On approximability of the Permanent of PSD matrices

Farzam Ebrahimnejad[1], Ansh Nagda[2], and Shayan Oveis Gharan[3]

[1]University of Washington, febrahim@cs.washington.edu
[2]UC Berkeley, anshnagda@gmail.com
[3]University of Washington, shayan@cs.washington.edu

April 16, 2024

Abstract

We study the complexity of approximating the permanent of a positive semidefinite matrix $A \in \mathbb{C}^{n\times n}$.

1. We design a new approximation algorithm for $\mathrm{per}(A)$ with approximation ratio $e^{(0.9999+\gamma)n}$, exponentially improving upon the current best bound of $e^{(1+\gamma-o(1))n}$ [Ana+17; YP22]. Here, $\gamma \approx 0.577$ is Euler's constant.
2. We prove that it is NP-hard to approximate $\mathrm{per}(A)$ within a factor $e^{(\gamma-\epsilon)n}$ for any $\epsilon > 0$. This is the first exponential hardness of approximation for this problem. Along the way, we prove optimal hardness of approximation results for the $\|\cdot\|_{2\to q}$ "norm" problem of a matrix for all $-1 < q < 2$.

1 Introduction

Given a matrix $A \in \mathbb{C}^{n\times n}$, the permanent of A is defined as

$$\mathrm{per}(A) = \sum_{\sigma\in\mathbb{S}_n} \prod_{i=1}^{n} A_{i,\sigma_i},$$

where the sum is over all permutations over n elements. It is well-known that the permanent of a matrix with non-negative entries can be approximated up to a $1+\epsilon$-multiplicative factor using the MCMC method [JSV04]. Recently there has been significant interest in studying permanent of Hermitian PSD matrices because of close connections to quantum optics and Boson sampling. A folklore algorithm is to simply take the product of the entries of the main diagonal to get an $n!$-approximation.

A few years ago, [Ana+17] obtained the first (deterministic) simply exponential approximation algorithms with approximation factor $e^{(\gamma+1)n}$. The algorithm proposed in [Ana+17] uses a basic SDP relaxation for the problem; many experts expected that perhaps by using higher-level SDP relaxations one can improve the approximation factor. Later on, several groups attempted to improve the approximation factor (see e.g., [Bar20]), but for the general case, only subexponential improvements to the approximation ratio were found [YP21; YP22]. Very recently, Meiburg showed that contrarily to the permanent of non-negative matrices, it is NP-Hard to approximate the permanent of a PSD matrix within a factor of $e^{-n^{1-\epsilon}}$ for any $\epsilon > 0$ [Mei23]. So, the MCMC method falls short of providing a $1+\epsilon$-approximation for PSD permanents.

It remained an open problem if, perhaps by using randomness or higher level SDP relaxations, one can obtain an $e^{-\epsilon n}$ approximation factor for ϵ arbitrarily small, or at the very least whether the $\gamma + 1$ factor in the exponent can be improved to a smaller constant. We answer both these questions in our work.

Our first result is an exponential improvement on the $e^{-(\gamma+1)n}$ approximation algorithm mentioned above.

Theorem 1.1 (Main Algorithmic Result)**.** *There is a deterministic polynomial time $e^{-(\gamma+0.9999)n}$-approximation algorithm for the permanent of a Hermitian PSD matrix $A \in \mathbb{C}^{n\times n}$.*

Our second result is the first exponential hardness of approximation for this problem. As a corollary of a general hardness of approximation result we prove (see Theorem 1.5 below), we show the following:

Theorem 1.2 (Main Hardness Result)**.** *For all $\epsilon > 0$, it is* NP*-hard to approximate the permanent of a Hermitian PSD matrix $A \in \mathbb{C}^{n\times n}$ within a factor $e^{-(\gamma-\epsilon)n}$.*

In particular, the above theorem shows that assuming NP $\neq$ RP *even using randomness* the approximation factor of [Ana+17; YP21] cannot be improved by more than a factor of e^n.

Maximizing Product of Linear Forms Our hardness techniques also apply to an optimization problem that happens to be related to the permanent of PSD matrices, called the "maximizing product of linear forms" problem, studied by Yuan and Parrilo [YP22; YP21]: Given a matrix $V \in \mathbb{C}^{n\times d}$ with rows $v_1, \ldots, v_n \in \mathbb{C}^d$, define

$$r(V) := \max_{x\in\mathbb{C}^n:\|x\|_2=1} \prod_{i=1}^{n} |\langle x, v_i\rangle|^2. \quad (1)$$

They design a polynomial time $O(e^{-\gamma n\cdot(1-o(1))})$-approximation algorithm for $r(V)$ using semidefinite programming, where $\gamma \approx 0.577$ is the Euler-Mascheroni constant. They also prove APX-hardness for this problem, and raise an open problem of finding the true approximability of this problem. Recently, Meiberg studied an equivalent problem under the name Approximate Quantum Maximum Likelihood Estimation, and showed NP-hardness of approximating it to within any constant factor [Mei23]. Our main technical hardness result, Theorem 1.5, immediately implies that the maximizing product of linear forms problem (and the Approximate Quantum Maximum Likelihood Estimation problem) does not admit a $e^{-\gamma n(1+\epsilon)}$-approximation for any constant $\epsilon > 0$, answering the question of Yuan and Parrilo up to sub-exponential factors in n.

1.1 Technical Contributions

1.1.1 Algorithmic results

In this part, we show the main ideas behind Theorem 1.1. We start by presenting algorithms used by previous work. Let $A = VV^\dagger$ be an $n \times n$ PSD matrix where $v_1, \ldots, v_n \in \mathbb{C}^n$ are the rows of V. Previous work ([Ana+17; YP22]) showed that the value of the following SDP gives a $e^{-(\gamma+1)n}$ approximation to $\mathrm{per}(A)$:

$$\mathrm{SDP}(V) := \max_{X\succeq 0, \mathrm{tr}(X)=n} \prod_{i\in[n]} v_i^\dagger X v_i.$$

$\mathrm{SDP}(V)$ might seem completely unrelated to the definition of $\mathrm{per}(A)$, but we remark that their relationship is a lot more clear when $\mathrm{per}(A)$ is rewritten using Wick's formula (Lemma 2.9). Notice that the objective function of $\mathrm{SDP}(V)$ is log-concave, so it can be optimized in polynomial time. It turns out that upon solving $\mathrm{SDP}(V)$, we can reduce to the case that the maximizer X^* of $\mathrm{SDP}(V)$ satisfies $v_i^\dagger X^* v_i = 1$ for all $i \in [n]$ (see Eq. (6)). This property simplifies matters enough that we will assume it for the rest of this section.

The above property implies that $A \preceq I$ (see Claim 3.1), so we immediately get $\mathrm{per}(A) \leq 1$. Conversely, [Ana+17; YP22] prove that

$$\mathrm{per}(A) \geq \frac{n!}{n^n} \cdot r(V) \geq e^{-n} \cdot r(V). \tag{2}$$

Noticing that $\mathrm{SDP}(V)$ is a semidefinite relaxation of $r(V)$, a simple Gaussian rounding argument ([YP22, Lemma 4.3], Lemma C.1) can be used to show

$$r(V) \geq e^{-\gamma n}\, \mathrm{SDP}(V) = e^{-\gamma n}. \tag{3}$$

Putting these together, one gets

$$e^{-(\gamma+1)n} \leq \mathrm{per}(A) \leq 1, \tag{4}$$

giving a $e^{-(\gamma+1)n}$ approximation factor.

We remark that both sides of the above inequality can be tight, in particular the upper bound is tight for the identity matrix and the lower bound is tight for a certain family of low rank projection matrices (see [Ana+17]). So one may expect that no improvement is possible along this line.

In our approach we exploit the fact these inequalities are tight for matrices of very different rank – the known tight examples for the upper/lower bounds have very high/low rank respectively. In order to make this intuition concrete, we will use $\mathrm{tr}(A)$ as a smooth analogue of rank. Our main technical results are improvements to both sides of Eq. (4).

Lemma 1.3 (Improved Upper Bound). *Let $\epsilon \in [0,1]$. Any matrix $0 \preceq A \preceq I$ with $\operatorname{tr}(A) \leq (1-\epsilon)n$ satisfies*

$$\operatorname{per}(A) \leq \left(1 - \frac{\epsilon^2}{20}\right)^n .$$

Lemma 1.4 (Improved Lower Bound). *Let $VV^\dagger = A \preceq I$, and assume that the maximizer X^* of $\mathrm{SDP}(V)$ satisfies $v_i^\dagger X^* v_i = 1$ for all i. For any $0 \leq \beta \leq 1$,*

$$\operatorname{per}(A) \geq e^{-n} \cdot r(V) \geq e^{-(\gamma+1)n} \cdot \exp\left(n \cdot \left(\ln(1-\beta) + \frac{\beta}{1-\beta} \cdot \frac{\operatorname{tr}(A)}{n} - \frac{0.273\beta^2}{(1-\beta)^2} \cdot \frac{n}{\operatorname{tr}(A)}\right)\right).$$

Our proof of Lemma 1.3 is inspired by an identity for $\operatorname{per}(A)$ appearing in [Bar20]. Our proof of Lemma 1.4 is based on an improved rounding procedure and is more technical, so we provide a proof overview in Section 1.2. We can now state our algorithm.

Algorithm: Given a PSD matrix $A = VV^\dagger$ where $v_1, \ldots, v_n$ are rows of V, first reduce to the case that the maximizer X^* of $\mathrm{SDP}(V)$ satisfies $v_i^\dagger X^* v_i = 1$ for all i (as described in Eq. (6)). Output $\left(1 - \frac{\epsilon^2}{20}\right)^n$, where ϵ is defined by $\operatorname{tr}(A) = (1-\epsilon)n$.

We will use this algorithm in our proof of Theorem 1.1, which is straightforward to analyze when equipped with Lemmas 1.3 and 1.4.

1.1.2 Hardness results

In this part we highlight the main technical contributions behind Theorem 1.2. Our proof broadly consists of two steps:

1. Show that $r(V)$ does not admit a $e^{-\gamma n(1+\epsilon)}$-approximation algorithm.
2. Give an approximation-preserving reduction from $r(V)$ to PSD permanents.

We start by elaborating on Item 1. In order to draw analogies to the existing hardness of approximation literature, we will first rephrase and generalize the optimization problem of $r(V)$. Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ be a field. For a vector $x \in \mathbb{F}^n$ and $p \in \mathbb{R} - \{0\}$, define

$$\|x\|_p = \left(\mathbb{E}_i |x_i|^p\right)^{1/p}.$$

We will be particularly interested in the case that $p = 0$, for which we define $\|x\|_0 = \lim_{p\to 0} \|x\|_p$. It is not too hard to see that for any vector x, $\|x\|_0$ equals $\prod_{i\in[n]} |x_i|^{1/n}$, the geometric mean of the magnitude of the entries of x. Note that in the case that $p < 1$, $\|\cdot\|_p$ is not a norm and not convex, but we will nevertheless refer to it as the p-norm.

Given a matrix $A \in \mathbb{F}^{m\times n}$, the $p \to q$ "norm" of A is defined as

$$\|A\|_{p\to q} = \max_{x\in\mathbb{F}^n : \|x\|_p = 1} \|Ax\|_q .$$

The connection of $\|A\|_{p\to q}$ to $r(V)$ is apparent: for any matrix $V \in \mathbb{C}^{n\times d}$, we have

$$r(V) = \|V\|_{2\to 0}^{2n}.$$

Over the last decade there has been significant interest in designing approximation algorithms or proving hardness of approximation for matrix $p \to q$ norms for $p, q \geq 1$ [Bar+12; BRS15; Bha+23]. Most notably, the $2 \to 4$ norm has been shown to be closely related to the Unique games and the small set expansion conjectures [Bar+12]. To the best of our knowledge, the problem is not well-studied when $q < 1$. We prove tight hardness of approximation (assuming P $\neq$ NP) for the $2 \to q$ norm when $-1 < q < 2$.

For $q > -1$ let

$$\gamma_{\mathbb{F},q} = \mathbb{E}_{g \sim \mathbb{F}\mathcal{N}(0,1)}[|g|^q]^{1/q}$$

be the q-norm of a standard (real/complex) normal random variable. Bhattiprolu, Ghosh, Guruswami, Lee, Tulsiani [Bha+23] showed that for any $1 \leq q < 2$, and any $\epsilon > 0$ it is NP-hard to approximate the $2 \to q$ norm of a real $m \times n$ matrix better than $\gamma_{\mathbb{R},p} + \epsilon$, matching known semidefinite relaxation-based approximation algorithms [Ste05]. In our main theorem, we build on their techniques and we extend their result to all $-1 < q < 2$.

