will show that the dual objective value for any feasible dual solution (λ, μ) is at least $\sum_{j=1}^{n-k} d_j(\theta)^2$ which will prove the theorem.

For any $0 \leq t \leq m$, consider the subspace S_t that is orthogonal to the vectors $\{\mathbf{a}_i\}_{i=1}^{k}$ and the vectors $\{\mathbf{v}_j\}_{j=1}^{t}$. Note that the dimension of S_t is at least $n-k-t$. Denote by B_t the matrix whose columns form an orthonormal basis of S_t. Taking the inner product of the dual constraint with $B_tB_t^T$, we obtain that:

$$\sum_{i=1}^{k} \mu_i \left(B_tB_t^T\right) \cdot \left(\mathbf{a}_i\mathbf{a}_i^T\right) + \sum_{j=1}^{m} \lambda_j \left(||\mathbf{v}_j||_2^2\right)^{-1} \left(B_tB_t^T\right) \cdot \left(\mathbf{v}_j\mathbf{v}_j^T\right) \geq \left(B_tB_t^T\right) \cdot I \ . \tag{2}$$

Let us focus first on the l.h.s of (2). Note that for every $1 \leq i \leq k$:

$$\left(B_tB_t^T\right) \cdot \left(\mathbf{a}_i\mathbf{a}_i^T\right) = \|B_t^T\mathbf{a}_i\|_2^2 \overset{\text{(i)}}{=} 0 \ .$$

Equality (i) is derived from the fact that the columns of B_t are orthogonal to $\{\mathbf{a}_i\}_{i=1}^{k}$. Similarly, one can show that $\left(B_tB_t^T\right) \cdot \left(\mathbf{v}_j\mathbf{v}_j^T\right) = 0$ for any $1 \leq j \leq t$. Additionally, for any $t+1 \leq j \leq m$, we have that:

$$\left(B_tB_t^T\right) \cdot \left(\mathbf{v}_j\mathbf{v}_j^t\right) = \|B_t^T\mathbf{v}_j\|_2^2 \overset{\text{(ii)}}{\leq} \|\mathbf{v}_j\|_2^2 \ .$$

Inequality (ii) follows since the columns of B_t form an orthonormal basis. Hence, we can conclude that the l.h.s of (2) is upper bounded by $\sum_{j=t+1}^{m} \lambda_j$. Let us focus now on the r.h.s of (2):

$$\left(B_tB_t^T\right) \cdot I = Tr(B_tB_t^T) = Tr(B_t^TB_t) \overset{\text{(iii)}}{\geq} n-k-t \ .$$

Inequality (iii) is derived from the fact that the columns of B_t are an orthonormal basis of dimension at least $n-k-t$. Thus, combining the upper bound on the l.h.s of (2) and lower bound on the r.h.s of (2), we obtain that for each $0 \leq t \leq m$:

$$\sum_{j=t+1}^{m} \lambda_j \geq n-k-t \ .$$

Our goal is to lower bound the value of any feasible solution for *(Dual-SDP)*. This can be done by considering the following linear program, whose variables are $\{\lambda_j\}_{j=1}^{m}$, and is a relaxation of *(Dual-SDP)*:

$$\begin{array}{lll} \min & \displaystyle\sum_{j=1}^{m} \lambda_j d_j(\theta)^2 & \\ s.t. & \displaystyle\sum_{j=t+1}^{m} \lambda_j \geq n-k-t & \forall 0 \leq t \leq m \\ & \lambda_j \geq 0 & \forall 1 \leq j \leq m \end{array}$$

Since *(Dual-SDP)* is a minimization problem, it suffices to lower bound the value of an optimal solution. Recall that $d_1(\theta) \leq \ldots \leq d_m(\theta)$ and all the λ_js are non-negative. Therefore, the optimal solution to the above linear program is $\lambda_j = 1$ for each $1 \leq j \leq n-k$ and $\lambda_j = 0$ for each $j > n-k$. Thus, we can conclude that $\sum_{j=1}^{m} \lambda_j d_j(\theta)^2 \geq \sum_{j=1}^{n-k} d_j(\theta)^2$ as claimed. □

Our analysis of Algorithm 1 actually requires the following corollary of Theorem 2. We choose the polytope $\mathcal{P}$ to correspond to the requirements we mentioned in Section 3: given a point θ, in addition to being contained in $\mathcal{P}^{\text{disc}}(t)$, the ellipsoid should also be contained in the subspace of all elements that are still active, i.e., $\{\mathbf{x} \in \mathbb{R}^n : x_i = \theta_i \; \forall i \in U \setminus \mathcal{C}^{\text{act}}(\theta)\}$, and the Euclidean unit sphere: $\text{Ball}(\theta, 1)$. The proof of the corollary appears in Appendix A.

Corollary 1. *For every $t \geq 0$ and every $\theta \in \mathcal{P}(t)$, there exists an ellipsoid*

$$E(\Sigma, \theta) \subseteq \mathcal{P}^{disc}(t) \cap \left\{\mathbf{x} \in \mathbb{R}^n : x_i = \theta_i \; \forall i \in U \setminus \mathcal{C}^{act}(\theta)\right\} \cap Ball(\theta, 1)$$

that satisfies: $Tr\,(\Sigma) \geq \min_{J \subseteq \{1,\ldots,m\} : |J| = |\mathcal{C}^{act}(\theta)|} \left\{\sum_{j \in J} \min\left\{1, d_j(\theta, t)^2\right\}\right\}$. Moreover, Σ can be computed in polynomial time.

4.2 Phases

The analysis of Algorithm 1 is done in *phases*, each comprising of several consecutive iterations of the algorithm. Denote the sequence of t values indicating the starting iteration of the ith phase by τ_i, where $\tau_i = {}^{(b \cdot i)}/_{\gamma^2}$ for some absolute constant b to be chosen later. Specifically, the ith phase of Algorithm 1, where $i = 0, 1, 2, \ldots$, corresponds to the following t values:

$$\tau_i = \frac{b \cdot i}{\gamma^2} \leq t < \frac{b \cdot (i+1)}{\gamma^2} = \tau_{i+1} \; .$$

We require the notion of success, which is made formal in the following definition.

