Matrix Sparsification and Applications

Arpon Basu Ştefan Tudose

June 13, 2025

Contents

1 The Main Theorem: Matrix Coloring

For any symmetric matrix $A \in \mathbb{R}^{n\times n}$ with eigenvalue decomposition $A = \sum_{i=1}^{n} \lambda_i v_i v_i^{\mathsf{T}}$, and any function $f : \mathbb{R} \to \mathbb{R}$, define $f(A) := \sum_{i=1}^{n} f(\lambda_i) v_i v_i^{\mathsf{T}}$. In particular, $|A| = \sum_{i=1}^{n} |\lambda_i| v_i v_i^{\mathsf{T}}$ is a PSD matrix.

Theorem 1.1 (Deterministic Matrix Partial Coloring). *Let $A_1, \ldots, A_m \in \mathbb{R}^{n\times n}$ be symmetric matrices such that $\sum_{i\leqslant m} |A_i| \preceq \mathrm{Id}$. Let $\mathcal{C} \subseteq \mathbb{R}^m$ be a set of good partial "colorings", i.e.*

$$\mathcal{C} := \left\{ x \in \mathbb{R}^m : \left\| \sum_{i\leqslant m} x_i A_i \right\|_2 \leqslant 16\sqrt{\frac{n}{m}} \right\}.$$

Further suppose $\mathcal{H} \subseteq \mathbb{R}^m$ is a subspace of $\mathbb{R}^m$ with dimension $\geqslant 0.8m$. Then there exists a deterministic polynomial time algorithm that returns a coloring $x \in [-1, 1]^m$ such that $x \in \mathcal{C} \cap \mathcal{H}$ and $|\{i \in [m] : x_i = \pm 1\}| \geqslant \Omega(m)$.

Remark. *A few remarks are due:*

1. *In a first reading of the theorem, it is convenient to assume that $\mathcal{H} = \mathbb{R}^m$. The reason we introduce the theorem with $\mathcal{H}$ is that often we would like our coloring to satisfy some additional properties, and we can use $\mathcal{H}$ to encode those properties. The key realization in the proof of the above theorem is that when a good partial coloring exists, actually an entire subspace (of dimension $\Omega(m)$) of good partial colorings exists, and thus as long as* codim$(\mathcal{H})$ *is small, we can impose the condition encoded by $\mathcal{H}$ "for free" on our coloring.*

2. *Note that in the setting of the matrix Spencer conjecture, we require $\|A_i\|_2 \leqslant 1$ for all $i \in [m]$, in which case $\|\sum_{i\leqslant m} |A_i|\|_2$ can be as large as m. Rescaling the matrices in the above theorem by m to make it commensurate with the setting of the matrix Spencer conjecture, we obtain a partial coloring with $\|\sum_{i\leqslant m} x_i A_i\|_2 \leqslant O(m \cdot \sqrt{n/m})$, while the matrix Spencer conjecture asks for a discrepancy of $O(\sqrt{m} \cdot \sqrt{\log(n/m)})$ (and matrix Chernoff provides $O(\sqrt{m \log m})$ very easily).*

3. *This theorem was first proven in [RR19] using methods from convex geometry. The proof we shall present below is due to [LWZ24], who removed all convex-geometric techniques, and gave a proof using the so-called* deterministic discrepancy walk*, introduced in the context of vector coloring by Pesenti and Vladu [PV23]. The discrepancy walk (and the potential function therein) should be seen as a direct generalization of the barrier function method in Batson, Spielman, and Srivastava [BSS09].*

We shall see a proof of the above theorem in Section 3. But before that, let's explore some applications of the above theorem.

2 Matrix Sparsification

Theorem 2.1 (Deterministic Matrix Sparsification). *Given PSD matrices $A_1, \ldots, A_m \in \mathbb{R}^{n\times n}$ such that $\sum_{i\leqslant m} A_i \preceq \mathrm{Id}$ and a subspace $\mathcal{H} \subseteq \mathbb{R}^m$ of dimension $m - O(n)$, there exists a deterministic polynomial time algorithm to construct $s \in \mathbb{R}^m_{\geqslant 0}$ such that $|\operatorname{supp}(s)| \leqslant O(n/\varepsilon^2)$ such that $s - \mathbf{1}_m \in \mathcal{H}$ and $\|\sum_{i\leqslant m} s_i A_i - \sum_{i\leqslant m} A_i\|_2 \leqslant \varepsilon$.*

As in [RR19], we can convert a partial coloring theorem into a matrix sparsification statement. We shall just sketch a proof here: Basically, we run an iterative algorithm, where at each step we invoke Theorem 1.1 to obtain a partial coloring $x_t \in \mathbb{R}^m$. We also maintain our sparsifier $s_t \in \mathbb{R}^m$ as $s_t(i) := s_{t-1}(i)(1 + x_t(i))$. Thus whenever $x_t(i)$ is set to -1, we zero out that particular coordinate in s_t. [1] In the next round, we invoke Theorem 1.1 to find a coloring x_{t+1} of $\{s_t(i)A_i\}$.