Theorem 1.5 (Main Technical Hardness Theorem). *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. For all $-1 < q < 2$ and $\epsilon > 0$, it is* NP*-hard to approximate $\|A\|_{2\to q}$ given a matrix $A \in \mathbb{F}^{m\times n}$ within a factor of $\gamma_{\mathbb{F},q} + \epsilon$.*

For the sake of completeness, in Appendix C we write down a semidefinite relaxation of $\|A\|_{2\to q}$ for all $-1 < q < 2$ and prove that it gives a $\gamma_{\mathbb{F},q}$-approximation to $\|A\|_{2\to q}$, matching the above hardness result. As $r(V) = \|V\|_{2\to 0}^{2n}$, we also get that it is NP-hard to approximate $r(V)$ within a factor of $(\gamma_{\mathbb{C},0} + \epsilon)^{2n} = e^{-\gamma n(1+\epsilon)}$.

Next, we elaborate on Item 2 – an approximation preserving reduction from $r(V)$ to $\operatorname{per}(A)$. Our main observation is that the permanent of a highly rank-deficient $n \times n$ PSD matrix $A = VV^\dagger$ is essentially (up to subexponential error) the same as $r(V)$. This is a consequence of Wick's formula (Lemma 2.9), which allows us to view the permanent of a PSD matrix as a squared absolute moment of a complex multivariate Gaussian. As a result, we are able to use Theorem 1.5 to prove Theorem 1.2, which we do in Section 4.2.

1.2 Overview of the proof of Lemma 1.4

Let us start by explaining the proof of Eq. (3), which Lemma 1.4 improves upon. For any distribution $\mathcal{D}$ over $\mathbb{C}^n$ with $\mathbb{E}[\|x\|_2^2] = 1$, we have the bound

$$r(V) = \max_{\|x\|_2=1} \prod_{i\in[n]} |\langle v_i, x\rangle|^2 \geq \left(\frac{\mathbb{E}_{x\sim\mathcal{D}} \prod_{i\in[n]} |\langle v_i, x\rangle|^{2/n}}{\mathbb{E}_{x\sim\mathcal{D}} \|x\|_2^2} \right)^n = \mathbb{E}_{x\sim\mathcal{D}} \left[\prod_{i\in[n]} |\langle v_i, x\rangle|^{2/n} \right]^n .$$

Using Jensen's inequality on the RHS, we get

$$r(V) \geq \exp\left(\sum_{i\in[n]} \mathbb{E}_{x\sim\mathcal{D}} \ln |\langle v_i, x\rangle|^2 \right) . \tag{5}$$

The basic Gaussian rounding scheme picks $x \sim \mathcal{D} = \mathbb{C}\mathcal{N}(0, X^*)$ (see Definition 4 for a definition of the complex Gaussian distribution). Notice that for each i, $\langle v_i, x\rangle \sim \mathbb{C}\mathcal{N}(0, v_i^\dagger X^* v_i) = \mathbb{C}\mathcal{N}(0, 1)$ by assumption of Lemma 1.4. Since $\mathbb{E}_{y\sim\mathbb{C}\mathcal{N}(0,1)} \ln |y|^2 = -\gamma$, we immediately get $r(V) \geq \exp(-n\gamma)$.

One can see that the analysis (in particular, the application of Jensen's inequality) is tight if $A = V = X^* = I$ for example. So to improve on Eq. (5), we must use a different rounding algorithm. Our first observation is that in the special case that $A = V = I$ where $v_i = e_i$, we can get the optimal lower bound by sampling independent Rademachers $s_1, \ldots, s_n \sim \{\pm 1\}$, and setting $x = \sum_{i\in[n]} s_i v_i$. With this choice, $\mathbb{E}_x \ln |\langle v_i, x\rangle|^2 = \mathbb{E}_x \ln 1 = 0$, implying $r(V) \geq 1$.

One could try to use a similar rounding scheme in the more general case of Lemma 1.4, i.e., when A is close to I in the sense that $\operatorname{tr}(A) \geq (1-\epsilon)n$. Unfortunately this strategy ends up failing, as when the v_i's are not exactly orthogonal, there could be a nonzero probability that $\langle v_i, x\rangle = 0$, which would imply $\mathbb{E}_x \ln |\langle v_i, x\rangle|^2 = -\infty$. Note that if A is close to I, this singularity is a very small probability event for most of the vectors v_i, so it is natural to try to avoid it by adding some noise to x. We do this by interpolating between the two rounding schemes. We pick a parameter $0 < \beta < 1$, and set $x = \sqrt{1-\beta} g + \sqrt{\beta} \sum_{i\in[n]} s_i v_i$, up to some normalization, where $g \sim \mathbb{C}\mathcal{N}(0, X^*)$.

On the technical side, this interpolation helps us analyze $\mathbb{E}_x \ln |\langle v_i, x\rangle|^2$ in terms of tractable quantities. We use a sharp bound on the expected log of the magnitude squared of a noncentral complex Gaussian (see Lemma 2.7): for any $c \in \mathbb{C}$,

$$\mathbb{E}_{g\sim\mathbb{C}\mathcal{N}(0,1)} \ln |g + c|^2 \geq -\gamma + |c|^2 - \frac{|c|^4}{4}.$$

As a result of this inequality, when β is bounded away from 1, we can effectively bound $\mathbb{E}_x \ln |\langle v_i, x\rangle|^2$ using only the second and fourth moments of the random variable $\sum_{i\in[n]} s_i v_i$, which are tractable.

1.3 Overview of the proof of Theorem 1.5

As alluded to in Section 1.1.2, prior to our work, optimal hardness results for the $2 \to q$ norm are already established for $q \geq 1$. Our first observation is that these results [Gur+16; BRS15; Bha+23] can be extended to all $-1 < q < 2$ (see Theorem 4.1), or even more generally, to 2-concave f-means (see Definition 2).

In particular, one can deduce the following theorem.

Theorem 1.6 (Informal version of Theorem 4.1)**.** *Let $q < 2$. Assume that there is a family $\{E_k\}$ of $k \times d_k$ gadget matrices such that $\|E_k\|_{2\to q} = 1$, but for all "smooth" unit vectors x ($\|x\|_\infty \ll 1$), $\|E_k x\|_q \leq \gamma$. Then for all $\epsilon > 0$, it is NP-Hard to distinguish between the following two cases given a matrix $A : \mathbb{C}^m \to \mathbb{C}^n$ with $\|A\|_{2\to 2} \leq 1$.*

1. *Completeness: $\|A\|_{2\to q} = 1$, or*
2. *Soundness: $\|A\|_{2\to q} \leq \gamma + \epsilon$.*

It remains to construct an appropriate family of gadget matrices $\{E_k\}$. We will use the following family, which was suggested in [BRS15]. For $k \geq 1$, define $E_k^{(\mathbb{C})} \in \mathbb{C}^{4^k \times k}$ as the matrix whose rows consist of the members of $\frac{1}{\sqrt{k}} \cdot \{-1, +1, -i, +i\}^k$ ordered arbitrarily.

It remains to show that the matrices $E_k^{(\mathbb{C})}$ satisfy the requirements of Theorem 1.6 with $\gamma \approx \gamma_{\mathbb{C},q}$. By construction, $\left\|E_k^{(\mathbb{C})}\right\|_{2\to q} = 1$.

To prove this, in Lemma B.1 we prove a Berry-Esseen type result for test functions of the form $|x|^q$ for $q \neq 0$ and $\log |x|$ otherwise, applied to a sum of independent random variables. In particular, the special case of interest to us (for Theorem 1.2) is $q = 0$. In that case, the test function is $\log|x|$ which has a singularity at $x = 0$, but we are nevertheless one can bound the right hand side of the lemma below by an arbitrarily small quantity as $\delta \to 0$.

Lemma 1.7 (Informal version of Lemma 4.2)**.** *Let $-1 < q < 2$ and $0 < \delta < 1$. For all "smooth" unit vectors x with $\|x\|_\infty \leq \delta$,*

$$f(\|E_k x\|_q) - \gamma_{\mathbb{C},q} \lesssim -\int_0^\delta \min(0, f(u)) + \delta \cdot \left(\max(0, f(2\sqrt{\log(1/\delta)})) + 2f'(1)\right),$$

where $f(x) = |x|^q$ for $q \neq 0$, and $f(x) = \log|x|$ for $q = 0$.

1.4 Future Directions

The most exciting open problem is to determine the correct approximability for PSD permanents. Improving the hardness result seems to be out of reach of current techniques, but our ideas provide a clear path to improving the algorithmic result. Any significant improvement to the algorithm along the lines of our ideas would require significantly better versions of Lemmas 1.3 and 1.4. In particular, Lemma 1.3 is currently the bottleneck to a better approximation ratio, specifically the $O(\epsilon^2)$ dependence. We conjecture that it can be improved to $O(\epsilon)$, which would yield better approximation ratios as a corollary.

Although we don't have concrete new applications of hardness of approximation of $\|A\|_{2\to q}$ for $q \neq 0$, we expect to find further applications of the machinery developed here in addressing counting and optimization of linear algebraic problems, e.g., in estimating mixed discriminant, sub-determinant maximization, Nash-welfare maximization, etc.

1.5 Paper Organization

In Section 2, we present preliminary definitions and results that we will use. In Section 3, we prove Theorem 1.1. In Section 4, we prove Theorems 1.2 and 1.5 (with some components of the proof appearing in Appendices A and B).

2 Preliminaries

2.1 Generalized Means and Norms

Although we mostly use p-norms that are defined using an expectation, it will be convenient to also define the counting version of 2-norm, which we denote as ℓ_2.

Definition 1 (ℓ_2-norm, Frobenius norm)**.** Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. For a vector $x \in \mathbb{F}^n$, define $\|x\|_{\ell_2} = \sqrt{\sum_{i\in[n]} |x_i|^2}$. For a matrix $A \in \mathbb{F}^{m\times n}$, define $\|A\|_F = \sqrt{\sum_{i,j\in[n]} |A_{i,j}|^2}$.

We work with a generalization of $\|\cdot\|_p$ using the framework of "f-means".

Definition 2. Let $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be continuous and injective. Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ be a field. For a vector $x \in \mathbb{F}^n$, define

$$[x]_f^f := \mathbb{E}_{i\sim[n]} f(|x_i|),$$

$$[x]_f := f^{-1}([x]_f^f) = f^{-1}(\mathbb{E}_{i\sim[n]} f(|x_i|)).$$

We will refer to $[x]_f$ as the f-*mean* of x. More generally, for a random variable X over $\mathbb{F}$ define

$$[X]_f^f = \mathbb{E}f(|X|), \quad [X]_f = f^{-1}(\mathbb{E}f(|X|)).$$

For a matrix $A \in \mathbb{F}^{n\times d}$, define

$$[A]_{2\to f} := \max_{x\in\mathbb{F}^n} \frac{[Ax]_f}{\|x\|_2}.$$

f-means provide a convenient and unified way to talk about $\|\cdot\|_p$, even for the case of $p = 0$.

Observation 1 (Power Means). *For all $p \in \mathbb{R}$, define*

$$f_p(x) = \begin{cases} x^p & \text{if } p > 0, \\ \log x & \text{if } p = 0, \\ -x^p & \text{if } p < 0. \end{cases}$$

Then, we have $[x]_{f_p} = \|x\|_p$ for any $p \in \mathbb{R}$.

We note that the observation would still hold if we used x^p instead of $-x^p$ in the third case, but it will be convenient for f_p to always be an increasing function.

We will mostly be concerned with f-means that are dominated by the 2-norm. This happens exactly when f is 2-concave:

Definition 3. A function $f : \mathbb{R}_{>0} \to \mathbb{R}$ is 2-concave if $x \to f(\sqrt{x})$ is concave.

Example 1. For any $p \leq 2$, f_p is 2-concave.

We make some useful observations about 2-concave functions.

Claim 2.1. *Let $f : \mathbb{R}_{>0} \to \mathbb{R}$ be a 2-concave increasing function. Then,*

1. *$[x]_f \leq \|x\|_2$ for any vector x. Equivalently, $[x]_f^f \leq f(\|x\|_2)$.*
2. *$f'(x) \leq \frac{f'(y)}{y} \cdot x$ for $0 < y \leq x$.*
3. *$f(x) - f(y) \leq \frac{f'(y)}{2y} \cdot x^2$ for $0 < y \leq x$.*

Proof.

1. Using Jensen's inequality on the concave function $x \to f(\sqrt{x})$,

$$[x]_f^f = \mathbb{E}f(|x_i|) = \mathbb{E}f\left(\sqrt{|x_i|^2}\right) \leq f(\sqrt{\mathbb{E}|x_i|^2}) = f(\|x\|_2).$$

2. Follows from the fact that $f'(\sqrt{x}) = \frac{f(\sqrt{x})}{2\sqrt{x}}$ is increasing in x.
3. Integrating the above from y to x,

$$f(x) - f(y) \leq \frac{f'(y)}{y} \cdot \frac{(x^2 - y^2)}{2} \leq \frac{f'(y)}{2y} \cdot x^2.$$

□

One could ask when $[x]_f$ is a homogeneous function of x. It turns out that this is exactly when $[x]_f$ is a p-norm.

Lemma 2.2 ([HLP52])**.** *$[\cdot]_f$ is 1-homogeneous (that is, $[ax]_f = a[x]_f$ for all scalars a and vectors x) if and only if it equals $[\cdot]_{f_p} = \|\cdot\|_p$ for some $p \in \mathbb{R}$.*

We will require another simple fact about f-means.

