UiO : Department of Mathematics
University of Oslo

# Joint operator convexity of a certain trace functional.

**Hans Henrich Neumann**

Master's Thesis, Spring 2021

This master's thesis is submitted under the master's programme *Mathematics*, with programme option *Mathematics*, at the Department of Mathematics, University of Oslo. The scope of the thesis is 30 credits.

The front page depicts a section of the root system of the exceptional Lie group $E_8$, projected into the plane. Lie groups were invented by the Norwegian mathematician Sophus Lie (1842–1899) to express symmetries in differential equations and today they play a central role in various parts of mathematics.

# 1 Introduction

A state of a system in quantum mechanics can be mathematically represented as a density matrix $\rho$, that is a positive definite matrix with $\text{Tr}(\rho) = 1$. A state is called *pure* if it can be represented as a vector $v$. The corresponding density matrix is then a rank one projection onto the span of $v$. A state that is not pure is called *mixed.* A mixed state $\rho$ can be written as a convex combination

$$\rho = \sum_i p_i \psi_i \psi_i^*,$$

where $*$ denotes the conjugate transpose, and the $\psi_i$ are pure states. The *entropy* $S$ of a quantum state $\rho$, as defined by von Neumann, is

$$S(\rho) := -\operatorname{Tr} \rho \log(\rho).$$

Assuming that the $\psi_i$ are orthogonal to each other in the formula $\rho = \sum_i p_i \psi_i \psi_i^*$ given above, noting that the $p_i$ are then the eigenvalues of $\rho$, and using that the trace is the sum of eigenvalues we see that

$$S(\rho) = -\sum_i p_i \log(p_i).$$

In this case the entropy of a pure state is 0, but strictly positive for a mixed state. The entropy gives a measure of the uncertainty in an inherently statistical quantum state. In a probability distribution where some outcomes have a higher probability than others has a lower uncertainty of the outcome than a probability distribution that is closer to being uniform. The average of two distributions is closer to uniform than either is separately, so the entropy of the average of two distributions should be higher than the average of the entropies of the two distributions. In other words, it is natural to expect the entropy of probability distributions and $S$ to be concave. This is indeed the case.

If $\rho_1$ and $\rho_2$ are two states and we let $\rho_{1,2}$ denote the state of their composite quantum system then we have the subadditivity inequality

$$S(\rho_{1,2}) \leq S(\rho_1) + S(\rho_2).$$

In the case that $\rho_{1,2,3}$ is the state of the composite of three subsystems whose states are $\rho_1$, $\rho_2$ and $\rho_3$ we get the following inequality known as strong subadditivity

$$S(\rho_{1,2,3}) + S(\rho_2) \leq S(\rho_{1,2}) + S(\rho_{2,3}).$$

See for example [Rus02] for details.

The *relative entropy* $S(\rho||\sigma)$ of two states $\rho$ and $\sigma$ is

$$S(\rho||\sigma) := \operatorname{Tr} \rho(\log(\rho) - \log(\sigma)).$$

The relative entropy gives a measure of how much two states differ. Two states that differ should have a higher relative entropy than two equal states, so it is natural to expect the relative entropy to be convex. The relative entropy is indeed jointly convex. The rest of this section lists results on convexity/concavity that doesn't necessarily have a direct interpretation in physics.

Lieb showed in [Lie73] the concavity of the map

$$A \mapsto \operatorname{Tr} A^r K^* A^p K$$

for a positive matrix $A$, thus proving the Wigner-Yanase-Dyson conjecture. He also showed that the two variable version

$$(A, B) \mapsto \operatorname{Tr} A^r K^* B^p K$$

is jointly concave in positive $A$ and $B$. In [Lie73] it was also shown that the map

$$A \mapsto \operatorname{Tr} \exp(L + \ln A)$$

is concave. Carlen, Frank and Lieb showed in [CFL16] that the map

$$(A, B) \mapsto \operatorname{Tr}(A^{\frac{q}{2}} B^p A^{\frac{q}{2}})^s$$

is jointly convex or jointly concave for varying values of $p$, $q$ and $s$.

A real function $f(x)$ is called *log-convex* if $\log\big(f(x)\big)$ is convex. Ando and Hiai showed in [AH11] that the map

$$A \mapsto \log \omega\big(f(A)\big),$$

where $f$ is an operator monotone decreasing function, and $\omega$ is a positive linear functional, is convex. They also showed that a function $f$ is operator monotone decreasing if and only if $f$ is operator log-convex, and similarly that a function $f$ is operator monotone increasing if and only if $f$ is operator log-concave.

The next two recent results, that grew out of these earlier developments, motivates our main theorem. Hiai showed in [Hia16] the joint convexity/concavity of the maps

$$(A, B) \mapsto \operatorname{Tr} g\big(\Phi(A^p)^{\frac{1}{2}} \Psi(B^q) \Phi(A^p)^{\frac{1}{2}}\big)$$

$$(A, B) \mapsto \operatorname{Tr} g\Big(\big(\Phi(A^{-p})^{\frac{1}{2}} \Psi(B^{-q}) \Phi(A^{-p})^{\frac{1}{2}}\big)^{-1}\Big).$$

Kirihata and Yamashita showed in [KY20] the convexity of the map

$$A \mapsto g\Big(\Phi\big(f(A)\big)\Big),$$

where $\Phi$ is a strictly positive linear operator. They also showed that the map

$$(A, B) \mapsto \tau\Big(\Phi\big(f_1(A)\big)^{\frac{1}{2}} \Psi\big(f_2(B)\big) \Phi\big(f_1(A)\big)^{\frac{1}{2}}\Big),$$

where $\tau$ is a tracial positive functional, is jointly convex.

The goal of this thesis is to show the joint convexity of the map

$$(A, B) \mapsto g\Big(\Phi\big(f_1(A)\big)^{\frac{1}{2}} \Psi\big(f_2(B)\big) \Phi\big(f_1(A)\big)^{\frac{1}{2}}\Big).$$

This would generalize the result in [Hia16] to more functions than power functions and in [KY20] from a functional $\tau$ to a functional of the form $A \mapsto \operatorname{Tr} g(A)$.

## 2 Preliminaries

This section is a review of some basic material in matrix analysis. See for example [Bha97] and [Bha07].

### Functional calculus.

**Definition 2.1.** Let $p(x) = c_n x^n + c_{n-1} x^{n-1} + \dots c_1 x + c_0$ be a real polynomial. For a hermitian matrix $A$ we define the *fuctional calculus* $p(A)$ of $p$ at $A$ to be

$$p(A) := c_n A^n + c_{n-1} A^{n-1} + \dots c_1 A + c_0 I.$$

Where $I$ denotes the identity matrix with the same dimensions as $A$.