[1] if there are more 1s than -1s among the frozen coordinates of x, we flip the sign of x

We terminate this iteration when $|\operatorname{supp}(s_t)| \leqslant O(n/\varepsilon^2)$. Since $\operatorname{supp}(s_t)$ decreases by a constant factor in every round, the above iteration terminates within $O(\log(\varepsilon^2 m/n))$ steps, leaving us with $s \in \mathbb{R}^m_{\geqslant 0}$ with $|\operatorname{supp}(s)| \leqslant O(n/\varepsilon^2)$.

We can use matrix sparsification to recover the graph sparsification result of [BSS09] by taking $A_e := L_G^{\dagger/2} L_e L_G^{\dagger/2}$, where $e \in E(G)$, L_e is the edge Laplacian, and L_G is the graph Laplacian. Furthermore, we can get a degree-preserving sparsifier by exploiting the subspace $\mathcal{H}$ in Theorem 2.1: Indeed, note that $\sum_{u:u\sim v} s_t(uv) = \sum_{u:u\sim v}(1 + x_t(uv)) \cdot s_{t-1}(uv) = \sum_{u:u\sim v} s_{t-1}(uv) + \sum_{u:u\sim v} x_t(uv)s_{t-1}(uv)$. Thus, at the t^{th} iteration, we can set $\mathcal{H} := \{x \in \mathbb{R}^m : \sum_{u:u\sim v} x_t(uv)s_{t-1}(uv) = 0\}$ to preserve the degree.

Although we won't talk much about it, matrix sparsification can also be used to prove stronger versions of graph sparsification: For example, **unit circle sparsification**, and **singular value sparsification** (see [APP+23] for the definition and applications). We briefly talk about unit circle sparsification below:

Definition 2.1 (Unit Circle Sparsification)**.** *Let G be a graph, let $L_G := D_G - A_G$ be the (signed) Laplacian of G, and let $U_G := D_G + A_G$ be the unsigned Laplacian of G. Then a graph G' is called a ε-unit-circle sparsifier of G if:*

1. *$(1-\varepsilon)L_G \preceq L_{G'} \preceq (1+\varepsilon)L_G$, i.e. G' is a spectral sparsifier in the usual sense,*
2. *$(1-\varepsilon)U_G \preceq U_{G'} \preceq (1+\varepsilon)U_G$, i.e. G' preserves the unsigned Laplacians too,*
3. *$D_G = D_{G'}$, i.e. G' preserves the degrees of G.*

Remark. *The definition of unit circle sparsifiers in [LWZ24, Definition 2.10] has something to do with complex numbers, and technically the conditions stated above* imply *their version of unit circle sparsification (see [LWZ24, Lemma 4.2]), but we found the above conditions to be more well-motivated.*

Then we have an extension of [BSS09] for unit circle sparsification (which follows from the general matrix sparsification Theorem 2.1):

Theorem 2.2. *Given an undirected graph G with n vertices, there exists a unit circle sparsifier with $O(n/\varepsilon^2)$ edges. Furthermore, it can be found in* $\operatorname{poly}(n, 1/\varepsilon)$ *time.*

Remark. *This theorem follows easily from Theorem 2.1 by taking $A_e := \begin{bmatrix} L_G^{\dagger/2} L_e L_G^{\dagger/2} & 0 \\ 0 & U_G^{\dagger/2} U_e U_G^{\dagger/2} \end{bmatrix}$, where U_e, U_G are the unsigned Laplacians of $e \in E(G)$ and the graph G respectively. This notion of sparsification was introduced by [AKM+22] who used it to study low-space Laplacian solvers for undirected and Eulerian directed graphs.*

3 Proof of Theorem 1.1

Consider the following general framework for a "deterministic discrepancy walk" with a potential function $\Phi(\cdot)$:

Deterministic Discrepancy Walk

1. Set $x_0 = 0 \in \mathbb{R}^m, H_1 = [m], t = 1, \alpha := 1/(2\eta)$. H_* should be thought of as the set of "active coordinates".
2. While $m_t := |H_t| \geqslant 3m/4$ is large,
   (a) Choose a **unit vector** $y_t \in x_{t-1}^{\perp}$ such that $\Phi(x_{t-1} + y_t) - \Phi(x_{t-1})$ is "small".
   (b) Let $\delta_t \leqslant \alpha$ be the largest step size such that $x_t \leftarrow x_{t-1} + \delta_t y_t$ lies in $[-1, 1]^m$.
   (c) Update $x_t \leftarrow x_{t-1} + \delta_t y_t, t \leftarrow t + 1, H_t \leftarrow \{i \in [n] : x_t(i) \neq \pm 1\}$.
3. Return x_T, where T is the number of iterations after which the above loop was broken.