Fact 2.3. *Let V be a matrix $V \in \mathbb{F}^{n\times d}$ and integer $k > 0$. Then,*

$$[V^{(k)}]_{2\to f} := \left[\begin{bmatrix} V \\ \vdots \\ V \end{bmatrix}\right]_{2\to f} = [V]_{2\to f},$$

where V is copied k times in the right hand side.

Proof. For any vector $x \in \mathbb{F}^d$, consider the two vectors $z = Vx$ and $z^{(k)} = V^{(k)}x$. Note that a uniformly random entry of z has the same distribution as a random entry of $z^{(k)}$, so $[z]_f = [z^{(k)}]_f$. The claim follows from the definition of $[\cdot]_{2\to f}$. □

2.2 Gaussians

We will consider both real and complex Gaussians.

Definition 4. Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. For the $n \times n$ identity matrix I_n, $\mathbb{F}\mathcal{N}(0, I_n)$ is defined to be the distribution over vectors $x \in \mathbb{F}^n$ given by the density function

$$p_{\mathbb{F}}(x) = \begin{cases} (2\pi)^{-n/2} \cdot \exp\left(-\|x\|_{\ell_2}^2/2\right) & \text{if } \mathbb{F} = \mathbb{R}, \\ \pi^{-n} \exp(-\|x\|_{\ell_2}^2) & \text{if } \mathbb{F} = \mathbb{C}. \end{cases}$$

More generally, given a positive semidefinite covariance matrix $\Sigma = AA^\dagger$ for $A \in \mathbb{F}^{n\times d}$, define $\mathbb{F}\mathcal{N}(0, \Sigma)$ to be distributed as Ax, where $x \sim \mathbb{F}\mathcal{N}(0, I_d)$. We will sometimes use $\mathcal{N}$ to denote $\mathbb{R}\mathcal{N}$.

More concretely, a complex Gaussian $g \sim \mathbb{C}\mathcal{N}(0, 1)$ can be sampled by sampling its real and imaginary parts independently from $\mathcal{N}(0, 1/2)$. There are formulas for the moments of univariate real and complex Gaussians in terms of the Gamma function.

Definition 5. For any $p \in \mathbb{R}$ and $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, define $\gamma_{\mathbb{F},p} = [g]_{f_p}$, where g is a random variable distributed as $\mathbb{F}\mathcal{N}(0, 1)$.

Fact 2.4. *For any $p \in (-1, \infty) - \{0\}$,*

$$\gamma_{\mathbb{R},p}^p = \mathbb{E}_{g\sim\mathcal{N}(0,1)}[|g|^p] = \frac{2^{p/2} \cdot \Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi}},$$

and for any $p \in (-2, \infty) - \{0\}$,

$$\gamma_{\mathbb{C},p}^p = \mathbb{E}_{g\sim CN(0,1)}[|g|^p] = \Gamma\left(\frac{p}{2} + 1\right).$$

In particular, this implies

$$\gamma_{\mathbb{R},1} = \sqrt{\frac{2}{\pi}}, \quad \gamma_{\mathbb{C},1} = \sqrt{\frac{\pi}{2}}, \quad \gamma_{\mathbb{R},0} = \lim_{p\to 0} \gamma_{\mathbb{R},p} = \sqrt{\frac{e^{-\gamma}}{2}}, \quad \gamma_{\mathbb{C},0} = \lim_{p\to 0} \gamma_{\mathbb{C},p} = \sqrt{e^{-\gamma}}.$$

Fact 2.5 (Moment Generating Function of $|g|^2$)**.** *Let $g \sim \mathbb{C}\mathcal{N}(0,1)$. For any $t < 1$, $\mathbb{E}[e^{t|g|^2}] = (1-t)^{-1}$.*

We will need sharp bounds on the expected value of $\ln|g+c|^2$ for Gaussian g and fixed c. First, we prove an estimate on the exponential integral function.

Fact 2.6. *For $x \geq 0$ it holds that*

$$\mathrm{Ei}(-x) = \int_{-\infty}^{-x} \frac{e^t}{t} dt = \gamma + \ln(x) + \sum_{k=1}^{\infty} \frac{(-x)^k}{k \cdot k!} \leq \gamma + \ln(x) - x + \frac{x^2}{4}.$$

Proof. The identity is due to Equation 5.1.11 in [AS48]. For the inequality, we must show that the function

$$f(x) = \sum_{k=3}^{\infty} \frac{(-x)^k}{k \cdot k!}$$

is nonpositive for $x \geq 0$. To do this, observe first that $f(0) = 0$, and

$$\begin{aligned} f'(x) &= \sum_{k=3}^{\infty} \frac{(-1)^k \cdot x^{k-1}}{k!} \\ &= \frac{1}{x} \sum_{k=3}^{\infty} \frac{(-x)^k}{k!} \\ &= \frac{e^{-x} - \left(1 - x + \frac{x^2}{2}\right)}{x} \\ &\leq 0. && (e^{-x} \leq 1 - x + x^2/2 \text{ for } x \geq 0) \end{aligned}$$

Therefore, $f(x) \leq 0$ for $x \geq 0$. □

Lemma 2.7. *Let $c \in \mathbb{C}$. Then, $\mathbb{E}_{g\sim\mathbb{C}\mathcal{N}(0,1)}[\ln|g+c|^2] \geq -\gamma + |c|^2 - |c|^4/4$.*

Proof. Define $x = |c|^2$. By [Mos20, Eqn. 35, Thm. 1] we have the identity

$$\mathbb{E}_{g\sim\mathbb{C}\mathcal{N}(0,1)}[\ln|g+c|^2] = -\ln(x) - \mathrm{Ei}(x).$$

By Fact 2.6, this is at least $-\gamma + x - x^2/4$, as desired. □

2.3 Permanent

For a matrix $A \in \mathbb{C}^{n\times n}$, its permanent is defined as

$$\mathrm{per}(A) := \sum_{\sigma \in S_n} \prod_{i=1}^{n} A_{i,\sigma(i)}.$$

On the domain of positive semidefinite matrices, the permanent has some nice properties. For example, it is monotone w.r.t. the Loewner order.

Lemma 2.8 (e.g., [Ana+17])**.** *If $A, B \in \mathbb{C}^{n\times n}$ are hermitian and $A \succeq B \succeq 0$, then*

$$\mathrm{per}(A) \geq \mathrm{per}(B).$$

Proof Sketch. The statement of the lemma follows, because $A \succeq B \succeq 0$ implies that $A^{\otimes n} \succeq B^{\otimes n} \succeq 0$. So, if 1_{S_n} is the indicator vector of all permutations in $\mathbb{R}^{n^{\otimes n}}$,

$$\mathrm{per}(A) = \frac{1}{n!} 1_{S_n}^{\dagger} A^{\otimes n} 1_{S_n} \geq \frac{1}{n!} 1_{S_n}^{\dagger} B^{\otimes n} 1_{S_n} = \mathrm{per}(B)$$

as desired. □

Lemma 2.9. *For any PSD matrix $VV^{\dagger}$ with $V \in \mathbb{R}^{n \times d}$, we have*

$$\mathbb{E}_{x \sim \mathcal{N}(0,I)} \left[\prod_{i \in [n]} |\langle v_i, x \rangle|^2 \right] = c_{n,d}^{\mathbb{R}} \cdot \mathbb{E}_{x \in \mathbb{R}^d, \|x\|_2 = 1} \left[\prod_{i \in [n]} |\langle v_i, x \rangle|^2 \right].$$

For any PSD matrix $VV^{\dagger}$ with $V \in \mathbb{C}^{n \times d}$, we have

$$\mathrm{per}(VV^{\dagger}) = \mathbb{E}_{x \sim \mathbb{C}\mathcal{N}(0,I)} \left[\prod_{i \in [n]} |\langle v_i, x \rangle|^2 \right] = c_{n,d}^{\mathbb{C}} \cdot \mathbb{E}_{x \in \mathbb{C}^d, \|x\|_2 = 1} \left[\prod_{i \in [n]} |\langle v_i, x \rangle|^2 \right].$$

Here, $v_1, \ldots, v_n$ are the rows of V. The proportionality constants above are defined by

$$c_{n,d}^{\mathbb{R}} = \frac{\Gamma(n + d/2)}{\Gamma(d/2) \cdot (d/2)^n}, \qquad c_{n,d}^{\mathbb{C}} = \frac{(d+n-1)!}{(d-1)! \cdot d^n}.$$

Proof. The first equality in the second conclusion follows from Isserlis' theorem/Wick's formula (see 3.1.4 in [Bar16]).

We prove the other equality in the real case, but the complex case can be proved similarly. Observe that

$$\begin{aligned}
\mathbb{E}_{x \sim \mathcal{N}(0,I)} \left[\prod_{i \in [n]} |\langle v_i, x \rangle|^2 \right] &= \mathbb{E}_{x \sim \mathcal{N}(0,I)} \|x\|_2^{2n} \cdot \left[\prod_{i \in [n]} \left| \left\langle v_i, \frac{x}{\|x\|_2} \right\rangle \right|^2 \right] \\
&= \mathbb{E}_{x \sim \mathcal{N}(0,I)} \|x\|_2^{2n} \cdot \mathbb{E}_{x \in \mathbb{R}^n, \|x\|_2 = 1} \left[\prod_{i \in [n]} |\langle v_i, x \rangle|^2 \right] \\
&= d^{-n} \mathbb{E}_{x \sim \mathcal{N}(0,I)} \|x\|_{\ell_2}^{2n} \cdot \mathbb{E}_{x \in \mathbb{R}^n, \|x\|_2 = 1} \left[\prod_{i \in [n]} |\langle v_i, x \rangle|^2 \right]
\end{aligned}$$

The first identity uses that for $x \sim \mathcal{N}(0, I)$, $\|x\|_2$ is independent from $x/\|x\|_2$. The second identity uses that fact that if $x \sim \mathcal{N}(0, I)$, then $x/\|x\|_2$ is distributed uniformly on a sphere of radius $\|x\|_2$. To conclude the proof, notice that $\mathbb{E}_{x \sim \mathcal{N}(0,I)} \|x\|_{\ell_2}^{2n}$ is the n^{th} moment of a chi-squared random variable with d-degrees of freedom, which is $2^n \frac{\Gamma(n+d/2)}{\Gamma(d/2)}$. □

We will also require the following formula for the permanent of the sum of two matrices.

Lemma 2.10 ([Per12, Page 2]). *For any two matrices $A, B \in \mathbb{C}^{n\times n}$,*

$$\operatorname{per}(A+B) = \sum_{S,T\subseteq[n],|S|=|T|} \operatorname{per}(A_{S,T}) \cdot \operatorname{per}(B_{\bar{S},\bar{T}}).$$

When $A = I$, this simplifies to

$$\operatorname{per}(I+B) = \sum_{S\subseteq[n]} \operatorname{per}(B_{S,S}).$$

Above, $A_{S,T}$ is the $|S| \times |T|$ submatrix of A containing rows only in S and columns only in T.

3 Algorithm

We start by expanding on the basic setup of the algorithms of [Ana+17; YP22], which we briefly introduced in Section 1.1.1. After this, we will show how Lemmas 1.3 and 1.4 imply Theorem 1.1. Later on, in Sections 3.1 and 3.2 respectively, we prove Lemmas 1.3 and 1.4.

Let $A = VV^\dagger$ be the PSD matrix whose permanent we wish to compute. Let $v_1, \dots, v_n \in \mathbb{C}^n$ be the rows of V, so by Lemma 2.9,

$$\operatorname{per}(A) = \mathbb{E}_{x\sim\mathbb{C}\mathcal{N}(0,I)}\left[\prod_{i\in[n]} |\langle x, v_i\rangle|^2\right].$$

Recall the log-concave maximization problem $\operatorname{SDP}(V)$ we associated with this problem:

$$\operatorname{SDP}(V) = \max_{X:X\succeq 0,\operatorname{tr}(X)=n} \prod_{i\in[n]} v_i^\dagger X v_i.$$

Let X^* be the optimal solution to $\operatorname{SDP}(V)$. Note that X^* can be found efficiently. It will be convenient to make a simplification to our problem. We will replace the matrix A by $\tilde{A} = D^{-1/2}AD^{-1/2}$, where D is a positive semidefinite diagonal matrix defined as $D_{i,i} = v_i^\dagger X^* v_i$. Since D is diagonal,

$$\operatorname{per}(A) = \operatorname{per}(\tilde{A}) \cdot \operatorname{per}(D) = \operatorname{per}(\tilde{A}) \cdot \operatorname{SDP}(V),$$

so it suffices to approximate $\operatorname{per}(\tilde{A})$ instead of $\operatorname{per}(A)$. Writing $\tilde{A} = \tilde{V}\tilde{V}^\dagger$ for $\tilde{V} = D^{-1/2}V$, we can see that the objective functions of $\operatorname{SDP}(V)$ and $\operatorname{SDP}(\tilde{V})$ are positive scalar multiples of each other, so $\operatorname{SDP}(\tilde{V})$ is also maximized by X^*. Note that $\tilde{A}$ enjoys the additional property $\tilde{v}_i^\dagger X^* \tilde{v}_i = 1$ for all $i \in [n]$, where $\tilde{v}_i = D^{-1/2} v_i$. Replacing A by $\tilde{A}$, we will henceforth assume that the maximizer X^* of $\operatorname{SDP}(V)$ satisfies

$$v_i^\dagger X^* v_i = 1 \text{ for all } i \in [n]. \tag{6}$$

In particular, this implies $\operatorname{SDP}(V) = 1$. Under this assumption, A satisfies an important property.

