A Matrix Expander Chernoff Bound

Ankit Garg, Yin Tat Lee, Zhao Song, Nikhil Srivastava

UC Berkeley
Thm [Hoeffding, Chernoff]. If $X_1, ..., X_k$ are independent mean zero random variables with $|X_i| \leq 1$ then

$$\mathbb{P} \left[ \left| \frac{1}{k} \sum_i X_i \right| \geq \epsilon \right] \leq 2\exp(-k\epsilon^2/4)$$
Vanilla Chernoff Bound

**Thm** [Hoeffding, Chernoff]. If $X_1, \ldots, X_k$ are independent mean zero random variables with $|X_i| \leq 1$ then

$$
P \left[ \left| \frac{1}{k} \sum_i X_i \right| \geq \epsilon \right] \leq 2 \exp\left(-k \epsilon^2 / 4\right)
$$

*Two Extensions:*

1. **Dependent** Random Variables
2. Sums of random matrices
Expander Chernoff Bound [AKS'87, G'94]

**Thm[Gillman’94]:** Suppose $G = (V, E)$ is a regular graph with transition matrix $P$ which has second eigenvalue $\lambda$. Let $f: V \to \mathbb{R}$ be a function with $|f(v)| \leq 1$ and $\sum_v f(v) = 0$. 
Expander Chernoff Bound [AKS’87, G’94]

**Thm[Gillman’94]:** Suppose $G = (V, E)$ is a regular graph with transition matrix $P$ which has second eigenvalue $\lambda$. Let $f : V \rightarrow \mathbb{R}$ be a function with $|f(v)| \leq 1$ and $\sum_v f(v) = 0$. Then, if $v_1, \ldots, v_k$ is a stationary random walk:

\[
\mathbb{P} \left[ \left| \frac{1}{k} \sum_i f(v_i) \right| \geq \epsilon \right] \leq 2 \exp(-c(1 - \lambda)k\epsilon^2)
\]

Implies walk of length $k \approx (1 - \lambda)^{-1}$ concentrates around mean.
Intuition for Dependence on Spectral Gap

\[ 1 - \lambda = \frac{1}{n^2} \]
Intuition for Dependence on Spectral Gap

$1 - \lambda = \frac{1}{n^2}$

Typical random walk takes $\Omega(n^2)$ steps to see both $\pm 1$. 
Thm[Gillman’94]: Suppose $G = (V, E)$ is a regular graph with transition matrix $P$ which has second eigenvalue $\lambda$. Let $f: V \to \mathbb{R}$ be a function with $|f(v)| \leq 1$ and $\Sigma_v f(v) = 0$. Then, if $v_1, \ldots, v_k$ is a stationary random walk:

$$\mathbb{P} \left[ \left| \frac{1}{k} \sum_i f(v_i) \right| \geq \epsilon \right] \leq 2 \exp(-c(1 - \lambda)k\epsilon^2)$$

To generate $k$ indep samples from $[n]$ need $k \log n$ bits. If $G$ is $3$-regular with constant $\lambda$, walk needs: $\log n + k \log(3)$ bits.
Derandomization Motivation

**Thm[Gillman’94]:** Suppose $G = (V, E)$ is a regular graph with transition matrix $P$ which has second eigenvalue $\lambda$. Let $f: V \rightarrow \mathbb{R}$ be a function with $|f(v)| \leq 1$ and $\sum_v f(v) = 0$. Then, if $v_1, ..., v_k$ is a stationary random walk:

$$\mathbb{P} \left[ \left| \frac{1}{k} \sum_i f(v_i) \right| \geq \epsilon \right] \leq 2\exp(-c(1 - \lambda)k\epsilon^2)$$

To generate $k$ indep samples from $[n]$ need $k \log n$ bits. If $G$ is 3-regular with constant $\lambda$, walk needs: $\log n + k \log(3)$ bits.