Definition 1. *Phase i is* successful *if at its end the algorithm has not aborted and:*

$$\left|\mathcal{C}^{act}\left(\mathbf{x}\left(\tau_{i+1}\right)\right)\right| \leq 2^{-(i+1)} n \; .$$

We also require that the absolute constants C, a and b satisfy the following three conditions: $a \geq 8b$, $C \cdot \left(1 - 2^{-b/(2a)}\right) \geq \sqrt{32b}$, and $b \geq 64$. It is important to note that these conditions are *sufficient* for the correctness of Algorithm 1, but might not be necessary (they were chosen for simplicity of presentation alone). The main lemma we prove is the following.

Lemma 2. *For every i, if phase $i-1$ is successful and $\gamma \leq {}^{b}/_{n^2}$, then phase i is successful with probability of at least ${}^{1}/_{4}$.*

In order to prove Lemma 2 we start the analysis by showing that with overwhelming probability Algorithm 1 never aborts. The following lemma states that in each iteration there is an exponentially small probability of aborting. Its proof appears in the Appendix B.

Lemma 3. *For every iteration $t \geq 0$: $\Pr\left[\mathbf{x}(t+1) \notin \mathcal{P}(t+1) | \mathbf{x}(t) \in \mathcal{P}(t)\right] \leq (2n+1) \cdot \left(1 - \Phi\left((\gamma \cdot n)^{-1}\right)\right)$. Moreover, if $\tau_i \leq t < \tau_{i+1}$, then the lemma holds also when conditioning that phase $i-1$ is successful.*

Consider the following random subset:

$$A(t) \triangleq \left\{j : d_j\left(\mathbf{x}(t), t\right) \leq 1\right\} \; .$$

The random subset $A(t)$ consists of all discrepancy constraints j which are *bad*, as such constraints are close to the location of Algorithm 1 at time t, i.e., $\mathbf{x}(t)$. Let us lower bound the expected trace of the ellipsoid for every iteration t using $A(t)$.

Lemma 4. *For every $t \geq 0$: $\mathbb{E}\left[Tr\left(B(\mathbf{x}(t),t)^2\right)\right] \geq \mathbb{E}\left[|\mathcal{C}^{act}(\mathbf{x}(t))|\right] - \mathbb{E}\left[|A(t)|\right]$. Moreover, if $\tau_i \leq t < \tau_{i+1}$, then the lemma holds also when conditioning that phase $i-1$ is successful.*

Proof.

$$\mathbb{E}\left[Tr\left(B\left(\mathbf{x}(t),t\right)^2\right)\right] \overset{\text{(i)}}{\geq} \mathbb{E}\left[\min_{J\subseteq\{1,\ldots,m\}:|J|=|\mathcal{C}^{\text{act}}(\mathbf{x}(t))|}\left\{\sum_{j\in J}\min\left\{1,d_j(\mathbf{x}(t),t)^2\right\}\right\}\right] \tag{3}$$

Inequality (i) is from Corollary 1. Let $J^*(t)$ be the random subset achieving the min value in the r.h.s of (3). Then,

$$\begin{aligned}\mathbb{E}\left[\min_{J\subseteq\{1,\ldots,m\}:|J|=|\mathcal{C}^{\text{act}}(\mathbf{x}(t))|}\left\{\sum_{j\in J}\min\left\{1,d_j(\mathbf{x}(t),t)^2\right\}\right\}\right] &= \mathbb{E}\left[\sum_{j\in J^*(t)}\min\left\{1,d_j(\mathbf{x}(t),t)^2\right\}\right]\\ &\geq \mathbb{E}\left[\sum_{j\in J^*(t)\setminus A(t)}\min\left\{1,d_j(\mathbf{x}(t),t)^2\right\}\right]\\ &\overset{\text{(ii)}}{\geq} \mathbb{E}\left[|J^*(t)\setminus A(t)|\right]\\ &\geq \mathbb{E}\left[|J^*(t)|\right] - \mathbb{E}\left[|A(t)|\right]\\ &\overset{\text{(iii)}}{=} \mathbb{E}\left[\left|\mathcal{C}^{\text{act}}(\mathbf{x}(t))\right|\right] - \mathbb{E}\left[|A(t)|\right]\end{aligned}$$

Inequality (ii) follows from the fact that if $j \notin A(t)$ then $d_j(\mathbf{x}(t),t) \geq 1$ (by the definition of $A(t)$). Equality (iii) is true since $|J^*(t)| = |\mathcal{C}^{\text{act}}(\mathbf{x}(t))|$ by the definition of $J^*(t)$. Note that the exact same proof holds also when conditioning that phase $i-1$ is successful. □

It is clear that one wishes that $|A(t)|$ be as small as possible. The following lemma states that for many of the iterations of phase i, the expected size of $A(t)$ is indeed small enough. Its proof appears in Appendix C.

Lemma 5. *For every $i \geq 0$ and iteration t, where $\tau_i + {}^{b}\!/\!{}_{(2\cdot\gamma^2)} \leq t < \tau_{i+1}$, if phase $i-1$ is successful then:*

$$\mathbb{E}\left[|A(t)|\right] \leq 2\cdot 2^{-\frac{1}{8b}C^2\left(1-2^{-b/(2a)}\right)^2}\cdot\left(2^{-i}\cdot n\right)\ .$$

We are now ready to prove the main Lemma.

Proof of Lemma 2. For simplicity of presentation, we omit in this proof the notations indicating that all events and probabilities are conditioned on the event that phase $i-1$ is successful. First, let us bound the probability that Algorithm 1 does not abort during phase i. Using union bound over all b/γ^2 iterations in phase i and applying Lemma 3, one can conclude that the probability of aborting during phase i is at most:

$$\frac{b}{\gamma^2}\cdot(2n+1)\cdot\left(1-\Phi\left(\frac{1}{\gamma\cdot n}\right)\right)\ .$$

Since $\gamma \leq b/n^2$, the application of standard gaussian tail bounds suffices to obtain a total aborting probability of at most: $e^{-n^2} \leq 1/4$.