The way this walk is analyzed is as follows: Suppose we want to find a ± 1-valued coloring which minimizes some quantity. Then $\Phi(\cdot)$ is supposed to represent some smooth proxy for the quantity we wish to minimize, such that finding the "update vector" y_t is not too difficult. Note that because our walk seeks to "push" the x_ts towards the boundary by choosing δ_t to be as large as possible so that x_t touches the boundary, at every step of the walk, hopefully we round a few more coordinates to ± 1, thus decreasing the number of active coordinates. Finally, the discrepancy of the solution returned by the algorithm is $\approx \Phi(x_T)$, which is small since all the increments $\Phi(x_{t-1} + \delta_t y_t) - \Phi(x_{t-1})$ are small.
We shall also run a discrepancy walk to prove Theorem 1.1. Following [AZLO15, LRR17, BLV22, PV23], set

$$\Phi(x) := \max_{M \in \Delta_n} \langle A(x), M \rangle + \frac{2}{\eta} \operatorname{tr}(M^{1/2}),$$

where $A(x) := \sum_{i \leqslant m} x_i A_i$, $\Delta_n := \{M \succeq 0 : \operatorname{tr}(M) = 1\}$ is the set of all *density matrices*, and $\eta := \sqrt{m}/4$. We also note a very easy proposition:

Proposition 3.1. *If $M \in \Delta_n$, then $\operatorname{tr}(M^{1/2}) \leqslant \sqrt{n}$, with equality being achieved when all eigenvalues of M are equal.*

Proof. Let the eigenvalues of M be $\lambda_1, \ldots, \lambda_n$. Then $1 = \operatorname{tr}(M) = \sum_i \lambda_i$, and $\operatorname{tr}(M^{1/2}) = \sum_i \lambda_i^{1/2}$. The claim now follows from Cauchy-Schwarz inequality. □

Note that $\max_{M \in \Delta_n} \langle A(x), M \rangle = \lambda_{\max}(A(x))$, which equals $\|A(x)\|_2$ (which is the quantity Theorem 1.1 seeks to minimize) if $A(x)$ is PSD. To remedy this slight discrepancy (no pun intended) between the maximum eigenvalue and the spectral norm, note that $\|A\|_2 = \lambda_{\max}\left(\begin{bmatrix} A & 0 \\ 0 & -A \end{bmatrix}\right)$, and thus by blowing up the dimension of our matrices by a factor of 2 if necessary, we can WLOG replace all spectral norms with maximum eigenvalues.

Intuition for $\Phi(\cdot)$

Note that as explained above, $\langle A(x), M\rangle$ extracts the maximum eigenvalue of $A(x)$, which we want. However, why is the *regularizing term* $\operatorname{tr}(M^{1/2})$ necessary? Note that the $\operatorname{tr}(M^{1/2})$ term promotes that the eigenvalues of M are "spread" by Proposition 3.1. In other words, $\operatorname{tr}(M^{1/2})$ forces M to be somewhat isotropic. This is important because note that by choosing a suitable y, one can abruptly change the direction along which $\langle A(x+y), M\rangle$ achieves its maximum eigenvalue. In such a situation, an isotropic M prevents $\Phi(\cdot)$ from changing too much. However, $\operatorname{tr}(M^{1/2})$ isn't the only regularizing term which works: Indeed, just as our choice of $\operatorname{tr}(M^{1/2})$ yields a matrix sparsification statement (see Theorem 2.1) implying the Batson-Spielman-Srivastava bound [BSS09], an "entropic" regularizing term of $-\langle M, \ln M\rangle$ coupled with the above discrepancy walk yields the Spielman-Srivastava bound [SS09, AZLO15], i.e. the above discrepancy walk can be seen as a derandomization of matrix Chernoff in this context.

Furthermore, by Proposition 3.1, for any $M \in \Delta_n$, $\operatorname{tr}(M^{1/2}) \leqslant \sqrt{n}$. Consequently,

$$\lambda_{\max}(A(x_T)) \leqslant \Phi(x_T) = \Phi(0) + \sum_{t=1}^{T} \left(\Phi(x_{t-1} + \delta_t y_t) - \Phi(x_{t-1})\right) \leqslant \frac{2\sqrt{n}}{\eta} + \sum_{t=1}^{T} \left(\Phi(x_{t-1} + \delta_t y_t) - \Phi(x_{t-1})\right). \tag{1}$$

We now quantify the exact difference $\Phi(x+y) - \Phi(y)$ to be used in our analysis later on:

Lemma 3.2. *For any symmetric matrices $A_1, \dots, A_m$ and vector $x \in \mathbb{R}^m$, there is a unique $M \in \Delta_n$ which maximizes $\langle A(x), M\rangle + \frac{2}{\eta}\operatorname{tr}(M^{1/2})$. Furthermore, there exists a unique $u \in \mathbb{R}$ such that $(u \cdot \mathrm{Id} - \eta A(x))^{-2}$ is the maximizer. Moreover, for any $y \in \mathbb{R}^m$, and for $M = \arg\max_{M \in \Delta_n} \langle A(x), M\rangle + \frac{2}{\eta}\operatorname{tr}(M^{1/2})$, if we have $\|M^{1/2}A(y)\|_2 \leqslant \frac{1}{2\eta}$, then:*

$$\Phi(x+y) - \Phi(x) \leqslant \langle A(x), M\rangle + 2\eta \operatorname{tr}(M^{1/2}A(y)M^{1/2}A(y)M^{1/2}).$$

See Appendix A for a proof of the above statement.
With an intent to apply Lemma 3.2 to (1), we note that

$$\|M^{1/2}A(\delta_t y_t)\|_2 \leqslant \|M^{1/2}\|_2 \cdot \|A(\delta_t y_t)\|_2 \leqslant \left\|\sum_{i=1}^{m} \delta_t y_t(i) A_i\right\|_2 \leqslant \delta_t \cdot \|y_t\|_\infty \cdot \left\|\sum_{i=1}^{m} |A_i|\right\|_2 \leqslant \delta_t \leqslant \alpha.$$