When $k = O(\log n)$ reduces randomness **quadratically**. Can completely derandomize in polynomial time.
Theorem [Rudelson’97, Ahlswede-Winter’02, Oliveira’08, Tropp’11...]. If $X_1, ..., X_k$ are independent mean zero random $d \times d$ Hermitian matrices with $||X_i|| \leq 1$ then

$$\mathbb{P}\left[\left\|\frac{1}{k} \sum_{i} X_i \right\| \geq \epsilon \right] \leq 2d \exp\left(-k\epsilon^2/4\right)$$
Matrix Chernoff Bound

**Thm** [Rudelson’97, Ahlswede-Winter’02, Oliveira’08, Tropp’11...]. If $X_1, ..., X_k$ are independent mean zero random $d \times d$ Hermitian matrices with $\|X_i\| \leq 1$ then

$$\mathbb{P} \left[ \left\| \frac{1}{k} \sum_i X_i \right\| \geq \epsilon \right] \leq 2d \exp\left( -k\epsilon^2 / 4 \right)$$

Factor $d$ is tight because of the diagonal case.

Very generic bound (no independence assumptions on the entries). Many applications + martingale extensions (see Tropp).
**Conj[Wigderson-Xiao’05]**: Suppose \( G = (V, E) \) is a regular graph with transition matrix \( P \) which has second eigenvalue \( \lambda \). Let \( f: V \to \mathbb{C}^{d \times d} \) be a function with \(||f(v)||| \leq 1 \) and \( \sum_v f(v) = 0 \). Then, if \( v_1, \ldots, v_k \) is a stationary random walk:

\[
P \left[ \left\| \frac{1}{k} \sum_i f(v_i) \right\| \geq \epsilon \right] \leq 2d \exp(-c(1 - \lambda)k\epsilon^2)
\]

Motivated by derandomized Alon-Roichman theorem.
Main Theorem

**Thm.** Suppose \( G = (V, E) \) is a regular graph with transition matrix \( P \) which has second eigenvalue \( \lambda \). Let \( f : V \to \mathbb{C}^{d \times d} \) be a function with \( \|f(v)\| \leq 1 \) and \( \sum_v f(v) = 0 \). Then, if \( v_1, \ldots, v_k \) is a stationary random walk:

\[
P \left[ \left\| \frac{1}{k} \sum_i f(v_i) \right\| \geq \epsilon \right] \leq 2d \exp(-c(1 - \lambda)k\epsilon^2)
\]

Gives black-box derandomization of any application of matrix Chernoff
1. Proof of Chernoff: reduction to mgf

**Thm [Hoeffding, Chernoff].** If $X_1, \ldots, X_k$ are independent mean zero random variables with $|X_i| \leq 1$ then

$$\mathbb{P}\left[\left|\frac{1}{k} \sum_{i} X_i \right| \geq \epsilon \right] \leq 2\exp(-k\epsilon^2/4)$$
Thm [Hoeffding, Chernoff]. If $X_1, \ldots, X_k$ are independent mean zero random variables with $|X_i| \leq 1$ then

$$\mathbb{P} \left[ \left| \frac{1}{k} \sum_i X_i \right| \geq \epsilon \right] \leq 2 \exp(-k\epsilon^2/4)$$

$$\mathbb{P} \left[ \sum_i X_i \geq k\epsilon \right] \leq e^{-tk\epsilon} \mathbb{E} \exp \left( t \sum_i X_i \right)$$

Markov
1. Proof of Chernoff: reduction to mgf

**Thm** [Hoeffding, Chernoff]. If $X_1, \ldots, X_k$ are independent mean zero random variables with $|X_i| \leq 1$ then

\[
\mathbb{P}\left[\left\{ \frac{1}{k} \sum_i X_i \geq \epsilon \right\} \right] \leq 2 \exp(-k \epsilon^2 / 4)
\]

\[
\mathbb{P}\left[ \sum_i X_i \geq k \epsilon \right] \leq e^{-tk\epsilon} \mathbb{E} \exp\left( t \sum_i X_i \right) = e^{-tk\epsilon} \prod_i \mathbb{E} e^{tX_i}
\]