With diagonalization $A = UDU^*$, for a diagonal matrix $D = \operatorname{diag}(\{\lambda_i\})$, we get

$$\begin{aligned} p(A) &= c_n A^n + c_{n-1} A^{n-1} + \dots + c_1 A + c_0 I \\ &= c_n (UDU^*)^n + c_{n-1} (UDU^*)^{n-1} + \dots + c_1 UDU^* + c_0 UIU^* \\ &= c_n UD^n U^* + c_{n-1} UD^{n-1} U^* + \dots + c_1 UDU^* + c_0 UIU^* \\ &= U \operatorname{diag}(\{c_n \lambda_i^n\}) U^* + U \operatorname{diag}(\{c_{n-1} \lambda_i^{n-1}\}) U^* + \dots \\ &\quad + U \operatorname{diag}(\{c_1 \lambda_i\}) U^* + U \operatorname{diag}(\{c_0\}) U^* \\ &= U \operatorname{diag}(\{p(\lambda_i)\}) U^*. \end{aligned}$$

Which we write as $Up(D)U^*$.
Many functions can be approximated by polynomials, so the functional calculus of a more general function is defined similarly. For a diagonal matrix:

**Definition 2.2.** Let $D = \operatorname{diag}(\lambda_i)$ be a diagonal matrix, and let $f$ be a function defined on a set that contains $\{\lambda_i\}$. Then we define the functional calculus $f(D)$ of $f$ at $D$ to be

$$f(D) := \operatorname{diag}(\{f(\lambda_i)\}).$$

For a general hermitian matrix:

**Definition 2.3.** Let $A$ be a hermitian matrix with diagonalization $A = UDU^*$, and let $f$ be a function defined on a set that contains the spectrum of $A$. Then we define the functional calculus $f(A)$ of $f$ at $A$ to be

$$f(A) := Uf(D)U^*.$$

**Example 2.4.** For $f(x) = e^x$ and

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

we get

$$f(A) = \begin{bmatrix} e^2 & 0 \\ 0 & e^3 \end{bmatrix}.$$

**Example 2.5.** For $f(x) = x^{\frac{1}{2}}$ and

$$A = \frac{1}{3} \begin{bmatrix} 19 & 8\sqrt{2}i \\ -8\sqrt{2}i & 11 \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} i & \sqrt{2} \\ -\sqrt{2} & -i \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix} \frac{1}{\sqrt{3}} \begin{bmatrix} -i & -\sqrt{2} \\ \sqrt{2} & i \end{bmatrix}$$

we get

$$f(A) = A^{\frac{1}{2}} = \frac{1}{\sqrt{3}} \begin{bmatrix} i & \sqrt{2} \\ -\sqrt{2} & -i \end{bmatrix} \begin{bmatrix} 1^{\frac{1}{2}} & 0 \\ 0 & 9^{\frac{1}{2}} \end{bmatrix} \frac{1}{\sqrt{3}} \begin{bmatrix} -i & -\sqrt{2} \\ \sqrt{2} & i \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 7 & 2\sqrt{2}i \\ -2\sqrt{2}i & 5 \end{bmatrix}.$$

**Positivity.**

**Definition 2.6.** A matrix $A$ is called *positive* if it is hermitian and all eigenvalues of $A$ are nonnegative, and *strictly positive* if the eigenvalues are positive. We write $0 \leq A$ in the former case and $0 < A$ in the latter case.

We denote the set of $n \times n$ matrices by $\mathbb{M}_{\mathrm{n}}$, the set of positive $n \times n$ matrices by $\mathbb{M}_{\mathrm{n}}^{+}$ and the set of strictly positive $n \times n$ matrices by $\mathbb{M}_{\mathrm{n}}^{++}$.
Positivity has many equivalent definitions. One of them states that $A$ is positive if and only if its quadratic form is positive semidefinite. That is, if

$$0 \leq x^* A x, \quad \text{for all } x.$$

Similarly $A$ is strictly positive if and only if its quadratic form is positive definite:

$$0 < x^* A x, \quad \text{for all } x \neq 0.$$

A third definition states that $A$ is positive if and only if $A = B^* B$ for some matrix $B$. $A$ is then strictly positive if and only if $B$ is invertible.
The notion of positivity also makes sense for linear operators on some Hilbert space. In this case a self-adjoint linear operator $\Phi$ is called positive if its spectrum $\sigma(\Phi)$ is a subset of $[0, \infty)$ and strictly positive if $\sigma(\Phi) \subset (0, \infty)$. We will, however, not need this in this text.
Positivity gives us a partial ordering on the set of hermitian matrices:

$$A \leq B, \quad \text{if} \quad 0 \leq B - A.$$

**Example 2.7.** We have

$$0 \leq \begin{bmatrix} 7 & 2\sqrt{2}i \\ -2\sqrt{2}i & 5 \end{bmatrix}$$

and

$$\begin{bmatrix} 7 & 2\sqrt{2}i \\ -2\sqrt{2}i & 5 \end{bmatrix} \leq \begin{bmatrix} 19 & 8\sqrt{2}i \\ -8\sqrt{2}i & 11 \end{bmatrix}.$$

**Example 2.8.** For

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

we neither have $A \leq B$ nor $B \leq A$.

**Example 2.9.** The matrix

$$\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$$

has only positive elements, so it may look positive at first glance. Its diagonalization, however, shows that this is a deception:

$$\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

**Definition 2.10.** A linear operator $\Phi\colon \mathbb{M}_{\mathrm{n}} \to \mathbb{M}_{\mathrm{m}}$ is called *positive* if $\Phi(A)$ is positive whenever $A$ is positive. $\Phi$ is called *strictly positive* if it is positive and $\Phi(A)$ is invertible whenever $A$ is invertible.

**Example 2.11.** The linear operator $\Phi(A) = XAX^* + YAY^*$ for some unitary matrices $X$ and $Y$ is strictly positive:
Writing $A = UDU^*$ we have $XAX^* = XUDU^*X^* = XUD(XU)^*$ and, since $XU(XU)^* = XUU^*X^* = XIX^* = I$ , we see that this is a hermitian matrix with positive eigenvalues whenever $A$ is. Similarly with $YAY^*$.

**Example 2.12.** The linear operator

$$\Phi(A) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} A \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

that acts on $2 \times 2$ matrices is positive, but not strictly positive:

$$\begin{aligned}
\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} A \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 & u_2 \\ u_3 & u_4 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \overline{u_1} & \overline{u_3} \\ \overline{u_2} & \overline{u_4} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 u_1 \overline{u_1} + \lambda_2 u_2 \overline{u_2} & 0 \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 \|u_1\|^2 + \lambda_2 \|u_2\|^2 & 0 \\ 0 & 0 \end{bmatrix}.
\end{aligned}$$

**Example 2.13.** The linear functional $\phi(A) = \operatorname{Tr} A$ is strictly positive as the trace of $A$ is the sum of the eigenvalues of $A$.