Here the first inequality follows by the sub-multiplicativity of $\|\cdot\|_2$, the second inequality follows since $M \in \Delta_n \implies \|M^{1/2}\|_2 = \lambda_{\max}(M^{1/2}) = \lambda_{\max}(M)^{1/2} \leqslant 1$, and $\|y_t\|_\infty \leqslant 1$ since y_t is a unit vector. Finally, from the hypothesis of Theorem 1.1, since $\sum_{i=1}^{m} |A_i| \preceq \mathrm{Id}, \alpha := 1/(2\eta)$, (2) holds:

$$\lambda_{\max}(A(x_T)) \leqslant \Phi(x_T) \leqslant \frac{2\sqrt{n}}{\eta} + \sum_{t=1}^{T} \delta_t \langle A(y_t), M_t\rangle + 2\eta\delta_t^2 \operatorname{tr}(M_t^{1/2}A(y_t)M_t^{1/2}A(y_t)M_t^{1/2}). \tag{2}$$

where $M_t := (u_t \, \mathrm{Id}_n - \eta A(x_{t-1}))^{-2}$, where $u_t \in \mathbb{R}$ is chosen according to Lemma 3.2 so that $M_t = \arg\max_{M \in \Delta_n} \langle A(x), M\rangle + \frac{2}{\eta}\operatorname{tr}(M^{1/2})$.
Now, note that

$$\operatorname{tr}(M_t^{1/2}A(y_t)M_t^{1/2}A(y_t)M_t^{1/2}) = \sum_{i,j} y_t(i)y_t(j) \operatorname{tr}(M_t^{1/2}A_iM_t^{1/2}A_jM_t^{1/2}) = \sum_{i,j \in H_t} y_t(i)y_t(j) \operatorname{tr}(M_t^{1/2}A_iM_t^{1/2}A_jM_t^{1/2})$$

$$= (y_t|_{H_t})^\top N_t (y_t|_{H_t})$$

where $y_t|_{H_t}$ is the restriction of y_t to the coordinates in H_t, and the matrix N_t is defined as $(N_t)_{ij} := \mathrm{tr}(M_t^{1/2} A_i M_t^{1/2} A_j M_t^{1/2})$. Note that in the second equality we use the fact that $\mathrm{supp}(y) \subseteq H_t$. Furthermore, $(N_t)_{ij} = \mathrm{tr}(M_t^{1/2} A_i M_t^{1/2} A_j M_t^{1/2}) = \langle M_t^{1/2} A_i M_t^{1/4}, M_t^{1/4} A_j M_t^{1/2} \rangle$, and thus N_t is a Gram matrix, and hence PSD. Let $N_t = \sum_{i=1}^{m_t} \lambda_i u_i u_i^\top$ be the eigenvalue decomposition of N_t with $0 \leqslant \lambda_1 \leqslant \cdots \leqslant \lambda_{m_t}$. At this point, we explain how to make the "choice" of y_t in Item 2a, we make a few definitions:

1. $U^0 := \{y_i = 0 : i \notin \mathrm{supp}(H_t)\}$: Forcing $y \in U^0$ ensures $\mathrm{supp}(y) \subseteq H_t$,
2. $U^1 := \{y \in \mathbb{R}^m : y \in x_{t-1}^\perp\}$: This is self-explanatory. Note that forcing y_t to be orthogonal to x_{t-1} implies that $x_t = x_{t-1} + \delta_t y_t$ satisfies $\|x_t\|^2 = \|x_{t-1}\|^2 + \delta_t^2$. By arguing that δ_t can't be too small too often, this allows use to bound the number of steps in our walk. Note that $\mathrm{codim}(U^1) \leqslant 1$.
3. $U^2 := \{y \in \mathbb{R}^m : \langle A(y), M_t \rangle = 0\}$: This is to ensure that the linear term in (2) vanishes, which simplifies the analysis. Note that $\mathrm{codim}(U^2) \leqslant 1$.
4. $U^3 := \{y \in \mathbb{R}^m : y|_{H_t} \in \mathrm{span}\{u_1, \ldots, u_{m_t/3}\}\}$: This is to ensure that y comes from a "low" eigenspace of N_t, which would imply that the $\mathrm{tr}(M_t^{1/2} A(y_t) M_t^{1/2} A(y_t) M_t^{1/2})$ term is small.

Write $\mathcal{U} := U^0 \cap U^1 \cap U^2 \cap U^3$. We choose y from $\mathcal{U}$, and note that $\dim(\mathcal{U}) \geqslant m_t - 2 - 2m_t/3 = m_t/3 - 2 \geqslant m/4 - 2$, where the last inequality follows from the fact that $m_t \geqslant 3m/4$. Thus, we always have plenty of choices for y. Now we need to argue that $y \in U^3$ implies that the quadratic term in (2) is actually small.