Second, let us prove that at the end of phase i: $\Pr\left[|\mathcal{C}^{\text{act}}(\mathbf{x}(\tau_{i+1}))| > 2^{-(i+1)}\cdot n\right] < 1/2$. Note that this concludes the proof since one can apply a union bound over the latter event and the event that Algorithm 1 did not abort during phase i, yielding a failure probability of at most $3/4$.

We now examine two cases, depending on the value of $\mathbb{E}\left[|\mathcal{C}^{\text{act}}(\mathbf{x}(\tau_{i+1}))|\right]$. If we are in the case that $\mathbb{E}\left[|\mathcal{C}^{\text{act}}(\mathbf{x}(\tau_{i+1}))|\right] \leq {}^{1}\!/_{2} \cdot \left(2^{-(i+1)} \cdot n\right)$ then Markov's inequality suffices. The reason for latter is that we get: $\Pr\left[|\mathcal{C}^{\text{act}}(\mathbf{x}(\tau_{i+1}))| > 2^{-(i+1)} \cdot n\right] < {}^{1}\!/_{2}$. Otherwise, let us assume that we are in the case where $\mathbb{E}\left[|\mathcal{C}^{\text{act}}(\mathbf{x}(\tau_{i+1}))|\right] > {}^{1}\!/_{2} \cdot \left(2^{-(i+1)} \cdot n\right)$. Note that $|\mathcal{C}^{\text{act}}(\mathbf{x}(t))|$ is a monotone non-decreasing function in t, and therefore for every iteration t in phase i we have that: $\mathbb{E}\left[|\mathcal{C}^{\text{act}}(\mathbf{x}(t))|\right] > {}^{1}\!/_{2} \cdot \left(2^{-(i+1)} \cdot n\right)$. Let us examine now the expected change in the Euclidean norm of $\mathbf{x}(t)$ during phase i.

$$\begin{aligned}
\mathbb{E}\left[||\mathbf{x}(\tau_{i+1}) - \mathbf{x}(\tau_i)||_2^2\right] &\overset{\text{(i)}}{=} \gamma^2 \cdot \mathbb{E}\left[\left|\left|\sum_{t=\tau_i}^{\tau_{i+1}-1} B(\mathbf{x}(t),t) \cdot \mathbf{g}(t)\right|\right|_2^2\right] \\
&\overset{\text{(ii)}}{=} \gamma^2 \cdot \mathbb{E}\left[\sum_{t=\tau_i}^{\tau_{i+1}-1} Tr\left(B(\mathbf{x}(t),t)^2\right)\right] \\
&\geq \gamma^2 \sum_{t=\tau_i + {}^{b}\!/(2\cdot\gamma^2)}^{\tau_{i+1}-1} \mathbb{E}\left[Tr\left(B(\mathbf{x}(t),t)^2\right)\right] \\
&\overset{\text{(iii)}}{\geq} \sum_{t=\tau_i + {}^{b}\!/(2\cdot\gamma^2)}^{\tau_{i+1}-1} \left(\mathbb{E}\left[\left|\mathcal{C}^{\text{act}}(\mathbf{x}(t))\right|\right] - \mathbb{E}\left[|A(t)|\right]\right) \\
&\overset{\text{(iv)}}{>} \gamma^2 \sum_{t=\tau_i + {}^{b}\!/(2\cdot\gamma^2)}^{\tau_{i+1}-1} \left(\frac{1}{4} - 2 \cdot 2^{-\frac{1}{8}\frac{C^2\left(1-2^{-\frac{b}{2a}}\right)^2}{b}}\right) \cdot \left(2^{-i} \cdot n\right) \\
&= \frac{b}{2} \cdot \left(\frac{1}{4} - 2 \cdot 2^{-\frac{1}{8}\frac{C^2\left(1-2^{-\frac{b}{2a}}\right)^2}{b}}\right) \cdot \left(2^{-i} \cdot n\right) \\
&\overset{\text{(v)}}{\geq} 4 \cdot \left(2^{-i} \cdot n\right)
\end{aligned}$$

Equality (i) is by the definition of Algorithm 1. Note that equality (ii) follows from the fact that all the $\mathbf{g}(t)$s are independent random standard gaussian vectors. Specifically, it is easy to show that for any matrix A, vector $\mathbf{z}$ and random gaussian vector $\mathbf{g}$:

$$\mathbb{E}\left[||\mathbf{z} + A\mathbf{g}||_2^2\right] = ||\mathbf{z}||_2^2 + Tr\left(A \cdot A^T\right) \ .$$

Lemma 4 yields inequality (iii). Inequality (iv) is derived from Lemma 5 and the current case assumption that $\mathbb{E}\left[|\mathcal{C}^{\text{act}}(\mathbf{x}(\tau_{i+1}))|\right] > {}^{1}\!/_{2} \cdot \left(2^{-(i+1)} \cdot n\right)$. Finally, inequality (v) follows from the conditions on the constants.

Note that $||\mathbf{x}(\tau_{i+1}) - \mathbf{x}(\tau_i)||_2^2$ can *never* exceed $4 \cdot |\mathcal{C}^{\text{act}}(\mathbf{x}(\tau_i))|$. This follows from the following two. First, Corollary 1, specifically that $E\left(B(\mathbf{x}(t),t)^2, \mathbf{x}(t)\right) \subseteq \{\mathbf{z} \in \mathbb{R}^n : z_i = x_i(t)\ \forall i \in U \setminus \mathcal{C}^{\text{act}}(\mathbf{x}(t))\}$, implies that all variables i that do not belong to $\mathcal{C}^{\text{act}}(\mathbf{x}(t))$ never change their value from iteration t onwards. Second, phase $i-1$ was successful, namely that $|\mathcal{C}^{\text{act}}(\mathbf{x}(\tau_i))| \leq 2^{-i} \cdot n$. Since we proved that $\mathbb{E}\left[||\mathbf{x}(\tau_{i+1}) - \mathbf{x}(\tau_i)||_2^2\right] > 4 \cdot 2^{-i} \cdot n$ we got a contradiction. Thus, it cannot happen that $\mathbb{E}\left[|\mathcal{C}^{\text{act}}(\mathbf{x}(\tau_{i+1}))|\right] > {}^{1}\!/_{2} \cdot \left(2^{-(i+1)} \cdot n\right)$ and we conclude the proof. □

We are now ready to prove the main result, Theorem 1.