### Convexity and monotonicity.

**Definition 2.14.** We say that a function $f$, defined on an interval $(a, b)$, is *operator convex* if

$$f\big((1-t)A + tB\big) \le (1-t)f(A) + tf(B), \quad t \in [0, 1]$$

holds for all hermitian matrices $A$ and $B$ whose eigenvalues is in $(a, b)$. Similarly, we say that $f$ is *operator concave* if the reverse inequality holds:

$$f\big((1-t)A + tB\big) \ge (1-t)f(A) + tf(B), \quad t \in [0, 1].$$

If $f$ is continuous then $f$ is operator convex if

$$f\left(\frac{A+B}{2}\right) \le \frac{f(A) + f(B)}{2}, \quad t \in [0, 1].$$

If $f$ is operator convex then $-f$ is operator concave, and if $f$ is operator concave then $-f$ is operator convex.

**Example 2.15.** $f(x) = x^2$ is operator convex.
We have

$$\begin{aligned}\frac{f(A)+f(B)}{2} - f\left(\frac{A+B}{2}\right) &= \frac{A^2+B^2}{2} - \left(\frac{A+B}{2}\right)^2 \\ &= \frac{2A^2+2B^2}{4} - \frac{(A^2+B^2+AB+BA)}{4} \\ &= \frac{A^2+B^2-AB-BA}{4} \\ &= \frac{(A-B)^2}{4}\end{aligned}$$

Which is positive since the eigenvalues of $A-B$ gets squared.

**Example 2.16.** The function $f(x) = x^3$ is convex on $[0,\infty)$, but not operator convex. With

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$$

we get

$$\frac{f(A)+f(B)}{2} - f\left(\frac{A+B}{2}\right) = \begin{bmatrix} 6 & 1 \\ 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 3-\sqrt{10} & 3+\sqrt{10} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3-\sqrt{10} & 0 \\ 0 & 3+\sqrt{10} \end{bmatrix} \begin{bmatrix} -\frac{1}{2\sqrt{10}} & \frac{1}{2}+\frac{3}{2\sqrt{10}} \\ \frac{1}{2\sqrt{10}} & \frac{1}{2}-\frac{3}{2\sqrt{10}} \end{bmatrix}.$$

**Definition 2.17.** We say that a function $f$ is *operator monotone increasing* (or simply *operator monotone*) on an interval $(a,b)$ if we have

$$A \leq B \Rightarrow f(A) \leq f(B).$$

for any hermitian matrices $A$ and $B$ with eigenvalues in $(a,b)$. Similarly we say that $f$ is *operator monotone decreasing* if

$$A \leq B \Rightarrow f(A) \geq f(B).$$

If $f$ is operator monotone increasing then $-f$ is operator monotone decreasing, and if $f$ is operator monotone decreasing then $-f$ is operator monotone increasing.
If $f(x)$ is operator monotone increasing then $f(x^{-1})$ and $f(x)^{-1}$ are operator monotone decreasing.

**Theorem 2.18.** *Let $f\colon (0,\infty) \to \mathbb{R}$ be a continuous, operator monotone increasing function. Then $f$ is operator concave.*

Similarly a continuous, operator monotone decreasing function $f\colon (0,\infty) \to \mathbb{R}$ is operator convex.

*Remark* 2.19. 2.18 can be proven by using the integral representation for operator monotone increasing functions introduced later, or see [Bha97, Theorem V.2.5].

**Example 2.20.** The linear operator $\Phi(A) = XAX^*$ is operator monotone increasing.
Let $A \leq B$, and let $C$ be such that $B - A = C^*C$. Then

$$XBX^* - XAX^* = X(B-A)X^* = XC^*CX^* = (CX^*)^*CX^*$$

which is positive.

**Example 2.21.** The function $f(x) = x^{-1}$ is operator monotone decreasing on $(0, \infty)$.
First note that if $I \leq A$ we get

$$\begin{aligned} I &\leq A \\ A^{-\frac{1}{2}} I A^{-\frac{1}{2}} &\leq A^{-\frac{1}{2}} A A^{-\frac{1}{2}} \\ A^{-1} &\leq I. \end{aligned}$$

Now let $0 < A \leq B$. Then

$$\begin{aligned} A &\leq B \\ I &\leq A^{-\frac{1}{2}} B A^{-\frac{1}{2}} \\ I &\geq (A^{-\frac{1}{2}} B A^{-\frac{1}{2}})^{-1} \\ I &\geq A^{\frac{1}{2}} B^{-1} A^{\frac{1}{2}} \\ A^{-1} &\geq B^{-1}. \end{aligned}$$

**Example 2.22.** Although $f(x) = x^2$ is monotone increasing on $[0, \infty)$ it is not operator monotone increasing. We have

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \leq \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

but

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^2 - \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}^2 = \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} \frac{1}{2}(3 - \sqrt{13}) & \frac{1}{2}(3 + \sqrt{13}) \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2}(3 - \sqrt{13}) & 0 \\ 0 & \frac{1}{2}(3 + \sqrt{13}) \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{13}} & \frac{1}{2} + \frac{3}{2\sqrt{13}} \\ \frac{1}{\sqrt{13}} & \frac{1}{2} - \frac{3}{2\sqrt{13}} \end{bmatrix}.$$

**Integral representation.** The source for this section is [Bha97, Section V.4].

**Theorem 2.23.** *Let $K$ be the set of all functions $f$ that are operator monotone increasing on $(-1, 1)$ and such that $f(0) = 0$ and $f'(0) = 1$. Then $K$ is convex, compact in the pointwise convergence topology, and its extreme points have the form*

$$f(x) = \frac{x}{1 - ax}, \quad \text{where } a = \frac{1}{2} f''(0).$$

The next theorem is a consequence of 2.23.

**Theorem 2.24.** *Let $f$ be a nonconstant operator monotone increasing function on $(-1, 1)$. Then $f$ has an integral representation*

$$f(x) = f(0) + f'(0) \int_{-1}^{1} \frac{x}{1 - \lambda x} d\mu(\lambda)$$

*for a unique probability measure $\mu$ on $[-1, 1]$.*

2.24 gives us an integral representation of operator convex functions.