Claim 3.1. *We have $A \preceq I$.*

Proof. Let $f(X) = \prod_{i\in[n]} v_i^\dagger X v_i$ be the objective function of $\operatorname{SDP}(V)$. We can compute

$$\nabla(\ln f)(X) = \sum_{i\in[n]} \frac{v_i v_i^\dagger}{v_i^\dagger X v_i}.$$

In particular, by Eq. (6), $\nabla(\ln f)(X^*) = \sum_{i\in[n]} v_i v_i^\dagger = V^\dagger V$.

The optimality conditions for X^* imply that for all symmetric matrices M with $\mathrm{tr}(M) = 0$ and $W_- \subseteq \mathrm{Range}(X^*)$ it holds that

$$\langle V^\dagger V, M\rangle = \langle \nabla(\ln f)(X^*), M\rangle \leq 0.$$

Here, W_- denotes the vector space spanned by the negative eigenvectors of M. Now, let $Q \succeq 0$ be an arbitrary PSD matrix, and set $M = Q - \frac{\mathrm{tr}(Q)}{n} X^*$. M satisfies both the conditions above, and therefore we have

$$\begin{aligned}
0 &\geq \langle V^\dagger V, M\rangle \\
&= \langle V^\dagger V, Q\rangle - \frac{\mathrm{tr}(Q)}{n}\langle V^\dagger V, X^*\rangle \\
&= \langle V^\dagger V, Q\rangle - \frac{\mathrm{tr}(Q)}{n}\sum_{i\in n} v_i^\dagger X^* v_i \\
&= \langle V^\dagger V, Q\rangle - \mathrm{tr}(Q). && \text{(by Eq. (6))}
\end{aligned}$$

In other words, $\langle V^\dagger V, Q\rangle \leq \mathrm{tr}(Q)$ for all $Q \succeq 0$, implying $V^\dagger V \preceq I$. Therefore $A = VV^\dagger \preceq I$. □

Claim 3.1 immediately implies $\mathrm{per}(A) \leq 1$. In [Ana+17; YP22], the authors prove the complimentary inequalities

$$\mathrm{per}(A) \geq \frac{n!}{n^n}\cdot r(V) \geq \exp(-\gamma n)\cdot\frac{n!}{n^n}\cdot \mathrm{SDP}(V) = \exp(-\gamma n)\cdot \frac{n!}{n^n} \gtrsim \exp(-(\gamma+1)n), \tag{7}$$

and together, the two inequalities above provide a $e^{-(1+\gamma)n}$ approximation for $\mathrm{per}(A)$.

Recall Lemmas 1.3 and 1.4, which (under Claim 3.1) improve the above inequalities to

$$e^{-(\gamma+1)n}\cdot \exp\left(n\cdot \ell\left(\frac{\mathrm{tr}(A)}{n}\right)\right) \leq \mathrm{per}(A) \leq \exp\left(n\cdot r\left(\frac{\mathrm{tr}(A)}{n}\right)\right). \tag{8}$$

Here, $\ell(x) = \max_{0\leq\beta\leq 1} \ln(1-\beta) + \frac{\beta x}{(1-\beta)} - \frac{0.273\beta^2}{(1-\beta)^2 x}$, and $r(x) = \ln\left(1 - \frac{(1-x)^2}{20}\right)$. We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. Let $A = VV^\dagger \succeq 0$, where V has rows $v_1, \ldots, v_n$. Our algorithm will first solve $\mathrm{SDP}(V)$ and use Eq. (6) and Claim 3.1 to reduce to the case that $0 \preceq A \preceq I$ and $v_i^\dagger X^* v_i = 1$ for all i, where X^* is the optimal solution to $\mathrm{SDP}(V)$. We will then output $\exp\left(n\cdot r\left(\frac{\mathrm{tr}(A)}{n}\right)\right) = \left(1 - \frac{\epsilon^2}{20}\right)^n$.

Eq. (8) implies that the approximation factor of this algorithm is at least $e^{-(\gamma+1-\alpha)n}$, where α is the minimum value of $r(x) - \ell(x)$ over all $x \in [0,1]$. Write $\ell(x) \geq \ell'(x) := \max(0, \ln(1-\beta^*) + \frac{\beta^* x}{(1-\beta^*)} - \frac{0.273(\beta^*)^2}{(1-\beta^*)^2 x}$ for $\beta^* = 0.34$. One can numerically determine that $\alpha \geq \min_{0\leq x\leq 1} r(x) - \ell'(x) \geq 10^{-4}$. □

3.1 Proof of Lemma 1.3

We first prove an inequality that we will require. The proof of this inequality is inspired by an identity of Barvinok [Bar20].

Lemma 3.2. *For any matrix* $0 \preceq B \prec I$, $\mathrm{per}(I + B) \leq \det\left((I - B)^{-1}\right)$.

Proof. Write $B = VV^\dagger$ for $V \in \mathbb{C}^{\times n}$. Let $v_1, \ldots, v_n$ be the rows of V.

$$\begin{aligned}
\mathrm{per}(I+B) &= \sum_{S\subseteq[n]} \mathrm{per}(B_{S,S}) && \text{(Lemma 2.10)} \\
&= \sum_{S\subseteq[n]} \mathbb{E}_{g\sim\mathbb{C}\mathcal{N}(0,I)}\left[\prod_{i\in S} |\langle v_i, g\rangle|^2\right] && \text{(Lemma 2.9)} \\
&= \mathbb{E}_{g\sim\mathbb{C}\mathcal{N}(0,I)}\left[\prod_{i\in[n]} \left(1 + |\langle v_i, g\rangle|^2\right)\right] \\
&\leq \mathbb{E}_{g\sim\mathbb{C}\mathcal{N}(0,I)}\left[\prod_{i\in[n]} e^{|\langle v_i, g\rangle|^2}\right] && (1 + x \leq e^x \text{ for all } x) \\
&= \mathbb{E}_{g\sim\mathbb{C}\mathcal{N}(0,I)}\left[\exp\left(g^\dagger V^\dagger V g\right)\right].
\end{aligned}$$

Let $\sigma_1, \ldots, \sigma_n$ be the eigenvalues of $V^\dagger V$. Since g is invariant under unitary transformations, we can rotate g into the eigenbasis of $V^\dagger V$ to get that

$$\begin{aligned}
\mathrm{per}(I+B) &\leq \mathbb{E}_{g\sim\mathbb{C}\mathcal{N}(0,I)}\left[\exp\left(\sum_{i\in[n]} \sigma_i |g_i|^2\right)\right] \\
&= \prod_{i\in[n]} \mathbb{E}_{g\sim\mathbb{C}\mathcal{N}(0,1)}\left[e^{\sigma_i |g|^2}\right] && \text{(Independence of } g_i\text{)} \\
&= \prod_{i\in[n]} \frac{1}{1-\sigma_i}. && \text{(Fact 2.5)}
\end{aligned}$$

Noting that the eigenvalues of $V^\dagger V$ match those of $B = VV^\dagger$, this is equal to $\det\left((I - B)^{-1}\right)$. □

Now, we are ready to prove Lemma 1.3. Let $0 \preceq A \preceq I$ be a matrix with $\mathrm{tr}(A) \leq (1-\epsilon)n$. Let $0 \leq \lambda_1 \leq \ldots \leq \lambda_n \leq 1$ be the eigenvalues of A, and let $v_1, \ldots, v_n$ be the corresponding eigenvectors. Let $t \in (1/2, 1]$ be a parameter we will set later, and let i_t be the smallest index i such that $\lambda_i > t$. For any parameter $t \in (1/2, 1]$, we can write

$$A \preceq tI + \sum_{i\geq i_t} (\lambda_i - t) v_i v_i^\dagger = t \cdot \left(I + \sum_{i\geq i_t} \frac{\lambda_i - t}{t} v_i v_i^\dagger\right).$$

Write $B = \sum_{i \geq i_t} \frac{\lambda_i - t}{t} v_i v_i^\dagger$. Since $t > 1/2$, $B \prec I$, so it satisfies the conditions of Lemma 3.2.

$$\begin{aligned}\mathrm{per}(A) &\leq t^n \cdot \mathrm{per}(I + B) && \text{(Lemma 2.8)}\\ &\leq \frac{t^n}{\det(I - B)}. && \text{(Lemma 3.2)}\end{aligned}$$

We now pick $t = 1 - \epsilon/5$. For this choice of t, we must have $i_t \geq \frac{\epsilon n}{2}$, since otherwise, $\mathrm{tr}(A) \geq t \cdot (n - i_t) \geq (1 - \epsilon/5) \cdot (1 - \epsilon/2)n > (1 - \epsilon)n$ contradicts the fact that $\mathrm{tr}(A) \leq (1 - \epsilon)n$. We compute

$$\det(I - B) = \prod_{i \geq i_t} \left(1 - \frac{\lambda_i - t}{t}\right) = \prod_{i \geq i_t} \left(2 - \frac{\lambda_i}{t}\right) \geq \left(2 - \frac{1}{t}\right)^{n - i_t}.$$

Plugging in the definition of t and our lower bound on i_t, this is at least

$$\left(2 - \frac{1}{1 - \epsilon/5}\right)^{(1-\epsilon/2)n} = \left(\frac{1 - 2\epsilon/5}{1 - \epsilon/5}\right)^{(1-\epsilon/2)n}.$$

We now have our upper bound on $\mathrm{per}(A)$:

$$\mathrm{per}(A) \leq (1 - \epsilon/5)^n \cdot \left(\frac{1 - \epsilon/5}{1 - 2\epsilon/5}\right)^{(1-\epsilon/2)n} = \left(\frac{(1 - \epsilon/5) \cdot (1 - \epsilon/5)^{1-\epsilon/2}}{(1 - 2\epsilon/5)^{1-\epsilon/2}}\right)^n.$$

To complete the proof, we use that $\frac{(1-\epsilon/5)\cdot(1-\epsilon/5)^{1-\epsilon/2}}{(1-2\epsilon/5)^{1-\epsilon/2}} \leq 1 - \frac{\epsilon^2}{20}$ for all $\epsilon \in [0, 1]$.

3.2 Proof of Lemma 1.4

By Eq. (7), it suffices to prove a lower bound on $r(V)$. Let X^* be the optimal solution to $\mathrm{SDP}(V)$. Consider the following randomized rounding scheme to a solution of $r(V)$: sample $g \sim \mathbb{C}\mathcal{N}(0, X^*)$ and $s_i \sim \{z \in \mathbb{C} : |z| = 1\}$ independently for all $i \in [n]$. Let $x = \sqrt{1 - \beta} g + \sqrt{\frac{\beta n}{\mathrm{tr}(A)}} \sum_{i \in [n]} s_i v_i$. We will use the bound

$$r(V)^{1/n} = \max_{\|x\|_2^2 = 1} \prod_{i \in [n]} |\langle v_i, x \rangle|^{2/n} \geq \frac{\mathbb{E}_x[\prod_{i \in n} |\langle v_i, x \rangle|^{2/n}]}{\mathbb{E}_x[\|x\|_2^2]}.$$

First we compute the denominator.

$$\begin{aligned}n \cdot \mathbb{E}[\|x\|_2^2] &= n \cdot \mathbb{E}[\|x\|_2^2]\\ &= \mathbb{E}\left[\left\|\sqrt{1 - \beta} g + \sqrt{\frac{\beta n}{\mathrm{tr}(A)}} \sum_{i \in [n]} s_i v_i\right\|_{\ell_2}^2\right]\\ &= (1 - \beta)\mathbb{E}[\|g\|_{\ell_2}^2] + \frac{\beta n}{\mathrm{tr}(A)} \mathbb{E}\left[\sum_{i,j} s_i s_j \langle v_i, v_j \rangle\right] + \sqrt{\frac{\beta n}{\mathrm{tr}(A)}}(1 - \beta)\mathbb{E}\left[\left\langle g, \sum_i s_i v_i \right\rangle\right]\\ &= (1 - \beta)\mathbb{E}[\|g\|_{\ell_2}^2] + \frac{\beta n}{\mathrm{tr}(A)} \sum_i \|v_i\|_{\ell_2}^2 && \text{(Independence)}\\ &= (1 - \beta) \cdot \mathrm{tr}(X^*) + \frac{\beta n}{\mathrm{tr}(A)} \cdot \sum_{i \in [n]} \|v_i\|_{\ell_2}^2 && (g \sim \mathbb{C}\mathcal{N}(0, X^*), \text{definition of } \|\cdot\|_{\ell_2})\\ &= n. && (\mathrm{tr}(X^*) = n, \textstyle\sum_{i \in [n]} \|v_i\|_{\ell_2}^2 = \mathrm{tr}(VV^\dagger) = \mathrm{tr}(A))\end{aligned}$$