Proof of Theorem 1. Let us calculate the probability that all phases Algorithm 1 makes are successful. First, denote by N the number of phases Algorithm 1 makes in case all phases are successful, and by T the total number of iterations in case all phases are successful. Definition 1 implies that $N = O(\log n)$, since

$|\mathcal{C}^{\text{act}}(\mathbf{x}(\tau_i))| \leq 2^{-i} \cdot n$. Second, Lemma 2 provides, conditioned on the success of the previous phase and by choosing $\gamma = b/n^2$, that the success probability of the current phase is at least $1/4$. Therefore, we can conclude that the probability that all phases of Algorithm 1 are successful is at least $(1/4)^N = 1/poly(n)$. Moreover, when all phases are successful the following two are implied:

1. Algorithm 1 never aborts (by Definition 1).
2. $\mathbf{x}(T) \in \mathcal{P}(T)$. Since $\mathcal{P}(T) \subseteq \mathcal{P}(\infty)$, we can conclude that for every $1 \leq j \leq m$:

$$\left|\mathbf{1}_{S_j} \cdot \mathbf{x}(T)\right| \leq C \cdot \sqrt{n \cdot \ln\left(\frac{2m}{n}\right)}\,.$$

Note that the rounding step of Algorithm 1 (step 10) might incur only an *additive* loss of 1 in the discrepancy. This concludes the proof. □

Choosing Parameters: It suffices to choose: $C = 2^7$, $a = 2^9$ and $b = 2^6$ in order to satisfy all the required conditions.

5 Conclusion and Open Problems

We devise the first algorithm for Spencer's theorem that directly computes a coloring, without recursively computing partial colorings. This naturally leads to several interesting questions.

- Can this technique for directly producing full colorings be used to make progress on the Beck-Fiala [6] conjecture? As a first step, can this approach give an algorithmic form of Banaszczyk's result [2]?
- Although our algorithm does not directly produce partial colorings, our analysis involves multiple phases, which are somewhat analogous to partial colorings. Can the analysis be refined to avoid the notion of phases?

References

[1] N. Alon and J. Spencer. *The Probabilistic Method*. John Wiley, 2000.

[2] Wojciech Banaszczyk. Balancing vectors and Gaussian measures of n-dimensional convex bodies. *Random Structures & Algorithms*, 12(4):351–360, 1998.

[3] N. Bansal and J. Spencer. Deterministic discrepancy minimization. *Algorithmica*, 67(4):451–471, 2013.

[4] Nikhil Bansal. Constructive algorithms for discrepancy minimization. In *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 3–10. IEEE, 2010.

[5] Nikhil Bansal, Moses Charikar, Ravishankar Krishnaswamy, and Shi Li. Better algorithms and hardness for broadcast scheduling via a discrepancy approach. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2014.

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

[7] József Beck and Vera T. Sós. Discrepancy theory. In R. Graham and M. Grötschel and L. Lovász, editor, *Handbook of Combinatorics*, pages 1405–1446. Elsevier Science B.V., 1995.

[8] Stephen P. Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

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

[10] Friedrich Eisenbrand, Dömötör Pálvölgyi, and Thomas Rothvoß. Bin packing via discrepancy of permutations. *ACM Transactions on Algorithms (TALG)*, 9(3):24, 2013.

[11] Apostolos Giannopoulos. On some vector balancing problems. *Studia Mathematica*, 122(3):225–234, 1997.

[12] Efim Davydovich Gluskin. Extremal properties of orthogonal parallelepipeds and their applications to the geometry of Banach spaces. *Mathematics of the USSR-Sbornik*, 64(1):85, 1989.

[13] Shachar Lovett and Raghu Meka. Constructive discrepancy minimization by walking on the edges. In *Proceedings of the 53rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 61–67. IEEE, 2012.

[14] Jiří Matoušek. *Geometric discrepancy: An illustrated guide*, volume 18. Springer, 1999.

[15] Thomas Rothvoß. Approximating bin packing within $O(\log \mathrm{OPT} \cdot \log\log \mathrm{OPT})$ bins. In *Proceedings of the 54th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 20–29. IEEE, 2013.

[16] Thomas Rothvoss. Constructive discrepancy minimization for convex sets. *arXiv preprint arXiv:1404.0339*, 2014.

[17] Joel Spencer. Six standard deviations suffice. *Transactions of the American Mathematical Society*, 289(2):679–706, 1985.

[18] Aravind Srinivasan. Improving the discrepancy bound for sparse matrices: better approximations for sparse lattice approximation problems. In *Proceedings of the 8th annual ACM-SIAM Symposium on Discrete Algorithms*, pages 692–701. Society for Industrial and Applied Mathematics, 1997.