**Theorem 2.25.** *Let $f$ be a nonlinear operator convex function on $(-1,1)$. Then $f$ has an integral representation*

$$f(x) = f(0) + f'(0)x + \frac{1}{2}f''(0)\int_{-1}^{1} \frac{x^2}{1-\lambda x} d\mu(\lambda)$$

*for a unique probability measure $\mu$ on $[-1,1]$.*

A function $f$ is operator monotone increasing on an interval $(a,b)$ if and only if $f\big(\frac{(b-a)t}{2} + \frac{b+a}{2}\big)$ is operator monotone increasing on $(-1,1)$, so the theorems above holds for functions on $(a,b)$ as well.
Denote by $H_+$ the upper halfplane $\{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ of the complex plane, and by $H_-$ the lower halfplane $\{z \in \mathbb{C} : \mathrm{Im}(z) < 0\}$.

**Definition 2.26.** A function $f\colon H_+ \to \overline{H_+}$ is called a *Pick function* if it is complex analytic. The set of all Pick functions will be denoted by P.

From *Nevanlinna's Theorem* we get the following theorem.

**Theorem 2.27.** *A function $f$ is in $P$ if and only if it has an integral representation*

$$f(z) = a + bz + \int_{-\infty}^{\infty} \frac{1}{\lambda - z} d\mu(\lambda) + C$$

*where $a \in \mathbb{R}$, $b \geq 0$, $\mu$ is a positive Borel measure on $\mathbb{R}$ and $C$ is the constant*

$$C := \int_{-\infty}^{\infty} \frac{\lambda}{\lambda^2 + 1} d\mu(\lambda).$$

The representation can be shown to be unique. Denote by $P(a,b)$ the subset of $P$ consisting of functions that can be analytically continued across the open interval $(a,b)$ to $H_-$ by reflection.
The results above leads to the following two theorems.

**Theorem 2.28.** *A function $f$ is in $P(a,b)$ if and only if the measure $\mu$ in its representation is such that $\mu(a,b) = 0$.*

**Theorem 2.29.** *A function $f$ is in $P(a,b)$ if and only if $f$ is operator monotone increasing on $(a,b)$.*

### Means.

**Definition 2.30.** For matrices $A$ and $B$ the *arethmetic mean* $A\triangledown B$ is defined to be

$$A\triangledown B := \frac{A+B}{2}.$$

**Definition 2.31.** For invertible matrices $A$ and $B$ the *harmonic mean* $A!B$ is defined to be

$$A!B := \left(\frac{A^{-1} + B^{-1}}{2}\right)^{-1}.$$

We have the relation

$$A!B = (A^{-1} \triangledown B^{-1})^{-1}.$$

**Definition 2.32.** For positive matrices $A$ and $B$ the *geometric mean* $A\#B$ is defined to be

$$A\#B := A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}}A^{\frac{1}{2}}.$$

**Theorem 2.33.** *Let $A$ and $B$ be positive matrices. We then have the order relation*

$$A!B \leq A\#B \leq A\nabla B.$$

*Proof.* Assume for contradiction that $A!B > A\#B$. Then we get

$$\begin{aligned}
\left(\frac{A^{-1}+B^{-1}}{2}\right)^{-1} &> A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}}A^{\frac{1}{2}} \\
A^{-1}+B^{-1} &< 2A^{-\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{-\frac{1}{2}}A^{-\frac{1}{2}} \\
I + A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}} &< 2(A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}})^{\frac{1}{2}} \\
I + A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}} &< 2A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}} \\
I &< A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}} \\
A^{-1} &< B^{-1}.
\end{aligned}$$

Since the means are symmetric we also get $B^{-1} < A^{-1}$. This is a contradiction, so $A!B \leq A\#B$. Assume for contradiction that $A\#B > A\nabla B$. Similarly to the above we get

$$\begin{aligned}
A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}}A^{\frac{1}{2}} &> \frac{A+B}{2} \\
2(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}} &> I + A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \\
2A^{-\frac{1}{2}}BA^{-\frac{1}{2}} &> I + A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \\
A^{-\frac{1}{2}}BA^{-\frac{1}{2}} &> I \\
B &> A.
\end{aligned}$$

Since the means are symmetric we also get $A > B$. This is a contradiction, so $A\#B \leq A\nabla B$. ■

## 3 Main Theorem

**Definition 3.1.** Define $\mathcal{F}$ to be the set of functions $f$ that are non-decreasing and concave on $(0, \infty)$ with $\lim_{x\to\infty} \frac{f(x)}{x} = 0$.

**Definition 3.2.** Let $f \in \mathcal{F}$. Define the function $\hat{f}$ by

$$\hat{f}(x) := \inf_{t\in(0,\infty)} \big(tx - f(t)\big).$$

Given a function $f$ we can find $\hat{f}$ by differentiating $tx - f(t)$ with respect to $t$ and finding its minimum.

**Example 3.3.** For $f(x) = x^{\frac{1}{2}}$ we get

$$\hat{f}(x) = \inf_{t\in(0,\infty)} (tx - t^{\frac{1}{2}}).$$

The derivative of $tx - t^{\frac{1}{2}}$ is

$$\frac{\partial}{\partial t}(tx - t^{\frac{1}{2}}) = x - \frac{1}{2\sqrt{t}}$$

which has the root $t = \frac{1}{4x^2}$. This gives us

$$\hat{f}(x) = -\frac{1}{4x}.$$

**Lemma 3.4** ([Hia16]). *Let $f \in \mathcal{F}$ and $B \in \mathbb{M}_{\mathrm{n}}^{++}$. Then*

$$\operatorname{Tr} f(B) = \inf_{A \in \mathbb{M}_{\mathrm{n}}^{++}} \big( \operatorname{Tr} AB - \operatorname{Tr} \hat{f}(A) \big).$$

*proof (sketch).* Writing $B = UDU^*$ we get, from the trace property,

$$\operatorname{Tr} f(B) = \operatorname{Tr} U f(D) U^* = \operatorname{Tr} UU^* f(D) = \operatorname{Tr} f(D)$$

and

$$\operatorname{Tr} AB = \operatorname{Tr} AUDU^* = \operatorname{Tr} AD.$$

We can therefore assume that $B$ is diagonal and write $B = \operatorname{diag}(b_1, b_2, \ldots, b_n)$ for $b_1 \geq b_2 \geq \ldots \geq b_n$. It can be shown that $\hat{\hat{f}} = f$, so we have $f(x) = \inf_{t \in (0,\infty)} \big( tx - \hat{f}(t) \big)$. Now we get

$$\begin{aligned}
\operatorname{Tr} f(B) &= \sum_{i=1}^{n} f(b_i) \\
&= \inf_{a_1, a_2, \ldots, a_n} \sum_{i=1}^{n} \big( a_i b_i - \hat{f}(a_i) \big) \\
&= \inf_{A := \operatorname{diag}(a_1, a_2, \ldots, a_n) \in \mathbb{M}_{\mathrm{n}}^{++}} \big( \operatorname{Tr} AB - \operatorname{Tr} \hat{f}(A) \big) \\
&\geq \inf_{A \in \mathbb{M}_{\mathrm{n}}^{++}} \big( \operatorname{Tr} AB - \operatorname{Tr} \hat{f}(A) \big).
\end{aligned}$$