Indep.
1. Proof of Chernoff: reduction to mgf

**Thm** [Hoeffding, Chernoff]. If \( X_1, \ldots, X_k \) are independent mean zero random variables with \( |X_i| \leq 1 \) then

\[
\mathbb{P} \left[ \left| \frac{1}{k} \sum_{i} X_i \right| \geq \epsilon \right] \leq 2 \exp \left( -k \epsilon^2 / 4 \right)
\]

\[
\mathbb{P} \left[ \sum_{i} X_i \geq k \epsilon \right] \leq e^{-tk\epsilon} \mathbb{E} \exp \left( t \sum_{i} X_i \right) = e^{-tk\epsilon} \prod_{i} \mathbb{E} e^{tX_i}
\]

\[
\leq e^{-tk\epsilon} (1 + t\mathbb{E}X_i + t^2)^k
\]

**Bounded**
1. Proof of Chernoff: reduction to mgf

**Thm** [Hoeffding, Chernoff]. If $X_1, \ldots, X_k$ are independent mean zero random variables with $|X_i| \leq 1$ then

\[
\mathbb{P} \left[ \left| \frac{1}{k} \sum_{i} X_i \right| \geq \epsilon \right] \leq 2 \exp(-k\epsilon^2/4)
\]

\[
\mathbb{P} \left[ \sum_{i} X_i \geq k\epsilon \right] \leq e^{-tk\epsilon} \mathbb{E} \exp \left( t \sum_{i} X_i \right) = e^{-tk\epsilon} \prod_{i} \mathbb{E} e^{tX_i}
\]

\[
\leq e^{-tk\epsilon} (1 + t^2)^k
\]

Mean zero
1. Proof of Chernoff: reduction to mgf

**Thm** [Hoeffding, Chernoff]. If $X_1, \ldots, X_k$ are independent mean zero random variables with $|X_i| \leq 1$ then

$$\mathbb{P}\left[ \left| \frac{1}{k} \sum_i X_i \right| \geq \epsilon \right] \leq 2\exp(-k\epsilon^2/4)$$

$$\mathbb{P}\left[ \sum_i X_i \geq k\epsilon \right] \leq e^{-tk\epsilon} \mathbb{E} \exp\left( t \sum_i X_i \right) = e^{-tk\epsilon} \prod_i \mathbb{E} e^{tX_i}$$

$$\leq e^{-tk\epsilon} \left( 1 + t^2 \right)^k \leq e^{-tk\epsilon + kt^2}$$
1. Proof of Chernoff: reduction to mgf

**Thm [Hoeffding, Chernoff].** If $X_1, \ldots, X_k$ are independent mean zero random variables with $|X_i| \leq 1$ then

$$\mathbb{P}\left[\left|\frac{1}{k} \sum_{i} X_i \right| \geq \epsilon \right] \leq 2\exp(-k\epsilon^2/4)$$

$$\mathbb{P}\left[\sum_{i} X_i \geq k\epsilon \right] \leq e^{-tk\epsilon} \mathbb{E} \exp\left(t \sum_{i} X_i\right) = e^{-tk\epsilon} \prod_{i} \mathbb{E}e^{tX_i}$$

$$\leq e^{-tk\epsilon} (1 + t^2)^k \leq e^{-tk\epsilon + kt^2} \leq e^{-\frac{k\epsilon^2}{4}}$$
2. Proof of Expander Chernoff