Appendix

A Proof of Corollary 1

Proof. The proof follows the same outline but differs slightly from the proof of Theorem 2. First observe that the equality constraint for $\mathcal{P}^{\text{disc}}(t) \cap \{\mathbf{x} \in \mathbb{R}^n : x_i = \theta_i \ \forall i \in U \setminus \mathcal{C}^{\text{act}}(\theta)\}$ are $\mathbf{e}_i \cdot \mathbf{x} = \theta_i$ for each $i \in U \setminus \mathcal{C}^{\text{act}}(\theta)$. Now, we write the inequalities to make sure the constraints are orthogonal to the equality constraint. Observe that $\mathbf{1}_{S_j} \cdot \mathbf{x} \leq c_j(t)$ is equivalent to the constraint $\mathbf{1}_{S_j^{\text{act}}(\theta)} \cdot \mathbf{x} \leq c_j(t) - \mathbf{1}_{S \setminus S_j^{\text{act}}(\theta)} \cdot \theta$. Similarly, we have $-\mathbf{1}_{S_j} \cdot \mathbf{x} \leq c_j(t)$ is equivalent to the constraint $-\mathbf{1}_{S_j^{\text{act}}(\theta)} \cdot \mathbf{x} \leq c_j(t) + \mathbf{1}_{S \setminus S_j^{\text{act}}(\theta)} \cdot \theta$.

Thus, as in Theorem 2, we obtain that it is enough to show that the objective of the following semidefinite program is more than $\min_{J \subseteq \{1,\ldots,m\} : |J| = |\mathcal{C}^{\text{act}}(\theta)|} \left\{ \sum_{j \in J} \min\left\{1, d_j(\theta,t)^2\right\} \right\}$. Here the constraint $\Sigma \preceq I$ follows from the fact $E(\Sigma, \theta) \subseteq Ball(\theta, 1)$.

$$
\begin{array}{llll}
\textit{(Primal-SDP2)} & \max & Tr(\Sigma) & \\
& s.t. & \langle \mathbf{e}_i \mathbf{e}_i^T, \Sigma \rangle = 0 & \forall i \in U \setminus \mathcal{C}^{\text{act}}(\theta) \\
& & \dfrac{\langle \mathbf{1}_{S_j^{\text{act}}(\theta)} \mathbf{1}_{S_j^{\text{act}}(\theta)}^T, \Sigma \rangle}{||\mathbf{1}_{S_j^{\text{act}}(\theta)}||_2^2} \leq \dfrac{c_j(t) - \mathbf{1}_{S \setminus S_j^{\text{act}}(\theta)} \cdot \theta - \mathbf{1}_{S^{\text{act}_j(\theta)}} \cdot \theta}{||\mathbf{1}_{S^{\text{act}_j(\theta)}}||_2^2} & \forall 1 \leq j \leq m \\
& & \dfrac{\langle \mathbf{1}_{S^{\text{act}_j(\theta)}} \mathbf{1}_{S^{\text{act}_j(\theta)}}^T, \Sigma \rangle}{||\mathbf{1}_{S^{\text{act}_j(\theta)}}||_2^2} \leq \dfrac{c_j(t) + \mathbf{1}_{S \setminus S_j^{\text{act}}(\theta)} \cdot \theta + \mathbf{1}_{S_j^{\text{act}}(\theta)} \cdot \theta}{||\mathbf{1}_{S_j^{\text{act}}(\theta)}||_2^2} & \forall 1 \leq j \leq m \\
& & \Sigma \preceq I & \\
& & \Sigma \succeq 0 &
\end{array}
$$

Simplifying, we obtain that the SDP is equivalent to

$$
\begin{array}{llll}
\textit{(Primal-SDP2)} & \max & Tr(\Sigma) & \\
& s.t. & \langle \mathbf{e}_i \mathbf{e}_i^T, \Sigma \rangle = 0 & \forall i \in U \setminus \mathcal{C}^{\text{act}}(\theta) \\
& & \left(||\mathbf{1}_{S_j^{\text{act}}(\theta)}||_2^2 \right)^{-1} \langle \mathbf{1}_{S_j^{\text{act}}(\theta)} \mathbf{1}_{S_j^{\text{act}}(\theta)}^T, \Sigma \rangle \leq d_j(\theta, t)^2 & \forall 1 \leq j \leq m \\
& & \Sigma \preceq I & \\
& & \Sigma \succeq 0 &
\end{array}
$$

Consider the dual of *Primal-SDP2*:

$$
\begin{array}{llll}
\textit{(Dual-SDP2)} & \min & \displaystyle\sum_{j=1}^{m} \lambda_j d_j(\theta)^2 + Tr(V) & \\
& s.t. & \displaystyle\sum_{i \in U \setminus \mathcal{C}^{\text{act}}(\theta)} \mu_i \left(\mathbf{e}_i \mathbf{e}_i^T \right) + \sum_{j=1}^{m} \lambda_j \frac{1}{|S_j^{\text{act}}(\theta)|} \left(\mathbf{1}_{S_j^{\text{act}}(\theta)} \mathbf{1}_{S_j^{\text{act}}(\theta)}^T \right) \succeq I - V & \\
& & \lambda_j \geq 0 & \forall 1 \leq j \leq m \\
& & V \succeq 0 &
\end{array}
$$

By renaming the constraints assume, without loss of generality, that $d_1(\theta, t) \leq \ldots \leq d_m(\theta, t)$ and r be the maximum value such that $d_r(\theta, t) \leq 1$. Assume that $r \leq |\mathcal{C}^{\text{act}}(\theta)|$, else the proof is identical to proof of Theorem 2. We will show that the dual objective value for any feasible dual solution (λ, μ, V) is at least $\sum_{j=1}^{r} d_j(\theta, t)^2 + |\mathcal{C}^{\text{act}}(\theta)| - r$ which will prove the theorem.