So, $\mathbb{E}[\|x\|_2^2] = 1$. It remains to lower bound the numerator. We start by applying Jensen's inequality to get

$$\mathbb{E}_x\left[\prod_{i\in n}|\langle v_i, x\rangle|^{2/n}\right] \geq \exp\left(\frac{1}{n}\sum_{i\in[n]}\mathbb{E}_x[\ln|\langle v_i, x\rangle|^2]\right). \tag{9}$$

We will bound each of the terms inside the sum. Fix some $i \in [n]$, and let $y_i = \sqrt{\frac{n}{\mathrm{tr}(A)}}\sum_{j\in[n]} s_j\langle v_i, v_j\rangle$ and $z_i = \langle g, v_i\rangle$, so $\langle v_i, x\rangle = \sqrt{1-\beta}z_i + \sqrt{\beta}y_i$. Notice that $z_i \sim \mathbb{C}\mathcal{N}(0, v_i^\dagger X^* v_i) = \mathbb{C}\mathcal{N}(0,1)$ by assumption. Let us bound

$$\begin{aligned}
\mathbb{E}[\ln|\langle v_i, x\rangle|^2] &= \mathbb{E}[\ln|\sqrt{1-\beta}z_i + \sqrt{\beta}y_i|^2] \\
&= \ln(1-\beta) + \mathbb{E}\left[\ln\left|z_i + \sqrt{\frac{\beta}{1-\beta}}y_i\right|^2\right] \\
&\geq -\gamma + \ln(1-\beta) + \frac{\beta}{1-\beta}\mathbb{E}[|y_i|^2] - \frac{\beta^2}{4(1-\beta)^2}\mathbb{E}[|y_i|^4].
\end{aligned}$$

(Lemma 2.7, $z_i \sim \mathbb{C}\mathcal{N}(0,1)$ and is independent of y_i)

We bound the second and fourth moments of y_i using the below claim, whose proof we defer to Section 3.2.1.

Claim 3.3. *For all $i \in [n]$,*

$$\mathbb{E}[|y_i|^2] = \frac{n}{\mathrm{tr}(A)}v_i^\dagger V^\dagger V v_i,$$

$$\mathbb{E}[|y_i|^4] \leq \frac{1.09n^2}{\mathrm{tr}(A)^2}\|v_i\|_{\ell_2}^2.$$

Plugging in the bounds from Claim 3.3 and summing over all i, we get

$$\begin{aligned}
\frac{1}{n}\sum_{i\in[n]}\mathbb{E}_x[\ln|\langle v_i, x\rangle|^2] &\geq -\gamma + \ln(1-\beta) + \frac{\beta}{(1-\beta)\,\mathrm{tr}(A)}\sum_{i\in[n]}v_i^\dagger V^\dagger V v_i - \frac{0.273\beta^2 n}{(1-\beta)^2\,\mathrm{tr}(A)^2}\sum_{i\in[n]}\|v_i\|_{\ell_2}^2 \\
&= -\gamma + \ln(1-\beta) + \frac{\beta}{1-\beta}\cdot\frac{\|A\|_F^2}{\mathrm{tr}(A)} - \frac{0.273\beta^2}{(1-\beta)^2}\cdot\frac{n}{\mathrm{tr}(A)} \\
&\geq -\gamma + \ln(1-\beta) + \frac{\beta}{1-\beta}\cdot\frac{\mathrm{tr}(A)}{n} - \frac{0.273\beta^2}{(1-\beta)^2}\cdot\frac{n}{\mathrm{tr}(A)}
\end{aligned}$$

($\|A\|_F^2 \geq \frac{\mathrm{tr}(A)^2}{n}$ by Jensen's inequality)

This completes the proof of Lemma 1.4.

3.2.1 Proof of Claim 3.3

Proof. We can directly compute

$$\frac{\mathrm{tr}(A)}{n}\mathbb{E}[|y_i|^2] = \sum_{j,k\in[n]} \mathbb{E}[s_j\overline{s_k}] \cdot \langle v_i, v_j\rangle\overline{\langle v_i, v_k\rangle}$$

$$= \sum_{j\in[n]} |\langle v_i, v_j\rangle|^2 = v_i^\dagger V^\dagger V v_i \qquad (s_j \text{ is independent from } s_k \text{ for } j \neq k)$$

Similarly,

$$\frac{\mathrm{tr}(A)^2}{n^2}\mathbb{E}[|y_i|^4] = \sum_{j,k,l,m\in[n]} \mathbb{E}[s_j s_k \overline{s_l s_m}]\langle v_i, v_j\rangle\langle v_i, v_k\rangle\overline{\langle v_i, v_l\rangle\langle v_i, v_m\rangle}$$

$$= \sum_{j\in[n]} |\langle v_i, v_j\rangle|^4 + 2\sum_{j\neq k} |\langle v_i, v_j\rangle\langle v_i, v_k\rangle|^2$$

$$= 2\left(\sum_{j\in[n]} |\langle v_i, v_j\rangle|\right)^2 - \sum_{j\in[n]} |\langle v_i, v_j\rangle|^4$$

$$= 2(v_i^\dagger V^\dagger V v_i)^2 - \sum_{j\in[n]} |\langle v_i, v_j\rangle|^4$$

$$\leq 2\|v_i\|_{\ell_2}^4 - \|v_i\|_{\ell_2}^8 \qquad (V^\dagger V \preceq I, \text{ since } A = VV^\dagger \preceq I)$$

$$\leq 1.09 \cdot \|v_i\|_{\ell_2}^2 \qquad (x^2 - x^4 \leq 1.09x \text{ for } x \geq 0)$$

The second equality is because $\mathbb{E}[s_j s_k \overline{s_l s_m}] = 0$ unless each index appears an equal number of times in $\{j, k\}$ and $\{l, m\}$. □

4 Hardness of Approximation

As mentioned in Section 1.1, we will first prove Theorem 1.5. Later on, in Section 4.2, we will use Theorem 1.5 to prove Theorem 1.2 using an approximation-preserving reduction to the permanent problem.

Our first result is a general inapproximability result for the f-mean version of $\|A\|_{2\to q}$ that is dependent on an appropriate family of gadgets $\{E_k\}$ as defined below.

Theorem 4.1. *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, and let $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be a continuous increasing 2-concave function such that $\lim_{x\to\infty} \frac{f(x)}{x^2} = 0$. Let $\delta, \gamma > 0$. Assume that for all k, there is a matrix $E_k : \mathbb{F}^k \to \mathbb{F}^{d_k}$ satisfying the following:*

1. $\|E_k\|_{2\to 2} = 1$.
2. *The entries of E_k have magnitude equal to $\frac{1}{\sqrt{k}}$.*
3. *for all vectors $x \in \mathbb{F}^k$ with $\|x\|_\infty \leq \delta \cdot \|x\|_{\ell_2}$, $[E_k x]_f \leq \gamma \cdot \|x\|_2$.*

Then for all $\epsilon > 0$, it is NP-Hard to distinguish between the following two cases given a matrix $A : \mathbb{F}^m \to \mathbb{F}^n$ with $\|A\|_{2\to 2} \leq 1$.

1. *Completeness:* $[A]_{2\to f} = 1$*, or*
2. *Soundness:* $[A]_{2\to f} \leq \gamma + \epsilon$*.*

The proof of the above theorem is in Section 4.1 and closely follows the arguments used in [BRS15; Bha+23]. In order to instantiate it, we will have to construct a family of gadgets $\{E_k\}$ that have $\|E_k\|_{2\to 2} = 1$, but at the same time have small $2 \to f$-norm when restricted to "smooth" vectors.

Definition 6. For $k \geq 1$, let us define $E_k^{(\mathbb{R})} \in \mathbb{R}^{2^k \times k}$ as the matrix whose rows consist of the members of $\frac{1}{\sqrt{k}} \cdot \{-1, +1\}^k$ ordered arbitrarily. Similarly, we define $E_k^{(\mathbb{C})} \in \mathbb{C}^{4^k \times k}$ as the matrix whose rows consist of the members of $\frac{1}{\sqrt{k}} \cdot \{-1, +1, -i, +i\}^k$ ordered arbitrarily.

Observe that these matrices are normalized so that $\|E_k^{(\mathbb{F})}\|_{2\to 2} = 1$. The following Lemma shows that condition 3 of Theorem 4.1 is satisfied with $\gamma \approx \gamma_{\mathbb{F},p}$.

Lemma 4.2. *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$*. Let* f *be an absolutely continuous* 2*-concave increasing function. Let* $x \in \mathbb{F}^k$*, and* $E = E_k^{(\mathbb{F})}$*. For all* $0 < \delta < 1$*, if* $\|x\|_\infty \leq \delta \|x\|_{\ell_2}$ *then*

$$\left[\frac{Ex}{\|x\|_2}\right]_f^f \leq [g]_f^f + C \cdot \left(- \int_0^{C\delta} \min(0, f(u)) + \delta \cdot \left(\max(0, f(2\sqrt{\log(1/\delta)})) + 2f'(1) \right) \right),$$

where $g \sim \mathbb{F}\mathcal{N}(0,1)$ *and* $C > 0$ *is a universal constant. In particular if* $f = f_p$ *for some* $-1 < p < 2$*, we have*

$$[Ex]_{f_p} \leq \|x\|_2 \cdot (\gamma_{\mathbb{F},p} + \epsilon_\delta),$$

where $\epsilon_\delta \to 0$ *as* $\delta \to 0$*.*

We prove Lemma 4.2 in Appendix A. The proof requires a Berry-Esseen type result for test functions of the form $f(\|.\|_2)$ applied to a sum of independent random vectors, which we prove in Appendix B.

With these results in hand, we can now prove Theorem 1.5.

Proof of Theorem 1.5. We pick δ to be such that Lemma 4.2 implies $\|Ex\|_q \leq \|x\|_2 \cdot (\gamma_{\mathbb{F},q} + \epsilon/2)$ for all x satisfying $\|x\|_\infty \leq \delta \|x\|_{\ell_2}$.

We apply Theorem 4.1 to the increasing 2-concave function $f = f_p$, gadget family $\{E_k^{(\mathbb{F})}\}$, and parameters δ, $\gamma = \gamma_{\mathbb{F},q} + \epsilon/2$, and $\epsilon/2$. By Lemma 4.2, the three conditions are satisfied, implying that it is NP-Hard to distinguish the case that $\|A\|_{2\to q} = 1$ and $\|A\|_{2\to q} \leq \gamma_{\mathbb{F},q} + \epsilon$. ☐

4.1 Proof of Theorem 4.1

We will closely follow the arguments used in [BRS15; Bha+23]. The starting point of our reduction will be the following result implicit in [BRS15], which informally says that it is NP-hard to find a sparse vector in a subspace, according to a certain block-wise notion of sparsity.

Theorem 4.3 ([BRS15])**.** *For all* $\epsilon, \delta, \alpha > 0$ *and* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$*, there is a* $k = \mathrm{poly}(1/\epsilon, 1/\delta, 1/\alpha)$ *such that given a subspace* $W \subseteq \mathbb{F}^{n\times k}$ *in the form of a projection matrix* $P \in \mathbb{F}^{(n\times k)\times(n\times k)}$*, it is* NP*-Hard to distinguish between the following:*

- *There is a vector $x \in W$ such that for all $i \in [n]$, the vector $x_i \in \mathbb{F}^k$ is in $\{e_1, \ldots, e_k\}$.*
- *For all vectors $x \in W$ with $\|x\|_{\ell_2}^2 = n$, the set*

  $$S = \{i \in [n] : \|x_i\|_{\ell_2} \leq 1/\alpha, \|x_i\|_\infty \geq \delta\}$$

  has size at most ϵn.

Let $0 < \epsilon', \alpha \leq 1$ be constant parameters depending on δ, γ, and ϵ that we will specify later. We prove hardness of the $[A]_{2\to f}$ problem by a reduction from the NP-Hard problem described in Theorem 4.3 with parameters ϵ', δ, and α. Theorem 4.3 implies that there is some $k = \text{poly}(1/\epsilon', 1/\delta, 1/\alpha)$ such that given a projection matrix $P \in \mathbb{F}^{(n\times k)\times(n\times k)}$ for a subspace $W \subseteq \mathbb{F}^{n\times k}$, it is NP-Hard to distinguish between the following:

- There is a vector $x \in W$ such that for all $i \in [n]$, the vector $x_i \in \mathbb{F}^k$ is in $\{e_1, \ldots, e_k\}$.
- For all vectors $x \in W$ with $\|x\|_{\ell_2}^2 = n$, the set

  $$S = \{i \in [n] : \|x_i\|_{\ell_2} \leq 1/\alpha, \|x_i\|_\infty \geq \delta\}$$

  has size at most ϵn.