**Goal:** Show $\mathbb{E} \exp(t \sum_i f(v_i)) \leq \exp(c_\lambda kt^2)$
2. Proof of Expander Chernoff

**Goal:** Show $\mathbb{E} \exp(t \sum_i f(v_i)) \leq \exp(c_\lambda kt^2)$

**Issue:** $\mathbb{E} \exp(\sum_i f(v_i)) \neq \prod_i \mathbb{E} \exp(tf(v_i))$

How to control the mgf without independence?
Step 1: Write mgf as quadratic form

\[ \mathbb{E} e^t \sum_{i \leq k} f(v_i) \]

\[ = \sum_{i_0, \ldots, i_k \in V} \mathbb{P}(v_0 = i_0) P(i_0, i_1) \ldots P(i_{k-1}, i_k) \exp \left( t \sum_{1 \leq j \leq k} f(i_j) \right) \]
Step 1: Write mgf as quadratic form

$$\mathbb{E}e^t \sum_{i \leq k} f(v_i)$$

$$= \sum_{i_0, \ldots, i_k \in V} \mathbb{P}(v_0 = i_0) P(i_0, i_1) \cdots P(i_{k-1}, i_k) \exp \left( t \sum_{1 \leq j \leq k} f(i_j) \right)$$

$$= \frac{1}{n} \sum_{i_0, \ldots, i_k \in V} P(i_0, i_1) e^{tf(i_1)} \cdots P(i_{k-1}, i_k) e^{tf(i_k)}$$
Step 1: Write mgf as quadratic form

\[ \mathbb{E} e^{t \sum_{i \leq k} f(v_i)} \]
\[ = \sum_{i_0, \ldots, i_k \in V} \mathbb{P}(v_0 = i_0) P(i_0, i_1) \cdots P(i_{k-1}, i_k) \exp \left( t \sum_{1 \leq j \leq k} f(i_j) \right) \]
\[ = \frac{1}{n} \sum_{i_0, \ldots, i_k \in V} P(i_0, i_1) e^{tf(i_1)} \cdots P(i_{k-1}, i_k) e^{tf(i_k)} \]
\[ = \langle u, (EP)^k u \rangle \text{ where } u = \left( \frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}} \right). \]
Step 2: Bound quadratic form

**Goal:** Show $\langle u, (EP)^k u \rangle \leq \exp(c\lambda kt^2)$

**Observe:** $||P - J|| \leq \lambda$ where $J =$ complete graph with self loops
So for small $\lambda$, should have $\langle u, (EP)^k u \rangle \approx \langle u, (EJ)^k u \rangle$
Step 2: Bound quadratic form

**Goal:** Show $\langle u, (EP)^k u \rangle \leq \exp(c_\lambda kt^2)$

**Observe:** $||P - J|| \leq \lambda$ where $J =$ complete graph with self loops

So for small $\lambda$, should have $\langle u, (EP)^k u \rangle \approx \langle u, (EJ)^k u \rangle$

**Approach 1.** Use perturbation theory to show $||EP|| \leq \exp(c_\lambda t^2)$
Step 2: Bound quadratic form

**Goal:** Show \( \langle u, (EP)^k u \rangle \leq \exp(c_\lambda k t^2) \)

**Observe:** \( ||P - J|| \leq \lambda \) where \( J \) = complete graph with self loops
So for small \( \lambda \), should have \( \langle u, (EP)^k u \rangle \approx \langle u, (EJ)^k u \rangle \)

**Approach 1.** Use perturbation theory to show \( ||EP|| \leq \exp(c_\lambda t^2) \)

**Approach 2.** (Healy’08) track projection of iterates along \( u \).
Simplest case: $\lambda = 0$
Simplest case: $\lambda = 0$
Simplest case: $\lambda = 0$

$E\hat{P}\hat{u}$
Simplest case: $\lambda = 0$

$PEPu$
Simplest case: $\lambda = 0$

$EPEP\nu$
Observations

Observe: $P$ shrinks every vector orthogonal to $u$ by $\lambda$.