For any $0 \leq t \leq |\mathcal{C}^{\text{act}}(\theta)|$, consider the subspace S_t that is orthogonal to the vectors $\{\mathbf{a}_i\}_{i \in U \setminus \mathcal{C}^{\text{act}}(\theta)}$ and the vectors $\{\mathbf{v}_j\}_{j=1}^{t}$. Note that the dimension of S_t is at least $|\mathcal{C}^{\text{act}}(\theta)| - t$. Denote by B_t the matrix whose columns form an orthonormal basis of S_t. Taking the inner product of the dual constraint with $B_t B_t^T$, we obtain that:

$$
B_t B_t^T \cdot \left(\sum_{i \in U \setminus \mathcal{C}^{\text{act}}(\theta)} \mu_i \left(\mathbf{e}_i \mathbf{e}_i^T \right) + \sum_{j=1}^{m} \lambda_j \frac{1}{|S_j^{\text{act}}(\theta)} \left(\mathbf{1}_{S_j^{\text{act}}(\theta)} \mathbf{1}_{S_j^{\text{act}}(\theta)}^T \right) \right) \succeq B_t B_t^T \cdot I - B_t B_t^T \cdot V \,. \quad (4)
$$

As in proof of Theorem 2, we conclude that l.h.s is upper bounded by $\sum_{j=t+1}^{m} \lambda_j$. Let us focus now on the r.h.s of (4):

$$\left(B_t B_t^T\right) \cdot I - B_t B_t^T \cdot V \geq Tr(B_t B_t^T) - Tr(V) \geq |\mathcal{C}^{\text{act}}(\theta)| - t - Tr(V) \ .$$

where we use the fact that $B_t B_t^T \cdot V \leq I \cdot V = Tr(V)$. Thus, we obtain that for each $1 \leq t \leq m$:

$$\sum_{j=t+1}^{m} \lambda_j \geq |\mathcal{C}^{\text{act}}(\theta)| - t - Tr(V) \ .$$

Consider the following linear program with variables λ_j for each $1 \leq j \leq m$ and variable $Tr(V)$ which is a relaxation of (Primal-SDP). Thus it enough to lower bound the value of optimum solution to this linear program.

$$\begin{array}{lll}
\min & \displaystyle\sum_{j=1}^{m} \lambda_j d_j(\theta)^2 + Tr(V) & \\
s.t. & \displaystyle\sum_{j=t+1}^{m} \lambda_j \geq |\mathcal{C}^{\text{act}}(\theta)| - t - Tr(V) \ . & 0 \leq t \leq |\mathcal{C}^{\text{act}}(\theta)| \\
& \lambda_j \geq 0 & \forall 1 \leq j \leq m
\end{array}$$

Since $d_1(\theta) \leq \ldots \leq d_m(\theta)$ and all the λ_js are non-negative and $d_j(\theta) > 1$ if $j > r$, the optimal solution to the above linear program is $\lambda_j = 1$ for each $1 \leq j \leq r$ and $\lambda_j = 0$ for each $j > r$ and $Tr(V) = |\mathcal{C}^{\text{act}}(\theta)| - r$. Thus we can conclude that $\sum_{j=1}^{m} \lambda_j d_j(\theta)^2 + Tr(V) \geq \sum_{j=1}^{r} d_j(\theta)^2 + |\mathcal{C}^{\text{act}}(\theta)| - r \geq \sum_{j=1}^{|\mathcal{C}^{\text{act}}(\theta)|} \min\{1, d_j(\theta)^2\}$ as claimed. □

B Proof of Lemma 3

Proof.

$$\begin{aligned}
\Pr\left[\mathbf{x}(t+1) \notin \mathcal{P}(t) | \mathbf{x}(t) \in \mathcal{P}(t)\right] &\overset{\text{(i)}}{=} \Pr\left[\mathbf{x}(t) + \gamma \cdot B(\mathbf{x}(t), t) \cdot \mathbf{g}(t) \notin \mathcal{P}(t) | \mathbf{x}(t) \in \mathcal{P}(t)\right] \\
&\overset{\text{(ii)}}{\leq} \Pr\left[\mathbf{x}(t) + \gamma \cdot B(\mathbf{x}(t), t) \cdot \mathbf{g}(t) \notin E\left(B(\mathbf{x}(t), t)^2, \mathbf{x}(t)\right) | \mathbf{x}(t) \in \mathcal{P}(t)\right] + \\
&\quad \Pr\left[||\mathbf{x}(t) + \gamma \cdot B(\mathbf{x}(t), t) \cdot \mathbf{g}(t)||_\infty > 1 | \mathbf{x}(t) \in \mathcal{P}(t)\right] \\
&= \Pr\left[\gamma \cdot B(\mathbf{x}(t), t) \cdot \mathbf{g}(t) \notin E\left(B(\mathbf{x}(t), t)^2, 0\right) | \mathbf{x}(t) \in \mathcal{P}(t)\right] + \\
&\quad \Pr\left[||\mathbf{x}(t) + \gamma \cdot B(\mathbf{x}(t), t) \cdot \mathbf{g}(t)||_\infty > 1 | \mathbf{x}(t) \in \mathcal{P}(t)\right] \\
&= \Pr\left[||\mathbf{g}(t)||_2^2 > 1/\gamma^2\right] + \\
&\quad \Pr\left[\exists i \in \mathcal{C}^{\text{act}}(\mathbf{x}(t)) \ s.t. \ |x_i(t) + \gamma \cdot (B(\mathbf{x}(t), t) \cdot \mathbf{g}(t))_i| > 1 | \mathbf{x}(t) \in \mathcal{P}(t)\right] \\
&\overset{\text{(iii)}}{\leq} \Pr\left[||\mathbf{g}(t)||_2^2 > 1/\gamma^2\right] + \\
&\quad \Pr\left[\exists i \in \mathcal{C}^{\text{act}}(\mathbf{x}(t)) \ s.t. \ \gamma \cdot |(B(\mathbf{x}(t), t) \cdot \mathbf{g}(t))_i| > 1/n | \mathbf{x}(t) \in \mathcal{P}(t)\right] \\
&\overset{\text{(iv)}}{\leq} (2n+1) \cdot \left(1 - \Phi\left(\frac{1}{\gamma \cdot n}\right)\right)
\end{aligned}$$