Our reduction will map the projection matrix $P \in \mathbb{F}^{(n\times k)\times(n\times k)}$ to the matrix $A = (I_n \otimes E_k) \cdot P$. Note that $A \in \mathbb{F}^{(n\times d_k)\times(n\times k)}$. To analyze the reduction, we must prove completeness and soundness.

4.1.1 Completeness

If there is a vector $x \in W$ such that for all $i \in [n]$, $x_i \in \{e_1, \ldots, e_k\}$, we need to show $[A]_{2\to f} = 1$.

Indeed, we can consider the vector $z := Ax = (E_k \otimes I_n) \cdot Px = (E_k \otimes I_n)x$. We have for all $i \in [n]$, $z_i = E_k x_i$. Since x_i is a standard basis vector, z_i must be equal to some column of E_k. So by Assumption 2, all entries of z have magnitude $1/\sqrt{k}$, implying $[z]_f = f^{-1}(f(\frac{1}{\sqrt{k}})) = \frac{1}{\sqrt{k}}$. Therefore $[A]_{2\to f} \geq \frac{[z]_f}{\|x\|_2} = 1$.

On the other hand, $[A]_{2\to f} \leq \|A\|_{2\to 2} \leq \|P\|_{2\to 2} \cdot \|E_k\|_{2\to 2} \leq 1$ by Claim 2.1.

4.1.2 Soundness

Assuming for all $x \in W$ with $\|x\|_{\ell_2}^2 \leq n$, the set

$$S = \{i \in [n] : \|x_i\|_{\ell_2} \leq 1/\alpha, \|x_i\|_\infty \geq \delta\}$$

has size at most $\epsilon' n$, we need to show $[A]_{2\to f} \leq \gamma + \epsilon$.

Let $y \in \mathbb{F}^{n\times k}$ be an arbitrary vector with $\|y\|_2 = 1$, and set $x = \frac{1}{k} \cdot Py$ and $z = Ay = k \cdot (E_k \otimes I_n)x$. Note that because P is a projection matrix, $\|x\|_{\ell_2}^2 = nk \cdot \|x\|_2^2 \leq n\|y\|_2^2 = n$, and $\|z\|_2 \leq \|y\|_2 = 1$. By virtue of the normalization on x, we have $\|x_i\|_{\ell_2} = \|z_i\|_2$ for each block $i \in [n]$.

We must show $[z]_f \leq \gamma + \epsilon$. We will upper bound the contribution of different indices $i \in [n]$ to $[z]_f$ separately. To do this, define the following partition of $[n]$:

$$V_0 := S = \{i \in [n] : \|x_i\|_{\ell_2} \leq 1/\alpha, \|x_i\|_\infty \geq \delta\},$$

$$V_1 := \{i \in [n] : \|x_i\|_{\ell_2} \le \alpha, \|x_i\|_\infty \le \delta\alpha\},$$
$$V_2 := \{i \in [n] : \|x_i\|_{\ell_2} \ge \alpha, \|x_i\|_\infty \le \delta\alpha\},$$
$$V_3 := \{i \in [n] : \|x_i\|_{\ell_2} > 1/\alpha\}.$$

For all $u \in \{0,1,2,3\}$, define $z^{(u)} \in \mathbb{F}^{V_u \times d_k}$ as the collection of $z_i \in \mathbb{F}^{d_k}$ for all $i \in V_u$. Note that

$$[z]_f^f = \sum_{u \in \{0,1,2,3\}} \frac{|V_u|}{|V|} \cdot [z^{(u)}]_f^f. \tag{10}$$

We will prove bounds on $[z^{(u)}]_f^f$ for $u \in \{0,1,2,3\}$.
For $u = 0$, Claim 2.1 applied to z_i implies

$$[z^{(0)}]_f^f = \mathbb{E}_{i \sim V_0}[z_i]_f^f \le \mathbb{E}_{i \sim V_0} f(\|z_i\|_2) = \mathbb{E}_{i \sim V_0} f(\|x_i\|_{\ell_2}) \le f(1/\alpha). \tag{11}$$

For $u = 1$, a similar application of Claim 2.1 implies

$$[z^{(1)}]_f^f = \mathbb{E}_{i \sim V_1}[z_i]_f^f \le \mathbb{E}_{i \sim V_1} f(\|z_i\|_2) = \mathbb{E}_{i \sim V_1} f(\|x_i\|_{\ell_2}) \le f(\alpha). \tag{12}$$

For $u = 2$, we have

$$[z^{(2)}]_f^f = \mathbb{E}_{i \sim V_2}[z_i]_f^f \underset{\text{Assumption 3}}{\le} \mathbb{E}_{i \sim V_2}\left[f(\gamma \cdot \|z_i\|_2)\right] \underset{\textit{Claim 2.1}}{\le} f(\gamma \cdot \|z^{(2)}\|_2) \underset{\|z\|_2 \le 1}{\le} f\left(\gamma \cdot \sqrt{\frac{|V|}{|V_2|}}\right). \tag{13}$$

Finally, for $u = 3$,

$$\begin{aligned} [z^{(3)}]_f^f &= \mathbb{E}_{i \sim V_3}[z_i]_f^f \\ &\underset{\textit{Claim 2.1}}{\le} \mathbb{E}_{i \sim V_3} f(\|z_i\|_2) \\ &= \mathbb{E}_{i \sim V_3}\left[\|z_i\|_2^2 \cdot \frac{f(\|z_i\|_2)}{\|z_i\|_2^2}\right] \\ &\le \sup_{w \ge 1/\alpha} \frac{f(w)}{w^2} \cdot \mathbb{E}_{i \sim V_3}\|z_i\|_2^2 \\ &= \sup_{w \ge 1/\alpha} \frac{f(w)}{w^2} \cdot \|z^{(3)}\|_2^2 \\ &\underset{\|z\|_2^2 = 1}{\le} \sup_{w \ge 1/\alpha} \frac{f(w)}{w^2} \cdot \frac{|V|}{|V_3|}. \end{aligned} \tag{14}$$

Now we are equipped to bound Eq. (10).

$$
\begin{aligned}
[z]_f^f &\le \sum_{u\in\{0,1,2\}} \frac{|V_u|}{|V\setminus V_3|}[z^{(u)}]_f^f + \sup_{w\ge 1/\alpha}\frac{f(w)}{w^2} && \text{(Eqs. (10) and (14))}\\
&\le f\left(\sqrt{\sum_{u\in\{0,1,2\}} \frac{|V_u|}{|V\setminus V_3|}[z^{(u)}]_f^2}\right) + \sup_{w\ge 1/\alpha}\frac{f(w)}{w^2} && \text{(Jensen's inequality for } x\to f(\sqrt{x})\text{)}\\
&\le f\left(\sqrt{\sum_{u\in\{0,1,2\}} \frac{|V_u|}{(1-\alpha^2)|V|}[z^{(u)}]_f^2}\right) + \sup_{w\ge 1/\alpha}\frac{f(w)}{w^2} && (|V_3|\le \alpha^2\cdot|V_0| \text{ by def of } V_3)\\
&\le f\left(\sqrt{\frac{\epsilon'/\alpha^2+\alpha^2+\gamma^2}{(1-\alpha^2)}}\right) + \sup_{w\ge 1/\alpha}\frac{f(w)}{w^2} && \text{(Eqs. (11) to (13) and } V_0\le \epsilon'|V|)\\
&\le f\left(\frac{\gamma+\sqrt{\epsilon'}/\alpha+\alpha}{\sqrt{1-\alpha^2}}\right) + \sup_{w\ge 1/\alpha}\frac{f(w)}{w^2} && (\sqrt{a+b}\le\sqrt{a}+\sqrt{b},\ f \text{ is monotone})\\
&= f\left(\frac{\gamma+2\alpha}{\sqrt{1-\alpha^2}}\right) + \sup_{w\ge 1/\alpha}\frac{f(w)}{w^2}. && (\text{Setting } \epsilon'=\alpha^4)
\end{aligned}
$$

Using the assumption that f is continuous and $\lim_{x\to\infty} f(x)/x^2 = 0$, we get that the limit of the right hand side as $\alpha \to 0$ is exactly $f(\gamma)$. Therefore, there exists some $\alpha > 0$ independent of n such that $[z]_f \le \gamma + \epsilon$. We choose α in our invocation of Theorem 4.3 accordingly, completing the proof that $[A]_{2\to f} \le \gamma + \epsilon$.

4.2 Proof of Theorem 1.2

In this section we prove Theorem 1.2. We use the following lemma which proves that the approximability of the permanent of highly rank-deficient $n \times n$ PSD matrices is essentially the same as the approximability of of the $2 \to 0$ norm.

Lemma 4.4. *Let $V \in \mathbb{C}^{n\times d}$. Then,*

$$
c_{n,d}^{\mathbb{C}} \cdot \binom{n+d-1}{d}^{-1} \cdot \|V\|_{2\to 0}^{2n} \le \mathrm{per}(VV^\dagger) \le c_{n,d}^{\mathbb{C}} \cdot \|V\|_{2\to 0}^{2n},
$$

where $c_{n,d}^{\mathbb{C}}$ is defined in Lemma 2.9.

Proof. Let $v_1, \ldots, v_n$ be the rows of V. For the upper bound, we can write

$$
\begin{aligned}
\mathrm{per}(VV^\dagger) &= \mathbb{E}_{x\sim\mathbb{C}N(0,I)} \prod_{i\in[n]} |\langle x, v_i\rangle|^2\\
&= c_{n,d}\cdot \mathbb{E}_{\|x\|_2=1} \prod_{i\in[n]} |\langle x, v_i\rangle|^2 && \text{(Lemma 2.9)}\\
&\le c_{n,d}\cdot \max_{\|x\|_2=1} \prod_{i\in[n]} |\langle x, v_i\rangle|^2\\
&= c_{n,d}\cdot \|V\|_{2\to 0}^{2n}.
\end{aligned}
$$

Next we prove the lower bound. Let $z \in \mathbb{C}^d$ be a vector with $\|z\|_2 = 1$ vector maximizing $\|Vz\|_0 = \prod_{i\in[n]} |\langle z, v_i\rangle|^{1/n}$. We have $zz^\dagger \preceq \|z\|_{\ell_2}^2 \cdot I = d \cdot I$, so $Vzz^\dagger V^\dagger \preceq d \cdot VV^\dagger$. By Lemma 2.8, this implies $\mathrm{per}(Vzz^\dagger V^\dagger) \le d^n \cdot \mathrm{per}(VV^\dagger)$. Since $Vzz^\dagger V^\dagger$ is rank 1, we can compute its permanent as $\mathrm{per}(Vzz^\dagger V^\dagger) = n! \cdot \prod_{i\in[n]} |\langle z, v_i\rangle|^2$.

$$
\begin{aligned}
\|Vz\|_0^{2n} &= \max_{\|x\|_2=1} \prod_{i\in[n]} |\langle x, v_i\rangle|^2 \\
&= \frac{1}{n!} \cdot \mathrm{per}(Vzz^\dagger V^\dagger) \\
&\le \frac{d^n}{n!} \cdot \mathrm{per}(VV^\dagger) \\
&= \binom{n+d-1}{d} \cdot c_{n,d}^{-1} \cdot \mathrm{per}(VV^\dagger).
\end{aligned}
$$

□

Proof of Theorem 1.2. We start from Theorem 1.5 for the case $\mathbb{F} = \mathbb{C}$ and $q = 0$, to get that it is NP-hard to approximate $\|A\|_{2\to 0}$ within a factor of $e^{-\gamma/2+\epsilon/4}$ (recall that $\gamma_{\mathbb{C},0} = e^{-\gamma/2}$ by Fact 2.4).

We reduce the problem of approximating $\|A\|_{2\to 0}$ for $A \in \mathbb{C}^{n\times d}$ to approximating the permanent of the positive semidefinite matrix $B = A^{(k)}(A^{(k)})^\dagger$, where $A^{(k)} \in \mathbb{C}^{nk\times d}$ is as in Fact 2.3. By Lemma 4.4, $\mathrm{per}(B)$ is proportional to $\|A^{(k)}\|_{2\to 0}^{2nk}$ up to a multiplicative error of

$$
\binom{nk+d-1}{d} \le e^{\epsilon kn/2},
$$

which holds for $k = O(\frac{d}{n\epsilon^2})$ and $\epsilon > 0$ small enough. Note that the reduction is efficient because k is polynomial in the size of A.

By Fact 2.3, we have $\|A^{(k)}\|_{2\to 0}^{2nk} = \|A\|_{2\to 0}^{2nk}$ which is hard to approximate within a factor of $e^{-kn(\gamma-\epsilon/2)}$. Therefore, it is NP-hard to approximate $\mathrm{per}(B)$ within a factor of $e^{-kn(\gamma-\epsilon)}$, where $B \in \mathbb{C}^{kn\times kn}$. □

References

[Ana+17] Nima Anari, Leonid Gurvits, Shayan Oveis Gharan, et al. "Simply Exponential Approximation of the Permanent of Positive Semidefinite Matrices". In: *FOCS*. Ed. by Chris Umans. IEEE Computer Society, 2017, pp. 914–925 (cit. on pp. 1, 2, 9, 11, 12).