Lemma 3.3. *Given symmetric matrices $A_1, \ldots, A_m \in \mathbb{R}^{n \times n}$ satisfying $\sum_{i \leqslant m} |A_i| \preceq \mathrm{Id}$, any unit vector $y \in U^0 \cap U^3$ satisfies $\mathrm{tr}(M_t^{1/2} A(y) M_t^{1/2} A(y) M_t^{1/2}) \leqslant 9\sqrt{n}/m_t^2$.*

Proof. Note that $\mathrm{tr}(M_t^{1/2} A(y) M_t^{1/2} A(y) M_t^{1/2}) = (y|_{H_t})^\top N_t (y|_{H_t}) \leqslant \lambda_{m_t/3}$ by the definition of U^3. To bound $\lambda_{m_t/3}$, we use the Cauchy interlacing theorem: Recall that the Cauchy interlacing theorem says that if $Y \in \mathbb{R}^{m \times m}$ is a principal sub-matrix of a symmetric matrix $X \in \mathbb{R}^{n \times n}$, and if the eigenvalues of X, Y are $\beta_1 \leqslant \cdots \leqslant \beta_n$ and $\gamma_1 \leqslant \cdots \leqslant \gamma_m$ respectively, then $\beta_i \leqslant \gamma_i$ for $1 \leqslant i \leqslant m$. More precisely, write $S := \{i \in H_t : \langle M_t^{1/2}, |A_i| \rangle \geqslant 3\,\mathrm{tr}(M_t^{1/2})/m_t\}$, and let N_t' be the principal sub-matrix induced by $H_t \setminus S$. Now, note that

$$\sum_{i \in H_t} \langle M_t^{1/2}, |A_i| \rangle = \mathrm{tr}(M_t^{1/2} \cdot \sum_{i \in H_t} |A_i|) \leqslant \mathrm{tr}(M_t^{1/2}) \cdot \| \sum_{i \in H_t} |A_i| \| \leqslant \mathrm{tr}(M_t^{1/2}),$$

and thus by Markov's inequality $|S| \leqslant m_t/3$. By Cauchy interlacing,

$$\lambda_{m_t/3} \leqslant \lambda_{m_t/3}(N_t') \leqslant \lambda_{|H_t \setminus S|/2}(N_t') \leqslant 2\,\mathrm{tr}(N_t')/|H_t \setminus S| \leqslant 3\,\mathrm{tr}(N_t')/m_t,$$

where the second last inequality also follows from Markov's inequality. However,

$$\mathrm{tr}(N_t') = \sum_{i \in H_t \setminus S} \mathrm{tr}(M_t^{1/2} A_i M_t^{1/2} A_i M_t^{1/2}) = \sum_{i \in H_t \setminus S} \mathrm{tr}(M_t A_i M_t^{1/2} A_i) \leqslant \sum_{i \in H_t \setminus S} \mathrm{tr}(M_t A_i) \cdot \mathrm{tr}(M_t^{1/2} A_i),$$

where the last inequality follows by noting that for symmetric matrices A, B, C with $A, B \succeq 0$, we have $\mathrm{tr}(ACBC) \leqslant \mathrm{tr}(A \cdot |C|)\,\mathrm{tr}(B \cdot |C|)$. [2] But by the definition of S,

$$\sum_{i \in H_t \setminus S} \mathrm{tr}(M_t A_i) \cdot \mathrm{tr}(M_t^{1/2} A_i) \leqslant \frac{3\,\mathrm{tr}(M_t^{1/2})}{m_t} \cdot \sum_{i \in H_t \setminus S} \mathrm{tr}(M_t |A_i|) \leqslant \frac{3\sqrt{n}}{m_t} \cdot \mathrm{tr}\left(M_t \sum_{i \in H_t \setminus S} |A_i| \right) \leqslant \frac{3\sqrt{n}}{m_t} \cdot \mathrm{tr}(M_t) = \frac{3\sqrt{n}}{m_t}.$$

[2] see [RR19, Lemma 10] for a proof

Here we use Proposition 3.1 and the fact that $\sum |A_i| \preceq \mathrm{Id}$. Consequently, $\lambda_{m_t/3} \leqslant 3\,\mathrm{tr}(N_t')/m_t \leqslant 9\sqrt{n}/m_t^2$, as desired. $\square$

We are finally ready to prove Theorem 1.1:

Proof. We apply the discrepancy walk to prove the theorem. Note that the choice of ys we had at any point was from the subspace $\mathcal{U}$. Since we also want our final coloring to belong to $\mathcal{H}$, we work with the subspace $\mathcal{U} \cap \mathcal{H}$. Note that $\dim(\mathcal{U} \cap \mathcal{H}) \geqslant \dim(\mathcal{U}) - \mathrm{codim}(\mathcal{H}) \geqslant m/4 - 2 - m/5 > 0$ for $m \geqslant \Omega(1)$. Now, rewriting (2) we obtain

$$\lambda_{\max}(A(x_T)) \leqslant \Phi(x_T) \leqslant \frac{2\sqrt{n}}{\eta} + \sum_{t=1}^{T} \delta_t \langle A(y_t), M_t \rangle + 2\eta\delta_t^2\,\mathrm{tr}(M_t^{1/2} A(y_t) M_t^{1/2} A(y_t) M_t^{1/2}).$$

Since $y \in \mathcal{U} \subseteq U_2$, $\langle A(y_t), M_t \rangle = 0$. Thus, simplifying the above inequality yields

$$\lambda_{\max}(A(x_T)) \leqslant \frac{2\sqrt{n}}{\eta} + 2\eta \sum_{t=1}^{T} \delta_t^2\,\mathrm{tr}(M_t^{1/2} A(y_t) M_t^{1/2} A(y_t) M_t^{1/2}).$$