Equality (i) is from the definition of Algorithm 1. Inequality (ii) is derived from the definition of $\mathcal{P}(t)$ and Corollary 1 since: $E\left(B(\mathbf{x}(t), t)^2, \mathbf{x}(t)\right) \subseteq \mathcal{P}^{\text{disc}}(t)$. Inequality (iii) is derived from the fact that $i \in \mathcal{C}^{\text{act}}(\mathbf{x}(t))$ implies that $|x_i(t)| < 1 - 1/n$. Finally, inequality (iv) is true since $|\mathcal{C}^{\text{act}}(\mathbf{x}(t))| \leq n$ and since Corollary 1 implies that $E\left(B(\mathbf{x}(t), t)^2, \mathbf{x}(t)\right) \subseteq \text{Ball}(\mathbf{x}(t), 1)$. The lemma now follows since $\mathcal{P}(t) \subseteq \mathcal{P}(t+1)$. Note that the exact same proof holds also when conditioning that phase $i - 1$ is successful. □

C Proof of Lemma 5

Proof. For simplicity, let us denote $t = \tau_i + s$ where ${}^{b}\!/\!{}_{(2\gamma^2)} \le s < {}^{b}\!/\!{}_{\gamma^2}$.

$$\begin{aligned}
\Pr\left[d_j\left(\mathbf{x}(t), t\right) \le 1\right] &\overset{\text{(i)}}{=} \Pr\left[\frac{c_j(t) - \left|\mathbf{1}_{S_j} \cdot \mathbf{x}(t)\right|}{||\mathbf{1}_{S_j^{\text{act}}\mathbf{x}(t)}||_2} \le 1\right] \\
&= \Pr\left[c_j(t) - ||\mathbf{1}_{S_j^{\text{act}}(\mathbf{x}(t))}||_2 \le \left|\mathbf{1}_{S_j} \cdot \left(\mathbf{x}(t) - \mathbf{x}(\tau_i)\right) + \mathbf{1}_{S_j} \cdot \mathbf{x}(\tau_i)\right|\right] \\
&\overset{\text{(ii)}}{\le} \Pr\left[\left|\mathbf{1}_{S_j} \cdot \left(\mathbf{x}(t) - \mathbf{x}(\tau_i)\right)\right| \ge c_j(t) - c_j(\tau_i) - ||\mathbf{1}_{S_j^{\text{act}}(\mathbf{x}(t))}||_2\right] . \qquad (5)
\end{aligned}$$

Equality (i) is by the definition of $d_j(\mathbf{x}(t), t)$. Inequality (ii) follows from the fact that phase $i-1$ is successful, and in particular Algorithm 1 did not abort (implying that $|\mathbf{x}(\tau_i)| \le c_j(\tau_i)$). Let us now lower bound the r.h.s of the event in (5). First,

$$\begin{aligned}
c_j(t) - c_j(\tau_i) &= C \cdot \sqrt{n \cdot \ln\left({}^{(2m)}\!/\!{}_{n}\right)} \cdot 2^{-\frac{b}{a} \cdot i} \cdot \left(1 - 2^{-\frac{s \cdot \gamma^2}{a}}\right) \\
&\overset{\text{(iii)}}{\ge} C \cdot \sqrt{n \cdot \ln\left({}^{(2m)}\!/\!{}_{n}\right)} \cdot 2^{-\frac{b}{a} \cdot i} \cdot \left(1 - 2^{-\frac{b}{2a}}\right) . \qquad (6)
\end{aligned}$$

Inequality (iii) is derived from the fact that we consider only iterations t in the second half of phase i, i.e., ${}^{b}\!/\!{}_{(2\gamma^2)} \le s$. Second,

$$||\mathbf{1}_{S_j^{\text{act}}(\mathbf{x}(t))}||_2 \le \sqrt{\left|S_j^{\text{act}}(\mathbf{x}(t))\right|} \le \sqrt{|\mathcal{C}^{\text{act}}(\mathbf{x}(t))|} \overset{\text{(iv)}}{\le} \sqrt{|\mathcal{C}^{\text{act}}(\mathbf{x}(\tau_i))|} \overset{\text{(v)}}{\le} 2^{-{}^{i}\!/\!{}_{2}} \cdot \sqrt{n} \,. \qquad (7)$$

Note that $|\mathcal{C}^{\text{act}}(\mathbf{x}(t))|$ is a monotone non-increasing function in t, and thus inequality (iv) follows. Since we assumed that phase $i-1$ is successful, i.e., $|\mathcal{C}^{\text{act}}(\mathbf{x}(\tau_i))| \le 2^{-i} \cdot n$ (see Definition 1), inequality (v) is true. Plugging (6) and (7) into the r.h.s of the event in (5) yields:

$$\begin{aligned}
c_j(t) - c_j(\tau_i) - ||\mathbf{1}_{S_j^{\text{act}}(\mathbf{x}(t))}||_2 &\ge C \cdot \sqrt{n \cdot \ln\left({}^{(2m)}\!/\!{}_{n}\right)} \cdot 2^{-\frac{b}{a} \cdot i} \cdot \left(1 - 2^{-\frac{b}{2a}}\right) - 2^{-{}^{i}\!/\!{}_{2}} \cdot \sqrt{n} \\
&\overset{\text{(vi)}}{\ge} \sqrt{n} \cdot 2^{-\frac{b}{a} \cdot i} \cdot \left(C \cdot \sqrt{\ln\left({}^{(2m)}\!/\!{}_{n}\right)} \cdot \left(1 - 2^{-\frac{b}{2a}}\right) - 1\right) \\
&\overset{\text{(vii)}}{\ge} \sqrt{n} \cdot 2^{-\frac{b}{a} \cdot i} \cdot \frac{C}{2} \cdot \sqrt{\ln\left({}^{(2m)}\!/\!{}_{n}\right)} \cdot \left(1 - 2^{-\frac{b}{2a}}\right) \qquad (8)
\end{aligned}$$

Inequalities (vi) and (vii) are both derived from the conditions we imposed on the constants. Specifically, inequality (vi) follows from the condition that $a \ge 8b$, whereas inequality (vii) follows from the condition that $C \cdot \left(1 - 2^{-{}^{b}\!/\!{}_{(2a)}}\right) \ge \sqrt{32b} \ge 4$ (since $b \ge 64$). Next, consider the l.h.s of the event in (5). Note that by the definition of Algorithm 1:

$$\mathbf{x}(t) - \mathbf{x}(\tau_i) = \gamma \sum_{r=\tau_i}^{\tau_i + s} B(\mathbf{x}(r), r) \cdot \mathbf{g}(r) \,.$$