$$\langle u, Eu \rangle = \frac{1}{n} \sum_{v \in V} e^{tf(v)} = 1 + t^2$$

by mean zero condition.
A small dynamical system

\[ m_\perp = \|Q_\perp v\| \]

\[ v = (EP)^j u \]

\[ m_\parallel = \langle v, u \rangle \]
A small dynamical system

\[ m_\perp = \|Q_\perp v\| \]

\[ v = (EP)^j u \]

\[ m_\parallel = \langle v, u \rangle \]

\[ e^{t^2} \]

\[ \lambda = 0 \]

\[ v \sim EPv \]
A small dynamical system

\[ m_\perp = ||Q_\perp v|| \]

\[ v = (EP)^j u \]

\[ m_\parallel = \langle v, u \rangle \]

\[ \lambda t \]

\[ e^{t^2} \]

\[ e^{t\lambda} \]

any \( \lambda \)

\[ v \sim EPv \]
A small dynamical system

\[ m_\perp = \|Q_\perp v\| \]

\[ v = (EP)^j u \]

\[ m_\parallel = \langle v, u \rangle \]

\[ t \leq \log \frac{1}{\lambda} \]

\[ \lambda t \leq 1 \]

\[ v \sim EPv \]
A small dynamical system

\[ m_\perp = ||Q_\perp v|| \]

\[ v = (EP)^j u \]

\[ m_\parallel = \langle v, u \rangle \]

Any mass that leaves gets shrunk by \( \lambda \)
Analyzing dynamical system gives

**Thm[Gillman’94]:** Suppose $G = (V, E)$ is a regular graph with transition matrix $P$ which has second eigenvalue $\lambda$. Let $f : V \to \mathbb{R}$ be a function with $|f(v)| \leq 1$ and $\sum_v f(v) = 0$. Then, if $v_1, \ldots, v_k$ is a stationary random walk:

$$\mathbb{P} \left[ \left| \frac{1}{k} \sum_i f(v_i) \right| \geq \epsilon \right] \leq 2 \exp(-c(1 - \lambda)k\epsilon^2)$$
Generalization to Matrices?

**Setup:** \( f: V \rightarrow \mathbb{C}^{d \times d} \), random walk \( v_1, \ldots, v_k \).

**Goal:**

\[
\mathbb{E}Tr \left[ \exp \left( t \sum_i f(v_i) \right) \right] \leq d \cdot \exp(ckt^2)
\]

where \( e^A \) is defined as a power series.
Generalization to Matrices?

Setup: $f : V \rightarrow \mathbb{C}^{d \times d}$, random walk $v_1, \ldots, v_k$.

Goal:

$$\mathbb{E} \text{Tr} \left[ \exp \left( t \sum_i f(v_i) \right) \right] \leq d \cdot \exp(ckt^2)$$

where $e^A$ is defined as a power series.

Main Issue: $\exp(A + B) \neq \exp(A) \exp(B)$ unless $[A, B] = 0$

can’t express $\exp(\text{sum})$ as iterated product.
The Golden-Thompson Inequality

Partial Workaround [Golden-Thompson’65]:

\[ \text{Tr}(\exp(A + B)) \leq \text{Tr}(\exp(A) \exp(B)) \]

Sufficient for \textbf{independent} case by induction.
The Golden-Thompson Inequality

Partial Workaround [Golden-Thompson’65]:

\[ \text{Tr}(\exp(A + B)) \leq \text{Tr}(\exp(A) \exp(B)) \]

Sufficient for **independent** case by induction.

For **expander** case, need this for \( k \) matrices.
The Golden-Thompson Inequality

Partial Workaround [Golden-Thompson’65]:

$$Tr(\exp(A + B)) \leq Tr(\exp(A) \exp(B))$$

Sufficient for **independent** case by induction.
For **expander** case, need this for $k$ matrices. False!

$$Tr(e^{A+B+C}) > 0 > Tr(e^A e^B e^C)$$
[Sutter-Berta-Tomamichel’16] If $A_1, \ldots, A_k$ are Hermitian, then

$$
\log Tr(e^{A_1+ \ldots + A_k}) \leq \int d\beta(b) \log Tr \left[ \left( e^{\frac{A_1(1+ib)}{2}} \ldots e^{\frac{A_k(1+ib)}{2}} \right) \left( e^{\frac{A_1(1+ib)}{2}} \ldots e^{\frac{A_k(1+ib)}{2}} \right)^* \right]
$$

where $\beta(b)$ is an explicit probability density on $\mathbb{R}$. 