By Lemma 3.3, the above inequality further simplifies to

$$\lambda_{\max}(A(x_T)) \leqslant \frac{2\sqrt{n}}{\eta} + 18\eta\sqrt{n} \sum_{t=1}^{T} \frac{\delta_t^2}{m_t^2} \leqslant \frac{2\sqrt{n}}{\eta} + \frac{32\eta\sqrt{n}}{m^2} \sum_{t=1}^{T} \delta_t^2.$$

where the last inequality follows since the algorithm only runs when $m_t \geqslant 3m/4$. On the other hand, note that since $x_T \in [-1,1]^m$, we have $m \geqslant \|x_T\|^2 = \|x_{T-1}\|^2 + \delta_T^2 \geqslant \cdots \geqslant \sum_{t=1}^{T} \delta_t^2$, where the second equality follows since y_t is perpendicular to x_{t-1}, and thus the norms just accumulate according to Pythagoras's theorem. Consequently, $\lambda_{\max}(A(x_T)) \leqslant \frac{2\sqrt{n}}{\eta} + \frac{32\eta\sqrt{n}}{m} \leqslant 16\sqrt{\frac{n}{m}}$, where the last inequality follows since $\eta := \sqrt{m}/4$.
Now, we only have to bound the run-time of the algorithm, which is $\mathrm{poly}(n,m) \cdot T$. Now, note that in Item 2c, either we're able to choose δ_t so that $x_{t-1} + \delta_t y_t$ hits the boundary of $[-1,1]^m$ (and consequently at least one coordinate gets frozen), or we get capped at $\delta_t := \alpha$. But, in the latter case, $\|x_t\|^2 = \|x_{t-1}\|^2 + \alpha^2$. Since $\|x_t\|^2 \leqslant m$, the latter situation can only occur $\leqslant m/\alpha^2$ times. The former situation (where a coordinate gets frozen) can obviously only occur $\leqslant m$ times. Consequently, $T \leqslant m + m/\alpha^2$. Recall that $\alpha = 1/(2\eta) = 2/\sqrt{m}$, and thus $T \leqslant m^2/4 + m = O(m^2)$, as desired. $\square$

A Appendix

Remark. *There is a small mistake in the proof of Lemma 3.2 in [LWZ24], in that they claim in [LWZ24, Lemma 3.1] that there exists a unique u such that $(uI_n - \eta A)^{-2} \in \Delta_n$, but this is manifestly not the case. The fact that there exist (at least) two u such that $(uI_n - \eta A)^{-2} \in \Delta_n$ should not come as a surprise: if we change max to min in the statement of the lemma, the proof goes through in the same way. However, small modifications to their proof makes it go through, which is what we present below.*

Proof of Lemma 3.2. Let us begin with the following:

Lemma A.1. *Given a symmetric $n \times n$ real-valued matrix A, then*

$$\arg\max_{M \in \Delta_n} \langle A, M \rangle + \frac{2}{\eta}\,\mathrm{tr}(M^{1/2}) = (uI_n - \eta A)^{-2}$$

for the unique real number $u \in (\eta\lambda_{\max}(A), \infty)$ for which $(uI_n - \eta A)^{-2} \in \Delta_n$.

Proof. A maximizer M exists because Δ_n is compact. Let B be a symmetric $n \times n$ matrix such that $\operatorname{tr} B = 0_n$. For ε small enough, $M + \varepsilon B \in \Delta_n$. The function

$$\varepsilon \mapsto \operatorname{tr}\left(A(M+\varepsilon B)\right) + \frac{2}{\eta}\operatorname{tr}\left((M+\varepsilon B)^{1/2}\right)$$

has a minimum at $\varepsilon = 0$, so the derivative at $\varepsilon = 0$ must be zero. We have

$$\frac{d}{d\varepsilon}\operatorname{tr}\left((M+\varepsilon B)^{1/2}\right) = \operatorname{tr}\left(\frac{d}{d\varepsilon}(M+\varepsilon B)^{1/2}\right) = \operatorname{tr}\left(\frac{d}{d\varepsilon}\left(M^{1/2}(I+\varepsilon M^{-1}B)^{1/2}\right)\right) = \frac{1}{2}\operatorname{tr}\left(M^{1/2}(I+\varepsilon M^{-1}B)^{-1/2}M^{-1}B\right)$$

We used that

$$(I+\varepsilon C)^{1/2} = \sum_{k=0}^{\infty}\binom{1/2}{k}(\varepsilon C)^k = \sum_{k=0}^{\infty}\frac{(-1)^{k-1}}{2^{2k-1}k}\binom{2k-2}{k-1}\varepsilon^k C^k$$

to be able to say that the function is differentiable when ε is close to zero.

Therefore the derivative at $\varepsilon = 0$ equals

$$\operatorname{tr}(AB) + \frac{1}{\eta}\operatorname{tr}(M^{-1/2}B)$$

so we have obtained that

$$\operatorname{tr}\left(\left(\eta A + M^{-1/2}\right)B\right) = 0$$

for any symmetric matrix B with $\operatorname{tr} B = 0_n$. Note that the matrices with zero trace are precisely the orthogonal complement (with respect to the Hilbert-Schmidt inner product) of the identity matrix. Since $\eta A + M^{-1/2}$ is orthogonal to all of them, it must be the case that $\eta A + M^{-1/2}$ belongs in the subspace generated by the identity and thus

$$\eta A + M^{-1/2} = u_* I_n$$

for some real number u_*.