It can be shown, for $A \in \mathbb{M}_{\mathrm{n}}^{++}$ with eigenvalues $a_1 \geq a_2 \geq \ldots \geq a_n$, that $\operatorname{Tr} AB \geq \sum_{i=1}^{n} a_i b_{n+1-i}$. With this we get

$$\operatorname{Tr} AB - \operatorname{Tr} \hat{f}(A) \geq \sum_{i=1}^{n} \big( a_i b_{n+1-i} - \hat{f}(a_i) \big) \geq \sum_{i=1}^{n} f(b_i) = \operatorname{Tr} f(B).$$

Taking infimum we get

$$\inf_{A \in \mathbb{M}_{\mathrm{n}}^{++}} \big( \operatorname{Tr} AB - \operatorname{Tr} \hat{f}(A) \big) \geq \operatorname{Tr} f(B).$$

With both

$$\operatorname{Tr} f(B) \geq \inf_{A \in \mathbb{M}_{\mathrm{n}}^{++}} \big( \operatorname{Tr} AB - \operatorname{Tr} \hat{f}(A) \big) \text{ and } \inf_{A \in \mathbb{M}_{\mathrm{n}}^{++}} \big( \operatorname{Tr} AB - \operatorname{Tr} \hat{f}(A) \big) \geq \operatorname{Tr} f(B)$$

we get

$$\operatorname{Tr} f(B) = \inf_{A \in \mathbb{M}_{\mathrm{n}}^{++}} \big( \operatorname{Tr} AB - \operatorname{Tr} \hat{f}(A) \big).$$

■

**Lemma 3.5** ([KY20]). *Let $f\colon (0,\infty) \to (0,\infty)$ be an operator monotone decreasing function, $g\colon (0,\infty) \to \mathbb{R}$ an operator monotone increasing function and $\Phi : \mathbb{M}_{\mathrm{n}}^{++} \to \mathbb{M}_{\mathrm{m}}^{++}$ a strictly positive linear operator.*
*Then the map*

$$\mathbb{M}_{\mathrm{n}}^{++} \to \mathbb{M}_{\mathrm{m}}^{++}, \quad A \mapsto g\Big(\Phi\big(f(A)\big)\Big)$$

*is convex.*

To prove this lemma we need two lemmas:

**Lemma 3.6** ([KY20]). *Let $f\colon (0,\infty) \to (0,\infty)$ be an operator monotone decreasing function. Then we have*

$$f(A \triangledown B) \leq f(A)!f(B)$$

*for $A, B \in \mathbb{M}_{\mathrm{n}}^{++}$.*

**Lemma 3.7** ([KY20]). *Let $\Phi\colon\ \mathbb{M}_{\mathrm{n}}^{++} \to \mathbb{M}_{\mathrm{m}}^{++}$ be a strictly positive linear operator. Then we have*

$$\Phi(A!B) \leq \Phi(A)!\Phi(B)$$

*for positive $A, B \in \mathbb{M}_{\mathrm{n}}^{++}$.*

*proof of 3.5.* As $g$ is operator monotone increasing we have, from 3.6, that

$$g\Big(\Phi\big(f(A\triangledown B)\big)\Big) \leq g\Big(\Phi\big(f(A)!f(B)\big)\Big).$$

Since $g$ is operator monotone increasing we have, from 3.7, that

$$g\Big(\Phi\big(f(A)!f(B)\big)\Big) \leq g\Big(\Phi\big(f(A)\big)!\Phi\big(f(B)\big)\Big).$$

Since $g$ is operator monotone increasing the function $h(x) := g(x^{-1})$ is operator monotone decreasing on $(0,\infty)$ and is therefore operator convex. Thus, for $C, D \in \mathbb{M}_{\mathrm{n}}^{++}$, we have

$$g\big((C\triangledown D)^{-1}\big) = h(C\triangledown D) \leq h(C)\triangledown h(D) = g(C^{-1})\triangledown g(D^{-1}).$$

Using this we get

$$\begin{aligned} g\Big(\Phi\big(f(A)\big)!\Phi\big(f(B)\big)\Big) &= g\bigg(\Big(\Phi\big(f(A)\big)^{-1}\triangledown\Phi\big(f(B)\big)^{-1}\Big)^{-1}\bigg) \\ &\leq g\bigg(\Big(\Phi\big(f(A)\big)^{-1}\Big)^{-1}\bigg)\triangledown g\bigg(\Big(\Phi\big(f(B)\big)^{-1}\Big)^{-1}\bigg) \\ &= g\Big(\Phi\big(f(A)\big)\Big)\triangledown g\Big(\Phi\big(f(B)\big)\Big). \end{aligned}$$

In summary we have

$$g\Big(\Phi\big(f(A\triangledown B)\big)\Big) \leq g\Big(\Phi\big(f(A)\big)\Big)\triangledown g\Big(\Phi\big(f(B)\big)\Big),$$

and $A \mapsto g\Big(\Phi\big(f(A)\big)\Big)$ is convex. ■

**Theorem 3.8.** *Let $f_1, f_2\colon (0,\infty) \to (0,\infty)$ be operator monotone decreasing functions, $g(x)$ an operator monotone increasing function on $(0,\infty)$ such that $\lim_{x\to 0} g(x)x = 0$ and $-\widehat{g(x^{-1})}$ is operator monotone increasing on $(0,\infty)$, and let $\Phi\colon \mathbb{M}_\mathrm{n}^{++} \to \mathbb{M}_\mathrm{k}^{++}$, $\Psi\colon \mathbb{M}_\mathrm{m}^{++} \to \mathbb{M}_\mathrm{k}^{++}$ be strictly positive linear operators. Then the map*

$$\mathbb{M}_\mathrm{n}^{++} \times \mathbb{M}_\mathrm{m}^{++} \to \mathbb{R}, \quad (A,B) \mapsto \operatorname{Tr} g\Big(\Phi\big(f_1(A)\big)^{\frac{1}{2}}\Psi\big(f_2(B)\big)\Phi\big(f_1(A)\big)^{\frac{1}{2}}\Big)$$