1. Matrix on RHS is always PSD.
2. Average-case inequality: $e^{A_{1/2}}$ are conjugated by unitaries.
3. Implies Lieb’s concavity, triple-matrix, ALT, and more.
Key Ingredient

[Sutter-Berta-Tomamichel’16] If $A_1, \ldots, A_k$ are Hermitian, then

$$\log \operatorname{Tr}(e^{A_1^+ + \cdots + A_k})$$

$$\leq \int d\beta(b) \log \operatorname{Tr} \left[ \left( e^{\frac{A_1(1+ib)}{2}} \cdots e^{\frac{A_k(1+ib)}{2}} \right) \left( e^{\frac{A_1(1+ib)}{2}} \cdots e^{\frac{A_k(1+ib)}{2}} \right)^* \right]$$

where $\beta(b)$ is an explicit probability density on $\mathbb{R}$.

1. Matrix on RHS is always PSD.
2. **Average-case** inequality: $e^{A_i/2}$ are conjugated by unitaries.
3. Implies Lieb’s concavity, triple-matrix, ALT, and more.
Proof of SBT: Lie-Trotter Formula

\[ e^{A+B+C} = \lim_{\theta \to 0^+} \left( e^{\theta A} e^{\theta B} e^{\theta C} \right)^{1/\theta} \]
Proof of SBT: Lie-Trotter Formula

\[ e^{A+B+C} = \lim_{\theta \to 0^+} \left( e^{\theta A} e^{\theta B} e^{\theta C} \right)^{1/\theta} \]

\[ \log \text{Tr} \ e^{A+B+C} = \lim_{\theta \to 0^+} 2 \log \| G(\theta) \|_{2/\theta / \theta} \]

For \( G(z) := e^{\frac{zA}{2}} e^{\frac{zB}{2}} e^{\frac{zC}{2}} \)
Complex Interpolation (Stein-Hirschman)

\[ \log \| G(\theta) \|_{2/\theta} \]
Complex Interpolation (Stein-Hirschman)

For each $\theta$, find analytic $F(z)$ st:

- $|F(it)| = 1$
- $|F(1 + it)| \leq \|G(1 + it)\|_2$
- $|F(\theta)| = \|G(\theta)\|_{2/\theta}$
Complex Interpolation (Stein-Hirschman)

For each $\theta$, find analytic $F(z)$ st:

\[
|F(it)| = 1
\]

\[
|F(1 + it)| \leq \|G(1 + it)\|_2
\]

\[
|F(\theta)| = \|G(\theta)\|_{2/\theta}
\]

\[
\log |F(\theta)| \leq \int \log |F(it)| + \int \log |F(1 + it)|
\]
Complex Interpolation (Stein-Hirschman)

For each $\theta$, find analytic $F(z)$ st:

- $|F(it)| = 1$
- $|F(1 + it)| \leq \|G(1 + it)\|_2$
- $|F(\theta)| = \|G(\theta)\|_{2/\theta}$
- $\lim_{\theta \to 0}$

$$\log|F(\theta)| \leq \int \log |F(it)| + \int \log |F(1 + it)|$$
[Sutter-Berta-Tomamichel’16] If $A_1, \ldots, A_k$ are Hermitian, then

$$\log \text{Tr} (e^{A_1^+ \ldots + A_k})$$

$$\leq \int d\beta(b) \log \text{Tr} \left[ \left( e^{\frac{A_1(1+ib)}{2}} \ldots e^{\frac{A_k(1+ib)}{2}} \right) \left( e^{\frac{A_1(1+ib)}{2}} \ldots e^{\frac{A_k(1+ib)}{2}} \right)^* \right]$$

where $\beta(b)$ is an explicit probability density on $\mathbb{R}$. 
Key Ingredient

[Sutter-Berta-Tomamichel’16] If $A_1, \ldots, A_k$ are Hermitian, then

$$\log \text{Tr}(e^{A_1+\ldots+A_k}) \leq \int d\beta(b) \log \text{Tr} \left[ \left( e^{\frac{A_1(1+ib)}{2}} \ldots e^{\frac{A_k(1+ib)}{2}} \right) \left( e^{\frac{A_1(1+ib)}{2}} \ldots e^{\frac{A_k(1+ib)}{2}} \right)^* \right]$$

where $\beta(b)$ is an explicit probability density on $\mathbb{R}$.