If $\lambda_1 \leq \lambda_2 \leq \ldots \lambda_n$ are the eigenvalues of A, note that the function

$$f(u) = \sum_{i=1}^{n}\frac{1}{(u-\eta\lambda_i)^2}$$

is strictly decreasing on $(\eta\lambda_n, \infty)$, approaches $+\infty$ as $u \to \eta\lambda_n$ and it approaches 0 as $u \to \infty$, so there exists a unique $u_1 > \eta\lambda_n$ such that $f(u_1) = 1$. There do exist other values[3] of u such that $f(u) = 1$ and all such values ensure that $(uI_n - \eta A)^{-2} \in \Delta_n$.

We now claim that $u_* = u_1$. A simple computation shows that for $M = (uI_n - \eta A)^{-2}$,

$$\operatorname{tr}(AM) + \frac{2}{\eta}\operatorname{tr}(M^{1/2}) = \frac{1}{\eta}\left(u + \sum_{i=1}^{n}\frac{1}{u-\eta\lambda_i}\right) = g(u)$$

so we need to show that

$$u_1 = \arg\max_{f(u)=1} g(u)$$

[3]The function f approaches $+\infty$ as $u \to \eta\lambda_1$ and it approaches 0 as $u \to -\infty$ so there exists some $u_3 < \eta\lambda_1$ such that $f(u_3) = 1$.

Let $u_2 < \eta\lambda_n < u_1$ be such that $f(u_2) = 1$. We will show that $g(u_2) < g(u_1)$. By Cauchy-Schwarz,

$$\left(\sum_{i=1}^{n} \frac{1}{(u_1 - \eta\lambda_i)(u_2 - \eta\lambda_i)}\right)^2 \leq f(u_1) \cdot f(u_2) = 1$$

but equality cannot be attained. Therefore

$$\eta\left(g(u_1) - g(u_2)\right) = \left(u_1 + \sum_{i=1}^{n} \frac{1}{u_1 - \eta\lambda_i}\right) - \left(u_2 + \sum_{i=1}^{n} \frac{1}{u_2 - \eta\lambda_i}\right) = (u_1 - u_2) \cdot \left(1 - \sum_{i=1}^{n} \frac{1}{(u_1 - \eta\lambda_i)(u_2 - \eta\lambda_i)}\right) > 0$$

□

Given $A(x) = \sum_{i=1}^{n} x_i A_i$, let us denote by u_x the unique real number in the interval $(\eta\lambda_{\max}(A(x)), \infty)$ for which

$$M_x = (u_x I_n - \eta A(x))^{-2} \in \Delta_n$$

By the previous lemma we know that

$$M_x = \underset{M \in \Delta_n}{\arg\max} \langle A(x), M \rangle + \frac{2}{\eta} \operatorname{tr}(M^{1/2})$$

We note that

$$\Phi(x) = \langle A(x), M_x \rangle + \frac{2}{\eta} \operatorname{tr}(M_x^{1/2}) = \frac{1}{\eta}\left(u_x + \operatorname{tr}(u_x I_n - \eta A(x))^{-2}\right)$$

Let us see how to control the increment in the potential function Φ:

Lemma A.2. *If y is chosen such that $\left\| M_x^{1/2} \eta A(y) \right\|_2 < 1$, then*

$$\Phi(x+y) - \Phi(x) \leq \frac{1}{\eta}\left(\operatorname{tr}(u_x I_n - \eta A(x+y))^{-1} - \operatorname{tr}(u_x I_n - \eta A(x))^{-1}\right)$$

Proof. We use the formula for the potential function to write

$$\Phi(x+y) - \Phi(x) = \frac{1}{\eta}\left(u_{x+y} - u_x + \operatorname{tr}(u_{x+y} I_n - \eta A(x+y))^{-1} - \operatorname{tr}(u_x I_n - \eta A(x))^{-1}\right)$$

Note that we have $u_{x+y} > \eta\lambda_{\max}(A(x+y))$ by the previous lemma and we have $\eta A(y) \prec M_x^{-1/2} = u_x I_n - \eta A(x)$ so $u_x > \eta\lambda_{\max}(A(x+y))$ as well. Since the function h given by

$$u \overset{h}{\mapsto} \operatorname{tr}(u I_n - \eta A(x+y))^{-1} = \sum_{i=1}^{n} \frac{1}{u - \eta\lambda_i(A(x+y))}$$

is convex on the interval $(\eta\lambda_{\max}(A(x+y)), \infty)$, we infer that

$$h(u_x) \geq h(u_{x+y}) + h'(u_{x+y})(u_x - u_{x+y})$$

Note that $h'(u_{x+y}) = -\sum_{i=1}^{n} \frac{1}{(u_{x+y} - \eta\lambda_i(A(x+y)))^2} = -\operatorname{tr} M_{x+y} = -1$, so the above gives

$$\operatorname{tr}(u_x I_n - \eta A(x+y))^{-1} \geq \operatorname{tr}(u_{x+y} I_n - \eta A(x+y))^{-1} + (u_{x+y} - u_x)$$