*is jointly convex.*

*Proof.* Define the function $h$ by $h(x) := -g(x^{-1})$, which is operator monotone increasing. We have

$$\operatorname{Tr} g\Big(\Phi\big(f_1(A)\big)^{\frac{1}{2}}\Psi\big(f_2(B)\big)\Phi\big(f_1(A)\big)^{\frac{1}{2}}\Big) = -\operatorname{Tr} h\Big(\Phi\big(f_1(A)\big)^{-\frac{1}{2}}\Psi\big(f_2(B)\big)^{-1}\Phi\big(f_1(A)\big)^{-\frac{1}{2}}\Big),$$

so we show the concavity of $\operatorname{Tr} h\Big(\Phi\big(f_1(A)\big)^{-\frac{1}{2}}\Psi\big(f_2(B)\big)^{-1}\Phi\big(f_1(A)\big)^{-\frac{1}{2}}\Big)$ instead.
By 3.4 we have

$$\begin{aligned}
&\operatorname{Tr} h\Big(\Phi\big(f_1(A)\big)^{-\frac{1}{2}}\Psi\big(f_2(B)\big)^{-1}\Phi\big(f_1(A)\big)^{-\frac{1}{2}}\Big)\\
&= \inf_{Y\in\mathbb{M}_\mathrm{k}^{++}} \Big(\operatorname{Tr} Y\Phi\big(f_1(A)\big)^{-\frac{1}{2}}\Psi\big(f_2(B)\big)^{-1}\Phi\big(f_1(A)\big)^{-\frac{1}{2}} - \operatorname{Tr}\hat{h}(Y)\Big)\\
&= \inf_{Y\in\mathbb{M}_\mathrm{k}^{++}} \Big(\operatorname{Tr} Y\Phi\big(f_1(A)\big)^{-1}\Psi\big(f_2(B)\big)^{-1} - \operatorname{Tr}\hat{h}(Y)\Big).
\end{aligned}$$

Define the positive matrix $X := \Big(Y\Phi\big(f_1(A)\big)^{-1}\Big)^{\frac{1}{2}}$ which gives $Y = X^2\Phi\big(f_1(A)\big)$. This gives us

$$\begin{aligned}
&\inf_{Y\in\mathbb{M}_\mathrm{k}^{++}} \Big(\operatorname{Tr} Y\Phi\big(f_1(A)\big)^{-1}\Psi\big(f_2(B)\big)^{-1} - \operatorname{Tr}\hat{h}(Y)\Big)\\
&= \inf_{X\in\mathbb{M}_\mathrm{k}^{++}} \Big(\operatorname{Tr} X^2\Psi\big(f_2(B)\big)^{-1} - \operatorname{Tr}\hat{h}\Big(X^2\Phi\big(f_1(A)\big)\Big)\Big)
\end{aligned}$$

We look at the term $\operatorname{Tr} X^2\Psi\big(f_2(B)\big)^{-1}$ first. Define the strictly positive linear operator $\Psi'(Z) := \Psi(Z)X^{-2}$, so that $\Psi'(Z)^{-1} = X^2\Psi(Z)^{-1}$. We have

$$\operatorname{Tr} X^2\Psi\big(f_2(B)\big)^{-1} = \operatorname{Tr}\Psi'\big(f_2(B)\big)^{-1}.$$

Define the operator monotone increasing function $k$ by $k(x) := -x^{-1}$. We get

$$\operatorname{Tr}\Psi'\big(f_2(B)\big)^{-1} = -\operatorname{Tr} k\Big(\Psi'\big(f_2(B)\big)\Big),$$

which is concave by 3.5.
Now we look at the term $-\operatorname{Tr}\hat{h}\Big(X^2\Phi\big(f_1(A)\big)\Big)$. Define the strictly positive linear operator $\Phi'(Z) := X^2\Phi(Z)$. This gives

$$-\operatorname{Tr}\hat{h}\Big(X^2\Phi\big(f_1(A)\big)\Big) = -\operatorname{Tr}\hat{h}\Big(\Phi'\big(f_1(A)\big)\Big).$$

By assumption $\hat{h}(x) = -\widehat{g(x^{-1})}$ is operator monotone increasing. The term $-\operatorname{Tr}\hat{h}\Big(\Phi'\big(f_1(A)\big)\Big)$ is therefore, by 3.5, operator concave.

Let $Z$ almost minimize $\inf_{X\in\mathbb{M}_k^{++}}\Big(\operatorname{Tr}X^2\Psi\big(f_2(B)\big)^{-1} - \operatorname{Tr}\hat{h}\Big(X^2\Phi\big(f_1(A)\big)\Big)\Big)$ for $A = A_1\nabla A_2$ and $B = B_1\nabla B_2$, and let $\varepsilon > 0$. Inserting $Z$ in the expression and using concavity we get

$$
\begin{aligned}
&\inf_{X\in\mathbb{M}_k^{++}}\Big(\operatorname{Tr}X^2\Psi\big(f_2(B)\big)^{-1} - \operatorname{Tr}\hat{h}\Big(X^2\Phi\big(f_1(A)\big)\Big)\Big) + \varepsilon\\
&\geq \operatorname{Tr}Z^2\Psi\big(f_2(B)\big)^{-1} - \operatorname{Tr}\hat{h}\Big(Z^2\Phi\big(f_1(A)\big)\Big)\\
&\geq \frac{1}{2}\Big(\operatorname{Tr}Z^2\Psi\big(f_2(B_1)\big)^{-1} + \operatorname{Tr}Z^2\Psi\big(f_2(B_2)\big)^{-1}\\
&\quad - \operatorname{Tr}\hat{h}\Big(Z^2\Phi\big(f_1(A_1)\big)\Big) - \operatorname{Tr}\hat{h}\Big(Z^2\Phi\big(f_1(A_2)\big)\Big)\Big)\\
&\geq \frac{1}{2}\Big(\inf_{Z_1\in\mathbb{M}_k^{++}}\Big(\operatorname{Tr}Z_1^2\Psi\big(f_2(B_1)\big)^{-1} - \operatorname{Tr}\hat{h}\Big(Z_1^2\Phi\big(f_1(A_1)\big)\Big)\Big)\\
&\quad + \inf_{Z_2\in\mathbb{M}_k^{++}}\Big(\operatorname{Tr}Z_2^2\Psi\big(f_2(B_2)\big)^{-1} - \operatorname{Tr}\hat{h}\Big(Z_2^2\Phi\big(f_1(A_2)\big)\Big)\Big)\Big)\\
&= \frac{1}{2}\Big(\operatorname{Tr}h\Big(\Phi\big(f_1(A_1)\big)^{-\frac{1}{2}}\Psi\big(f_2(B_1)\big)^{-1}\Phi\big(f_1(A_1)\big)^{-\frac{1}{2}}\Big)\\
&\quad + \operatorname{Tr}h\Big(\Phi\big(f_1(A_2)\big)^{-\frac{1}{2}}\Psi\big(f_2(B_2)\big)^{-1}\Phi\big(f_1(A_2)\big)^{-\frac{1}{2}}\Big)\Big),
\end{aligned}
$$