**Issue.** SBT involves integration over unbounded region, bad for Taylor expansion.
Bounded Modification of SBT

**Solution.** Prove bounded version of SBT by replacing strip with half-disk.

**[Thm]** If $A_1, \ldots, A_k$ are Hermitian, then

$$\log \text{Tr}(e^{A_1} + \ldots + e^{A_k}) \leq \int d\beta(b) \log \text{Tr} \left[ \left( e^{\frac{A_1e^{ib}}{2}} \ldots e^{\frac{A_ke^{ib}}{2}} \right) \left( e^{\frac{A_1e^{ib}}{2}} \ldots e^{\frac{A_ke^{ib}}{2}} \right)^* \right]$$

where $\beta(b)$ is an explicit probability density on $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

**Proof.** Analytic $F(z) +$ Poisson Kernel $+$ Riemann map.
Handling Two-sided Products

Issue. Two-sided rather than one-sided products:

$$\text{Tr} \left[ \left( e \frac{tf(v_1)e^{ib}}{2} \ldots e \frac{tf(v_k)e^{ib}}{2} \right) \left( e \frac{tf(v_1)e^{ib}}{2} \ldots e \frac{tf(v_k)e^{ib}}{2} \right)^* \right]$$
Handling Two-sided Products

Issue. Two-sided rather than one-sided products:

\[
Tr \left[ \left( \frac{tf(v_1)e^{ib}}{2} \ldots \frac{tf(v_k)e^{ib}}{2} \left( \frac{tf(v_1)e^{ib}}{2} \ldots \frac{tf(v_k)e^{ib}}{2} \right)^* \right) \right]
\]

Solution.
Encode as one-sided product by using \(Tr(AXB) = (A \otimes B^T)vec(X)\):

\[
\langle e \frac{tf(v_1)e^{ib}}{2} \otimes e \frac{tf(v_1)^*T e^{ib}}{2} \ldots e \frac{tf(v_k)^*T e^{-ib}}{2} \otimes e \frac{tf(v_k)^*T e^{-ib}}{2} \rangle vec(I_d), vec(I_d)
\]
Carry out a version of Healy’s argument with $P \otimes I_{d^2}$ and:

$$E = \begin{bmatrix} \frac{tf(1)e^{ib}}{2} \otimes \frac{tf(1)^*T e^{ib}}{2} \\ e \end{bmatrix} \ldots \begin{bmatrix} \frac{tf(n)e^{ib}}{2} \otimes \frac{tf(n)^*T e^{-ib}}{2} \\ e \end{bmatrix}$$

And $\text{vec}(I_d) \otimes u$ instead of $u$.

This leads to the additional $d$ factor.
Main Theorem

Thm. Suppose $G = (V, E)$ is a regular graph with transition matrix $P$ which has second eigenvalue $\lambda$. Let $f : V \to \mathbb{C}^{d \times d}$ be a function with $\|f(v)\| \leq 1$ and $\sum_v f(v) = 0$. Then, if $v_1, \ldots, v_k$ is a stationary random walk:

$$\mathbb{P}\left[\left\| \frac{1}{k} \sum_i f(v_i) \right\| \geq \epsilon \right] \leq 2d \exp(-c(1 - \lambda)k\epsilon^2)$$
Open Questions

Other matrix concentration inequalities
  (multiplicative, low-rank, moments)
Other Banach spaces
  (Schatten norms)
More applications of complex interpolation