Plugging this in the formula for $\Phi(x+y) - \Phi(x)$ gives the claim of the lemma. □

To further control the density increment, we will use the following

Lemma A.3. *If A is positive definite and $\left\|A^{-1}B\right\|_2 \leq 1/2$ then*

$$\left|\operatorname{tr}((A-B)^{-1}) - \operatorname{tr} A^{-1} - \operatorname{tr}(A^{-1}BA^{-1})\right| \leq 2\operatorname{tr}(A^{-1}BA^{-1}BA^{-1})$$

Proof. We write

$$\operatorname{tr}((A-B)^{-1}) = \operatorname{tr}\left(\left(A(I-A^{-1}B)\right)^{-1}\right) = \operatorname{tr}(A^{-1}(I-A^{-1}B)^{-1}) = \operatorname{tr}\left(A^{-1}\sum_{i=0}^{\infty}(A^{-1}B)^i\right)$$

and therefore

$$\left|\operatorname{tr}((A-B)^{-1}) - \operatorname{tr} A^{-1} - \operatorname{tr}(A^{-1}BA^{-1})\right| = \left|\operatorname{tr}\left(A^{-1}\sum_{i=2}^{\infty}(A^{-1}B)^i\right)\right| = \left|\operatorname{tr}\left(A^{-1}(A^{-1}B)^2(I-A^{-1}B)^{-1}\right)\right|$$

We now use the following simple inequality[4]: if C is positive definite, then $|\operatorname{tr}(CD)| \leq \operatorname{tr}(C)\cdot\|D\|_2$. Using this for $C = A^{-1}(A^{-1}B)^2$ and $D = (I - BA^{-1})^{-1}$ we obtain

$$\left|\operatorname{tr}((A-B)^{-1}) - \operatorname{tr} A^{-1} - \operatorname{tr}(A^{-1}BA^{-1})\right| \leq \operatorname{tr}\left(A^{-1}(A^{-1}B)^2\right)\cdot\left\|(I-A^{-1}B)^{-1}\right\|_2 \leq 2\operatorname{tr}(A^{-1}BA^{-1}BA^{-1})$$

since $\left\|A^{-1}B\right\|_2 \leq 1/2$. ☐

Using this lemma with $A = u_x I_n - \eta A(x)$, $B = \eta A(y)$, we obtain that if y is chosen such that $\left\|(u_x I_n - A(x))^{-1}\eta A(y)\right\|_2 < 1/2$, then

$$\begin{aligned}\Phi(x+y) - \Phi(x) &\leq \frac{1}{\eta}\left(\operatorname{tr}(u_x I_n - \eta A(x+y))^{-1} - \operatorname{tr}(u_x I_n - \eta A(x))^{-1}\right)\\ &\leq \operatorname{tr}(M_x^{1/2}A(y)M_x^{1/2}) + 2\eta\operatorname{tr}(M_x^{1/2}A(y)M_x^{1/2}A(y)M_x^{1/2})\\ &= \operatorname{tr}(M_x A(y)) + 2\eta\operatorname{tr}(M_x^{1/2}A(y)M_x^{1/2}A(y)M_x^{1/2})\end{aligned}$$

which finishes the proof of Lemma 3.2. ☐

[4]This follows immediately by writing $C = \sum_{j=1}^{n}\mu_j w_j w_j^T$ for an orthonormal eigenbasis $\{w_j\}_{j\in[n]}$ and applying the triangle inequality.

References

[AKM⁺22] AmirMahdi Ahmadinejad, Jonathan Kelner, Jack Murtagh, John Peebles, Aaron Sidford, and Salil Vadhan. High-precision estimation of random walks in small space, 2022.

[APP⁺23] AmirMahdi Ahmadinejad, John Peebles, Edward Pyne, Aaron Sidford, and Salil Vadhan. Singular value approximation and sparsifying random walks on directed graphs, 2023.

[AZLO15] Zeyuan Allen-Zhu, Zhenyu Liao, and Lorenzo Orecchia. Spectral sparsification and regret minimization beyond matrix multiplicative updates. STOC '15, New York, NY, USA, 2015. Association for Computing Machinery.

[BLV22] Nikhil Bansal, Aditi Laddha, and Santosh Vempala. A Unified Approach to Discrepancy Minimization. In Amit Chakrabarti and Chaitanya Swamy, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2022)*, volume 245 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 1:1–1:22, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[BSS09] Joshua Batson, Daniel A. Spielman, and Nikhil Srivastava. Twice-ramanujan sparsifiers, 2009.

[LRR17] Avi Levy, Harishchandra Ramadas, and Thomas Rothvoss. Deterministic discrepancy minimization via the multiplicative weight update method, 2017.

[LWZ24] Lap Chi Lau, Robert Wang, and Hong Zhou. Spectral sparsification by deterministic discrepancy walk, 2024.

[PV23] Lucas Pesenti and Adrian Vladu. Discrepancy minimization via regularization. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1734–1758, 2023.

[RR19] Victor Reis and Thomas Rothvoss. Linear size sparsifier and the geometry of the operator norm ball, 2019.

[SS09] Daniel A. Spielman and Nikhil Srivastava. Graph sparsification by effective resistances, 2009.