By letting $\varepsilon$ go to zero we thus get joint concavity of

$$
\operatorname{Tr}h\Big(\Phi\big(f_1(A)\big)^{-\frac{1}{2}}\Psi\big(f_2(B)\big)^{-1}\Phi\big(f_1(A)\big)^{-\frac{1}{2}}\Big).
$$

■

Hiai showed in [Hia16] the following theorem:

**Theorem 3.9** ([Hia16])**.** *Let $g$ be a non-increasing function on $(0,\infty)$, $\Phi\colon \mathbb{M}_n^{++}\to\mathbb{M}_k^{++}$, $\Psi\colon \mathbb{M}_m^{++}\to\mathbb{M}_k^{++}$ be strictly positive linear operators and $p,q\in[0,1]$. If either $g(x^{1+p})$ or $g(x^{1+q})$ is convex on $(0,\infty)$ then the maps*

$$
\mathbb{M}_n^{++}\times\mathbb{M}_m^{++}\to\mathbb{R}\ ,\quad (A,B)\mapsto \operatorname{Tr}g\big(\Phi(A^p)^{\frac{1}{2}}\Psi(B^q)\Phi(A^p)^{\frac{1}{2}}\big)
$$

$$
\mathbb{M}_n^{++}\times\mathbb{M}_m^{++}\to\mathbb{R},\quad (A,B)\mapsto \operatorname{Tr}g\Big(\big(\Phi(A^{-p})^{\frac{1}{2}}\Psi(B^{-q})\Phi(A^{-p})^{\frac{1}{2}}\big)^{-1}\Big)
$$

*are jointly convex.*

We see that the map $(A,B)\mapsto \operatorname{Tr}g\Big(\Phi\big(f_1(A)\big)^{\frac{1}{2}}\Psi\big(f_2(B)\big)\Phi\big(f_1(A)\big)^{\frac{1}{2}}\Big)$ in 3.8 is more general than the maps in 3.9, but that this comes at the cost of assuming $g(x)$ and $-\widehat{g(x^{-1})}$ are operator monotone increasing instead of $g(x^{1+p})$ or $g(x^{1+q})$ being convex.

**Example 3.10.** With $g(x) = x^{\frac{1}{2}}$, $f_1(x) = f_2(x) = x^{-1}$ and $\Phi(A) = \Psi(A) = A$ we get that

$$\begin{aligned}\operatorname{Tr} g\Big(\Phi\big(f_1(A)\big)^{\frac{1}{2}}\Psi\big(f_2(B)\big)\Phi\big(f_1(A)\big)^{\frac{1}{2}}\Big) &= \operatorname{Tr}\big((A^{-1})^{\frac{1}{2}}B^{-1}(A^{-1})^{\frac{1}{2}}\big)^{\frac{1}{2}} \\ &= \operatorname{Tr}(BA)^{-\frac{1}{2}}\end{aligned}$$

is jointly convex.

**Example 3.11.** With $g(x) = \log(x)$, $f_1(x) = f_2(x) = x^{-\frac{1}{2}}$, $\Phi(A) = XA$ and $\Psi(A) = YB$ for some $X \in \mathbb{M}_{\mathrm{n}}^{++}$, $Y \in \mathbb{M}_{\mathrm{m}}^{++}$ we get that

$$\operatorname{Tr} g\Big(\Phi\big(f_1(A)\big)^{\frac{1}{2}}\Psi\big(f_2(B)\big)\Phi\big(f_1(A)\big)^{\frac{1}{2}}\Big) = \operatorname{Tr}\log\big((XA^{-\frac{1}{2}})^{\frac{1}{2}}YB^{-\frac{1}{2}}(XA^{-\frac{1}{2}})^{\frac{1}{2}}\big)$$

is jointly convex.

## 4 Closeness to convexity.

We saw earlier that the function $f(x) = x^3$, although convex on $(0, \infty)$ is not operator convex. This only means that the inequality $f\left(\frac{A+B}{2}\right) \leq \frac{f(A)+f(B)}{2}$ does not hold for general positive matrices $A$ and $B$, not that it breaks for all $A$ and $B$. It is therefore possible to develop a notion of "closeness to convexity" by asking the following question: If we generate random positive matrices $A$ and $B$, how often does the convexity inequality hold? First we need a way to generate random matrices. We will use the Gaussian Unitary Ensemble.

**Definition 4.1.** Let $A \in \mathbb{M}_{\mathrm{n}}$ have the entries

$$A_{k,l} := \frac{a_{k,l}^1 + a_{k,l}^2 i}{\sqrt{2n}}, \quad 1 \leq k < l \leq n$$

above the main diagonal, and

$$A_{k,k} := \frac{a_{k,k}}{\sqrt{n}}, \quad 1 \leq k \leq n$$

on the main diagonal, where $a_{k,l}^1$, $a_{k,l}^2$ and $a_{k,k}$ are independent, centered Gaussian variables with variance 1. Then $A$ is called the *Gaussian Unitary Ensemble.*

A matrix $A$ generated this way isn't necessarily positive, but we can define the matrix $B$ by

$$B := (A^2)^{\frac{1}{2}}$$

and use this instead.

**The cube function.** With $f(x) = x^3$, $A, B$ randomly generated positive matrices and checking the convexity inequality we get for dimensions $n = 2, \ldots, 10$ with 10000 trials the graph in 0.1. We see that the convexity inequality holds more often than not in dimension 2, but drops off quickly as the dimension increases.

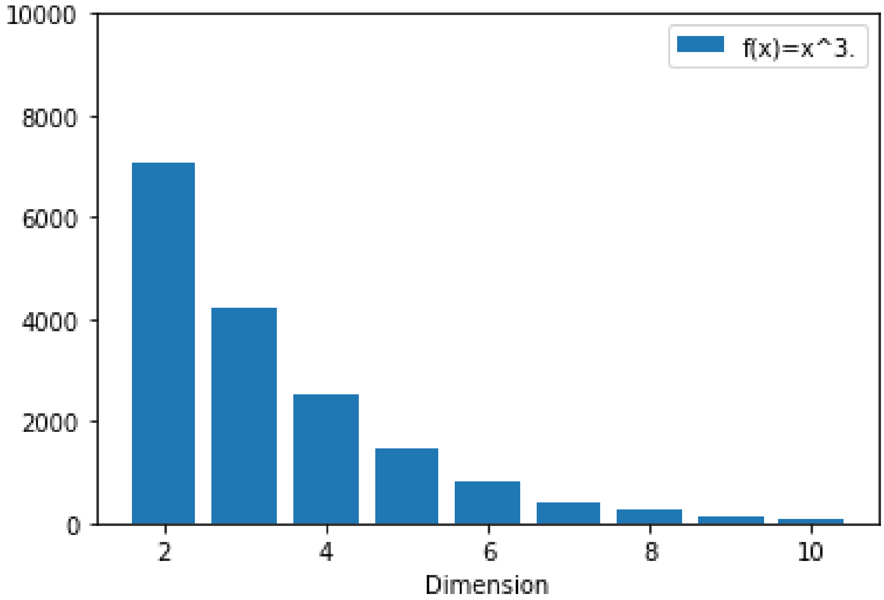


Figure 0.1: The cube function.