It can be verified that given the random choices of the algorithm in the first $r-1$ iterations, for any $\tau_i \le r \le \tau_i + s$, the random variable $\gamma \mathbf{1}_{S_j} \cdot (B(\mathbf{x}(r), r) \cdot \mathbf{g}(r))$ is a normal random variable with mean 0 and standard deviation σ, where:

$$\sigma^2 = \gamma^2 \mathbf{1}_{S_j}^T B(\mathbf{x}(r), r)^2 \mathbf{1}_{S_j} \overset{\text{(viii)}}{=} \gamma^2 \mathbf{1}_{S_j^{\text{act}}(\mathbf{x}(r))}^T B(\mathbf{x}(r), r)^2 \mathbf{1}_{S_j^{\text{act}}(\mathbf{x}(r))} \overset{\text{(ix)}}{\le} \gamma^2 \left|S_j^{\text{act}}(\mathbf{x}(r))\right| \overset{\text{(x)}}{\le} \gamma^2 \cdot 2^{-i} n \,. \qquad (9)$$

Equality (viii) is derived from the property that

$$E(B(\mathbf{x}(r), r)^2, \mathbf{x}(r)) \subseteq \left\{\mathbf{z} \in \mathbb{R}^n : z_i = x_i(r)\ \forall i \in U \setminus \mathcal{C}^{\text{act}}(\mathbf{x}(r))\right\}\ ,$$

as guaranteed by Corollary 1. Note that inequality (ix) follows again from Corollary 1, specifically that $B(\mathbf{x}(r), r)^2 \preceq I$ (or equivalently that $E(B(\mathbf{x}(r), r)^2, \mathbf{x}(r)) \subseteq \text{Ball}(\mathbf{x}(r), 1)$). Finally, recall that $|\mathcal{C}^{\text{act}}(\mathbf{x}(r))|$ is a monotone non-increasing function in r. Therefore, inequality (x) is true since we assume that phase $i-1$ was successful, i.e., $|\mathcal{C}^{\text{act}}(\mathbf{x}(\tau_i))| \leq 2^{-i} \cdot n$ (see Definition 1). Applying the concentration bound of [4] for (5) results in:

$$\begin{aligned}
\Pr\left[d_j(\mathbf{x}(t), t) \leq 1\right] &\leq \Pr\left[\left|\mathbf{1}_{S_j} \cdot (\mathbf{x}(t) - \mathbf{x}(\tau_i))\right| \geq c_j(t) - c_j(\tau_i) - ||\mathbf{1}_{S_j^{\text{act}}(\mathbf{x}(t))}||_2\right] \\
&\overset{\text{(xi)}}{\leq} 2 \cdot exp\left(-\frac{1}{2} \cdot \frac{n \cdot \frac{1}{4} \cdot C^2 \cdot 2^{-\frac{2b}{a} \cdot i} \cdot \left(1 - 2^{-\frac{b}{2a}}\right)^2 \cdot \ln\left(\frac{2m}{n}\right)}{\left(\gamma \cdot 2^{-i/2} \cdot \sqrt{n}\right)^2 \cdot s}\right) \\
&\overset{\text{(xii)}}{\leq} 2 \cdot exp\left(-\frac{1}{8} \cdot \frac{C^2 \cdot \left(1 - 2^{-\frac{b}{2a}}\right)^2}{b} \cdot 2^{\left(1 - \frac{2b}{a}\right) \cdot i} \cdot \ln\left(\frac{2m}{n}\right)\right) \\
&= 2 \cdot \left(\frac{n}{2m}\right)^{\frac{1}{8} \cdot \frac{C^2 \cdot \left(1 - 2^{-\frac{b}{2a}}\right)^2}{b} \cdot 2^{\left(1 - \frac{2b}{a}\right) \cdot i}} \\
&\overset{\text{(xiii)}}{\leq} 2 \cdot \left(\frac{n}{2m}\right)^{\frac{1}{8} \cdot \frac{C^2 \cdot \left(1 - 2^{-\frac{b}{2a}}\right)^2}{b}} \cdot 2^{-2^{\left(1 - \frac{2b}{a}\right) \cdot i}} \\
&\overset{\text{(xiv)}}{\leq} 2 \cdot \left(\frac{n}{2m}\right)^{\frac{1}{8} \cdot \frac{C^2 \cdot \left(1 - 2^{-\frac{b}{2a}}\right)^2}{b}} \cdot 2^{-i} \\
&\overset{\text{(xv)}}{\leq} 2 \cdot 2^{-\frac{1}{8} \cdot \frac{C^2 \cdot \left(1 - 2^{-\frac{b}{2a}}\right)^2}{b}} \cdot \frac{n}{m} \cdot 2^{-i}
\end{aligned}$$

Inequality (xi) is obtained by plugging into the tail bound of [4] inequalities (9) and (8) for any $\tau_i \leq r \leq \tau_i + s$. Inequality (xii) follows since $s \leq {}^{b}/_{\gamma^2}$. Inequality (xiii) is derived from the conditions on the constants, specifically that $C \cdot \left(1 - 2^{-b/(2a)}\right) \geq \sqrt{32b}$, and the fact that $m \geq n$ (and hence ${}^{n}/_{(2m)} \leq {}^{1}/_{2}$). Inequality (xiv) is derived again from the conditions on the constants, specifically that $a \geq 8b$, which implies that $2^{\left(1 - \frac{2b}{a}\right) \cdot i} \geq i$ for every $i \geq 0$. Inequality (xv), similarly to (xiii), is derived from $C \cdot \left(1 - 2^{-b/(2a)}\right) \geq \sqrt{32b}$. Linearity of expectation concludes the proof since $|\mathcal{S}| = m$. □