[AS48] Milton Abramowitz and Irene A Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Vol. 55. US Government printing office, 1948 (cit. on p. 9).

[Bar+12] Boaz Barak, Fernando G.S.L. Brandao, Aram W. Harrow, et al. "Hypercontractivity, sum-of-squares proofs, and their applications". In: *STOC*. ACM, 2012, pp. 307–326 (cit. on p. 4).

[Bar16] Alexander Barvinok. *Combinatorics and complexity of partition functions*. Vol. 30. Springer, 2016 (cit. on p. 10).

[Bar20] Alexander Barvinok. "A remark on approximating permanents of positive definite matrices". arxiv. 2020. URL: https://arxiv.org/abs/2005.06344 (cit. on pp. 1, 3, 13).

[Ben05] Vidmantas Bentkus. "A Lyapunov-type bound in Rd". In: *Theory of Probability & Its Applications* 49.2 (2005), pp. 311–323 (cit. on p. 24).

[Bha+23] Vijay Bhattiprolu, Mrinal Kanti Ghosh, Venkatesan Guruswami, et al. "Inapproximability of Matrix $p \to q$ Norms". In: *SIAM Journal on Computing* 52.1 (2023), pp. 132–155. DOI: 10.1137/18M1233418 (cit. on pp. 4, 5, 17).

[BRS15] Jop Briët, Oded Regev, and Rishi Saket. "Tight hardness of the non-commutative Grothendieck problem". In: *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*. IEEE. 2015, pp. 1108–1122 (cit. on pp. 4, 5, 17).

[DG03] Sanjoy Dasgupta and Anupam Gupta. "An elementary proof of a theorem of Johnson and Lindenstrauss". In: *Random Structures & Algorithms* 22.1 (2003), pp. 60–65 (cit. on p. 25).

[Gur+16] Venkatesan Guruswami, Prasad Raghavendra, Rishi Saket, et al. "Bypassing UGC from Some Optimal Geometric Inapproximability Results". In: *ACM Transactions on Algorithms* 12.1 (2016), pp. 1–25 (cit. on p. 5).

[HLP52] Godfrey Harold Hardy, John Edensor Littlewood, and George Pólya. *Inequalities*. Cambridge university press, 1952 (cit. on p. 8).

[JSV04] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. "A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries". In: *Journal of the ACM (JACM)* 51.4 (2004), pp. 671–697 (cit. on p. 1).

[Mei23] A Meiburg. "Inapproximability of Positive Semidefinite Permanents and Quantum State Tomography". In: *Algorithmica* 85 (2023), pp. 3828–3854. DOI: https://doi.org/10.1007/s00453-023-01169-1 (cit. on pp. 1, 2).

[Mos20] Stefan M Moser. "Expected Logarithm and Negative Integer Moments of a Noncentral χ 2-Distributed Random Variable". In: *Entropy* 22.9 (2020), p. 1048 (cit. on p. 9).

[Per12] Jerome K Percus. *Combinatorial methods*. Vol. 4. Springer Science & Business Media, 2012 (cit. on p. 11).

[Pin] Iosif Pinelis. "An Approach to Inequalities for the Distributions of Infinite-Dimensional Martingales". In: *Progress in Probability* 30 (), pp. 128–134 (cit. on p. 24).

[Ste05] Daureen Steinberg. "Computation of matrix norms with applications to robust optimization". In: *Research thesis, Technion-Israel University of Technology* 2 (2005) (cit. on p. 4).

[YP21] Chenyang Yuan and Pablo A. Parrilo. "Semidefinite Relaxations of Products of Nonnegative Forms on the Sphere". CoRR abs/2102.13220. 2021. eprint: 2102.13220. URL: https://arxiv.org/abs/2102.13220 (cit. on p. 1).

[YP22] Chenyang Yuan and Pablo A. Parrilo. "Maximizing products of linear forms, and the permanent of positive semidefinite matrices". In: *Math. Program.* 193.1 (2022), pp. 499–510 (cit. on pp. 1, 2, 11, 12).

A Proof of Lemma 4.2

In this section we prove Lemma 4.2. We use the following corollary which we will prove in Appendix B.

Corollary A.1. *Let $f : \mathbb{R}_{>0} \to \mathbb{R}$ be an absolutely continuous 2-concave increasing function with $0 \in \mathrm{Range}(f)$. Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, and let $z \in \mathbb{F}^n$ be a vector with $\|z\|_{\ell_2} = 1$ and $\|z\|_\infty \leq \delta$. For each i, let σ_i be an independently and uniformly sampled member of $\{-1, +1\}$ if $\mathbb{F} = \mathbb{R}$ and be an independently and uniformly sampled member of $\{-1, +1, -i, +i\}$ otherwise. Then there exists a universal $C > 0$ such that $Z = \sum_{i \in [n]} \sigma_i z_i$ satisfies*

$$[Z]_f^f \leq [g]_f^f + C \cdot \left(- \int_0^{C\delta} \min(0, f(u)) + \delta \cdot \left(f(2\sqrt{\log(1/\delta)}) + f'(1) \right) \right),$$

where $g \sim \mathbb{F}\mathcal{N}(0, 1)$.

Now we are ready to prove the main result of this section.

Proof of Lemma 4.2. Let $z = x/\|x\|_2$ and $y = Ez = Ex/\|x\|_2$. Let m be the number of rows of E. Note that for $i \in [m]$, we have

$$y_i = \sum_{j \in [n]} E_{i,j} \cdot z_j.$$

Let $Z_j = E_{i,j} \cdot z_j$ be the random variable where i chosen uniformly at random from $[m]$. We invoke Corollary A.1 on $z/\sqrt{k}$. We verify its conditions: First, $\left\| \frac{z}{\sqrt{k}} \right\|_{\ell_2} = \|z\|_2 = 1$. Second, by definition of $E_k^{(\mathbb{F})}$ we can write $Z_j = \sigma_j \cdot \frac{z_j}{\sqrt{k}}$, where the random variables $\sigma_j \in \{+1, -1\}$ when $\mathbb{F} = \mathbb{R}$ and $\sigma_j \in \{+1, -1, +i, -i\}$ chosen independently and uniformly at random Lastly,

$$\frac{|z_j|}{\sqrt{k}} = \frac{|x_j|}{\sqrt{k}\,\|x\|_2} = \frac{|x_j|}{\|x\|_{\ell_2}} \leq \frac{\|x\|_\infty}{\|x\|_{\ell_2}} \leq \delta.$$

Now by invoking Lemma B.1 on the random variables $Z_1, \ldots, Z_k$, we get

$$\left[\frac{Ex}{\|x\|_2} \right]_f^f = [y]_f^f = \mathbb{E}_{i \sim [m]} f(y_i) = \mathbb{E}\big[f\big(\sum_{j \in [n]} Z_j \big) \big] \leq [g]_f^f + C \cdot \left(- \int_0^{C\delta} \min(0, f(u)) + \delta \cdot \left(f(2\sqrt{\log(1/\delta)}) + f'(1) \right) \right)$$

as desired.

Now assume $f = f_p$ for some $-1 < p < 2$. By Lemma 2.2, $\| \cdot \|_p$ is homogeneous, and we get

$$\begin{aligned} [Ex]_f &= \|x\|_2 \cdot \left[\frac{Ex}{\|x\|_2} \right]_f \\ &\leq \|x\|_2 \cdot f^{-1} \left(\gamma_{\mathbb{F},p}^p + C \cdot \left(- \int_0^{C\delta} \min(0, f(u)) + \delta \cdot \left(f(2\sqrt{\log(1/\delta)}) + f'(1) \right) \right) \right) \end{aligned}$$

The second term is 0 if $p > 0$, otherwise it is equal to $\frac{\delta^{p+1}}{p+1}$. The third term is 2δ if $p < 0$, otherwise it is bounded by $4\delta(\log(1/\delta) + 1)$ for δ small enough. Therefore, the limit of the right hand side as $\delta \to 0$ is $\|x\|_2 \cdot \gamma_{F,p}$. □

B Proof of Corollary A.1

Corollary A.1 follows directly as a special case of the following Lemma, for $k = 1$ if $\mathbb{F} = \mathbb{R}$ and for $k = 2$ if $\mathbb{F} = \mathbb{C}$.

Lemma B.1. *Let $f : \mathbb{R}_{>0} \to \mathbb{R}$ be an absolutely continuous 2-concave increasing function with $0 \in \overline{\mathrm{Range}(f)}$. Let $k \geq 1$ and let $X_1, \ldots, X_n$ are bounded independent random variables in $\mathbb{R}^k$ with $\mathbb{E}[X_i] = 0$, $Cov(\sum_i X_i) = I_k/k$, and $\|X_i\|_{\ell_2} \leq \delta_i$ such that $\sum_i \delta_i^2 \leq 1$ and $\delta_i \leq \delta$ for some $0 < \delta < 1$, then there exists $\eta_k > 0$ such that*

$$\left[\| \sum_{i \in [n]} X_i\|_{\ell_2}\right]_f^f \leq [\|g\|_{\ell_2}]_f^f - e\int_0^{C_k\delta/e} \min(0, f(u))du + \delta \cdot \left(C_k \cdot \max(0, f(2\sqrt{\log(1/\delta)})) + 2f'(1)\right),$$

where $g \sim N(0, I_k/k)$, and C_k is the constant in Theorem B.2.

Before proving Lemma B.1, we state some probabilistic tools we will require in the proof.

Theorem B.2 (Multivariate Berry-Esseen [Ben05])**.** *Let $k \geq 1$ and let $X_1, \ldots, X_n$ be independent random variables in $\mathbb{R}^k$ satisfying $\mathbb{E}[X_i] = 0$ for $1 \leq i \leq n$. Define $X = X_1 + \cdots + X_n$ and suppose $\mathbb{E}[XX^T] = I_k$. Further let $g \sim \mathcal{N}(0, I_k)$. Then there exists $C_k > 0$ such that for all convex sets $U \subseteq \mathbb{R}^k$ it holds that*

$$|\Pr[g \in U] - \Pr[X \in U]| \leq C_k \cdot \mathbb{E}_{i \in [n]}[\|X_i\|_{\ell_2}^3].$$

In particular, for all $u \geq 0$ it holds that

$$|\Pr[\|g\|_{\ell_2} \leq u] - \Pr[\|X\|_{\ell_2} \leq u]| \leq C_k \cdot \mathbb{E}_{i \in [n]}[\|X_i\|_{\ell_2}^3].$$

Theorem B.3 (Multivariate Hoeffding [Pin, Theorem 3])**.** *Let $k \geq 1$ and let $X_1, \ldots, X_n$ be independent random variables in $\mathbb{R}^k$ satisfying $\mathbb{E}[X_i] = 0$ for $1 \leq i \leq n$. Further suppose $\|X_i\|_{\ell_2} \leq \delta_i$ almost surely for $\delta_1, \ldots, \delta_n > 0$. Define $X = X_1 + \cdots + X_n$.*

$$|\Pr[\|X\|_{\ell_2} \geq u]| \leq 2 \cdot e^{-u^2/2\sum_{i \in [n]} \delta_i^2}.$$

Fact B.4. *Let $f : \mathbb{R}_{>0} \to \mathbb{R}$ be an absolutely continuous function. If $\int_0^b f(u)$ is bounded then $\lim_{u \to 0} uf(u) = 0$.*

Claim B.5. *Let X be a random variable over $\mathbb{R}_{>0}$, and let $f : \mathbb{R}_{>0} \to \mathbb{R}$ be an absolutely continuous function with $f(a) = 0$ for some $a \in \mathbb{R}_{>0}$. Then, if $\mathbb{E}[f(X)]$ exists,*

$$\mathbb{E}[f(X)] = -\int_0^a f'(u)\Pr[X \leq u]du + \int_a^\infty f'(u)\Pr[X \geq u]du + \lim_{u \to 0} f(u)\Pr[X \leq u] - \lim_{u \to \infty} f(u)\Pr[X \geq u].$$

Proof. Let $\mu : \mathbb{R}_{>0} \to \mathbb{R}$ denote the PDF of the random variable X.

$$\begin{aligned}
\mathbb{E}[f(X)] &= \int_0^\infty f(u) d\mu(u) \\
&= \int_0^a f(u) d\mu(u) + \int_a^\infty f(u) d\mu(u) \\
&= f(u) \Pr[X \le u]\Big|_0^a - \int_{I_1} f'(u) \Pr[X \le u] - f(u) \Pr[X \ge u]\Big|_a^\infty + \int_{I_2} f'(u) \Pr[X \ge u] \\
&= \lim_{u \to 0} f(u) \Pr[X \le u] - \lim_{u \to \infty} f(u) \Pr[X \ge u] - \int_{I_1} f'(u) \Pr[X \le u] + \int_{I_2} f'(u) \Pr[X \ge u] \\
& \qquad (f(a) = 0.)
\end{aligned}$$

□

Fact B.6. *Let $k \ge 1$ and let $g \sim N(0, I_k/k)$. Then for all $0 \le u \le 1$, $\Pr[\|g\|_{\ell_2} \le u] \le eu$.*

Proof. $k \cdot \|g\|_{\ell_2}^2$ is distributed as a Chi-squared random variable with k degrees of freedom. In [DG03, Lemma 2.2], the authors show that

$$\Pr[\|g\|_{\ell_2} \le u] \le (u^2 e^{1-u^2})^{k/2} \le e \cdot u,$$

as desired. □

With these facts in hand, we are ready to prove Lemma B.1.

Proof of Lemma B.1. Define the random variable $X = \|\sum_i X_i\|_{\ell_2}$. We split the domain of f into $I_1 = f^{-1}([-\infty, 0])$ and $I_2 = f^{-1}([0, \infty])$, where $I_1 = [0, a]$ and $I_2 = [a, \infty]$. Since $0 \in \overline{\text{Range}(f)}$, we have $f(a) = 0$. We use Claim B.5 to write the left hand side as

$$\begin{aligned}
\mathbb{E}[f(X)] &= -\int_0^a f'(u) \Pr[X \le u] + \int_a^\infty f'(u) \Pr[X \ge u] + \lim_{u \to 0} f(u) \Pr[X \le u] - \lim_{u \to \infty} f(u) \Pr[X \ge u] \\
&\le -\int_0^a f'(u) \Pr[X \le u] + \int_a^\infty f'(u) \Pr[X \ge u]. \qquad (f(0) \le f(a) = 0,\ f(\infty) \ge f(a) = 0.)
\end{aligned}$$

We bound each of the above integrals separately.

Claim B.7. *We have*

$$\int_0^a f'(u) \Pr[X \le u] \ge \int_0^a f'(u) \Pr[\|g\|_{\ell_2} \le u] + e \int_0^{C_k \delta / e} \min(0, f(u)).$$

Claim B.8. *We have*

$$\int_a^\infty f'(u) \Pr[X \ge u] \le \int_a^\infty f'(u) \Pr[|g| \ge u] + \delta \cdot \left(C_k \cdot \max(0, f(2\sqrt{\log(1/\delta)})) + 2f'(1) \right).$$

We will complete the proof of Lemma B.1 using these two claims, and prove the claims later.

$$\begin{aligned}\mathbb{E}[f(X)] &\le -\int_0^a f'(u)\Pr[X \le u] + \int_a^\infty f'(u)\Pr[X \ge u] \\ &\le -\int_0^a f'(u)\Pr[\|g\|_{\ell_2} \le u] - e\int_0^{C_k\delta/e} \min(0, f(u)) \\ &\quad + \int_a^\infty f'(u)\Pr[|g| \ge u] + \delta\cdot\left(C_k\cdot f(2\sqrt{\log(1/\delta)}) + 2f'(1)\right) \\ &= \mathbb{E}[f(\|g\|_{\ell_2})] - \int_0^{C_k\delta} \min(0, f(u)) + \delta\cdot\left(C_k\cdot \max(0, f(2\sqrt{\log(1/\delta)})) + 2f'(1)\right).\end{aligned}$$

Here, the last equality is by applying Claim B.5 on the random variable $\|g\|_{\ell_2}$:

$$\begin{aligned}\mathbb{E}[\|g\|_{\ell_2}] = &-\int_0^a f'(u)\Pr[\|g\|_{\ell_2} \le u] + \int_a^\infty f'(u)\Pr[|g| \ge u] \\ &+ \lim_{u\to 0} f(u)\Pr[\|g\|_{\ell_2} \le u] - \lim_{u\to\infty} f(u)\Pr[\|g\|_{\ell_2} \ge u]\end{aligned}$$

Observe that by Fact B.4, the third term above is zero, and by Claim 2.1,

$$\lim_{u\to\infty} f(u)\Pr[\|g\|_{\ell_2} \ge u] \le \lim_{u\to\infty}(f'(1)\cdot u^2 + f(1))\cdot\Pr[\|g\|_{\ell_2} \ge u] = 0.$$

This completes the justification of the last inequality. It remains to prove Claim B.7 and Claim B.8. We begin by writing an inequality which will be used in both proofs. By Theorem B.2, for any $u \ge 0$,

$$\begin{aligned}|\Pr[X \le u] - \Pr[\|g\|_{\ell_2} \le u]| &\le C_k\cdot\mathbb{E}_{i\in[n]}[\|X_i\|_{\ell_2}^3] \\ &\le C_k\cdot\sum_{i\in[n]}[\|X_i\|_{\ell_2}^2]\cdot\max_{i\in[n]}\|X_i\|_{\ell_2} \\ &\le C_k\cdot\delta. \qquad (15)\end{aligned}$$

Proof of Claim B.7. For a parameter b with $0 \le b \le a$, we write

$$\begin{aligned}
\int_{I_1} f'(u)\Pr[X \le u] &\ge \int_b^a f'(u)\Pr[X \le u] && (f'(u) \ge 0)\\
&\ge \int_b^a f'(u)(\Pr[\|g\|_{\ell_2} \le u] - C_k\delta) && (\text{Eq. (15)})\\
&= \int_0^a f'(u)\Pr[\|g\|_{\ell_2} \le u] - \int_0^b f'(u)\Pr[\|g\|_{\ell_2} \le u] - C_k\delta(f(a)-f(b))\\
&\ge \int_0^a f'(u)\Pr[\|g\|_{\ell_2} \le u] - e\int_0^b f'(u)u + C_k\delta f(b)\\
& && (\text{using } f(a)=0, \text{ and Fact B.6})\\
&= \int_0^a f'(u)\Pr[\|g\|_{\ell_2} \le u] - euf(u)\Big|_0^b + e\int_0^b f(u) + C_k\delta f(b)\\
&= \int_0^a f'(u)\Pr[\|g\|_{\ell_2} \le u] - ebf(b) + e\int_0^b f(u) + C_k\delta f(b) && (\text{Fact B.4})\\
&= \int_0^a f'(u)\Pr[\|g\|_{\ell_2} \le u] + e\int_0^{\min(a,C_k\delta/e)} f(u) && (\text{Setting } b = \min(a, C_k\delta/e))\\
&= \int_0^a f'(u)\Pr[\|g\|_{\ell_2} \le u] + e\int_0^{C_k\delta/e} \min(f(u),0).
\end{aligned}$$

The penultimate equality is because if $\min(a, C_k\delta/e) = a$, then $f(b) = 0$, and if $\min(a, C_k\delta/e) = C_k\delta/e$, then $bf(b) = C_k\delta f(b)/e$.

□

Proof of Claim B.8. Applying Theorem B.3 on X,

$$\Pr[X \ge u] \le 2 \cdot e^{-u^2/2\sum \delta_i^2} \le 2 \cdot e^{-u^2/2}. \tag{16}$$

For $\max(a,1) < c < \infty$ we write

$$\begin{aligned}
\int_{I_2} f'(u)\Pr[X \ge u] &= \int_a^c f'(u)\Pr[X \ge u] + \int_c^\infty f'(u)\Pr[X \ge u]\\
&\le \int_a^c f'(u)(\Pr[|g| \ge u] + C_k\delta) + \int_c^\infty 2f'(u)e^{-u^2/2} && (\text{Eq. (15), Eq. (16)})\\
&\le \int_a^c f'(u)\Pr[|g| \ge u] + C_k\delta \cdot (f(c) - f(a)) + 2f'(1) \cdot \int_c^\infty ue^{-u^2/2}\\
& && (\text{using } f \text{ is 2-concave, Claim 2.1, and } c \ge 1)\\
&\le \int_a^\infty f'(u)\Pr[|g| \ge u] + C_k\delta \cdot f(c) + 2f'(1) \cdot \int_c^\infty ue^{-u^2/2}\\
& && (f \text{ is increasing, } f(a)=0)\\
&= \int_a^\infty f'(u)\Pr[|g| \ge u] + C_k\delta \cdot f(c) + 2f'(1) \cdot e^{-c^2/2} && (17)
\end{aligned}$$

Setting $c = \max(a, 2\sqrt{\log(1/\delta)})$, we get

$$\int_{I_2} f'(u)\Pr[X \ge u] \le \int_a^\infty f'(u)\Pr[|g| \ge u] + \delta \cdot \left(C_k \cdot \max(0, f(2\sqrt{\log(1/\delta)})) + 2f'(1)\right). \tag{18}$$

The inequality above is by bounding $e^{-c^2/2} \leq \delta$, and if $\max(a, 2\sqrt{\log(1/\delta)}) = a$, then $f(c) = 0$. ☐

☐

C Algorithm for approximating $2 \to q$ norm

Let $q < 2$. Given a matrix $A \in \mathbb{F}^{n \times d}$ with rows $\{a_i\}_{i \in [n]}$, we write an expression for $\|A\|_{2\to q}$:

$$\|A\|_{2\to q} = f_q^{-1}\left(\max_{x \in \mathbb{F}^d, \|x\|_2 = 1}\left[\sum_{i \in [n]} f_q\left(|\langle a_i, x\rangle|\right)\right]\right) = f_q^{-1}\left(\max_{x \in \mathbb{F}^d, \|x\|_2 = 1}\left[\sum_{i \in [n]} f_q\left(\sqrt{a_i^\dagger x x^\dagger a_i}\right)\right]\right).$$

Notice that $xx^\dagger$ is a rank-1 PSD matrix with $\mathrm{tr}(xx^\dagger) = x^\dagger x = \|x\|_{\ell_2} = d \cdot \|x\|_2 = d$. We can relax the rank 1 constraint to obtain a relaxation of $\|A\|_{2\to q}$ which we denote by $\mathrm{SDP}_{2\to q}(A)$:

$$\mathrm{SDP}_{2\to q}(A) := f_q^{-1}\left(\max_{X \succeq 0, \mathrm{tr}(X) = d}\left[\sum_{i \in [n]} f_q\left(\sqrt{a_i^\dagger X a_i}\right)\right]\right),$$

where the maximum is taken over all matrices in $\mathbb{F}^{d \times d}$. Note that the objective function being maximized is concave for all $q \leq 2$ because the function $X \to a_i^\dagger X a_i$ is linear and f_q is 2-concave.

Lemma C.1. *For any matrix $A \in \mathbb{F}^{n \times d}$ and any $-1 < q \leq 2$,*

$$\|A\|_{2\to q} \leq \mathrm{SDP}_{2\to q}(A) \leq \gamma_{\mathbb{F},q}^{-1} \cdot \|A\|_{2\to q}.$$

Proof. The first inequality is because $\mathrm{SDP}_{2\to q}(A)$ is a relaxation of $\|A\|_{2\to q}$. For the second inequality, we describe a rounding procedure for the relaxation.

Let $X^* \in \mathbb{F}^{d \times d}$ be the optimal solution to $\mathrm{SDP}_{2\to q}(A)$. We sample a vector $x \sim \mathbb{F}\mathcal{N}(0, X^*)$. Clearly,

$$\|A\|_{2\to q}^2 = \max_{x \in \mathbb{F}^n, x \neq 0} \frac{\|Ax\|_q^2}{\|x\|_2^2} \geq \frac{\mathbb{E}\|Ax\|_q^2}{\mathbb{E}\|x\|_2^2}. \tag{19}$$

First we calculate the denominator of the right hand side: $\mathbb{E}\|x\|_2^2 = \frac{1}{d} \cdot \mathrm{tr}(X) = 1$. For the numerator, we bound

$$
\begin{aligned}
\mathbb{E}\|Ax\|_q^2 &= \mathbb{E}\left(f_q^{-1}\left(\sum_{i\in[n]} f_q(|\langle a_i, x\rangle|)\right)^2\right) \\
&\geq f_q^{-1}\left(\sum_{i\in[n]} \mathbb{E} f_q(|\langle a_i, x\rangle|)\right)^2 && \text{(Jensen's inequality, and } x \to f_q^{-1}(x)^2 \text{ is convex)} \\
&= f_q^{-1}\left(\sum_{i\in[n]} \mathbb{E}_{g\sim\mathbb{F}\mathcal{N}(0,1)} f_q\left(\sqrt{a_i^\dagger X^* a_i}\cdot |g|\right)\right)^2 \\
& && (\langle a_i, x\rangle \text{ is a Gaussian with variance } a_i^\dagger X^* a_i) \\
&= f_q^{-1}\left(\mathbb{E}_{g\sim\mathbb{F}\mathcal{N}(0,1)} f_q(|g|)\right)^2 \cdot f_q^{-1}\left(\sum_{i\in[n]} f_q\left(\sqrt{a_i^\dagger X^* a_i}\right)\right)^2 && \text{(Homogeneity)} \\
&= \gamma_{\mathbb{F},q}^2 \cdot \mathrm{SDP}_{2\to q}(A)^2.
\end{aligned}
$$

Together with Eq. (19), this implies $\|A\|_{2\to q} \geq \gamma_{\mathbb{F},q} \cdot \mathrm{SDP}_{2\to q}(A)$, completing the proof. □