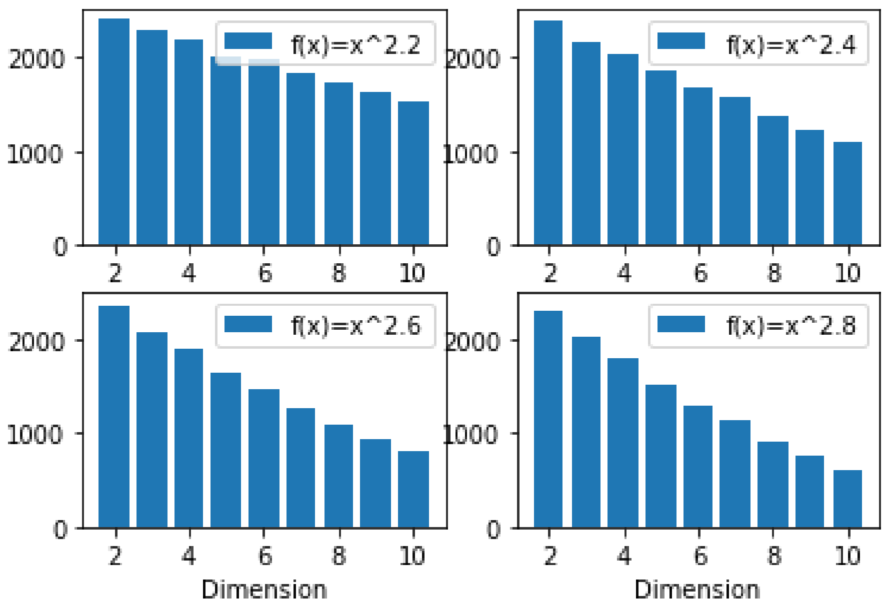


Figure 0.2: Some power functions.

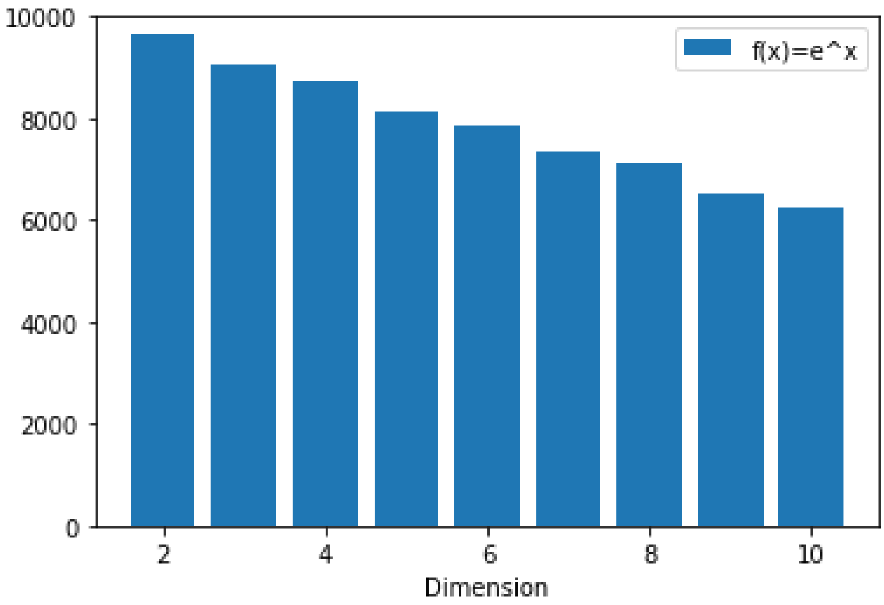


Figure 0.3: The exponential function.

**Some power functions.** The function $f(x) = x^3$ may not be operator convex, but $g(x) = x^2$ is. The graph in 0.2 shows the power function $h(x) = x^r$ for some values between 2 and 3 with 2500 trials. We see that the closeness to convexity drops off in every dimension as $r$ increases.

**The exponential function.** The exponential function $f(x) = e^x$ is convex as a real function, but it is not operator convex. The graph in 0.3 shows that it starts off close to convex for dimension 2 and drops off slowly as the dimension increases.

# 5 Appendix.

Code used to produce 0.1, 0.2 and 0.3:

```
import numpy as np
import scipy as sp
from scipy.linalg import fractional_matrix_power, expm, eigvals
from math import sqrt
import scipy.stats as stats
import matplotlib.pyplot as pp


def GUE(N):
    A = np.matrix([[stats.norm.rvs(loc=0, scale=1)/sqrt(2*N)
        + stats.norm.rvs(loc=0, scale=1)*1j/sqrt(2*N)
        for i in range(N)] for j in range(N)])
    for i in range(N):
        for j in range(N):
            if i<j:
                A[i, j] = np.conj(A[j, i])
            if i==j:
                A[i, j] = stats.norm.rvs(loc=0, scale=1)/sqrt(N)
    A = fractional_matrix_power(A**2, .5)
    return A


# f(x) = x^3

n = 10
its = 10000
dim_size = range(2, n+1)

convex_plot = []
for N in dim_size:
    convex_counter = 0
    for i in range(its):
        A = GUE(N)
        B = GUE(N)
        if min(eigvals((A**3+B**3)/2-((A+B)/2)**3)) >= 0:
            convex_counter += 1
    convex_plot.append(convex_counter)
pp.ylim([0, its])
pp.xlabel("Dimension")
pp.bar(dim_size, convex_plot)
pp.legend(["f(x)=x^3."])


# f(x) = e^x

n = 10
```

```
its = 10000
dim_size = range(2, n+1)

convex_plot = []
for N in dim_size:
    convex_counter = 0
    for i in range(its):
        A = GUE(N)
        B = GUE(N)
        if min(eigvals((expm(A)+expm(B))/2-expm((A+B)/2))) >= 0:
            convex_counter += 1
    convex_plot.append(convex_counter)
pp.ylim([0, its])
pp.xlabel("Dimension")
pp.bar(dim_size, convex_plot)
pp.legend(["f(x)=e^x"])


# f(x) = x^r

n = 10
its = 2500
exponents = [2.2, 2.4, 2.6, 2.8]
dim_size = range(2, n+1)

fig, axs = pp.subplots(2, 2)
first_axis = 0
second_axis = 0

for exp in exponents:
    def f(A):
        return fractional_matrix_power(A, exp)

    convex_plot = []
    for N in dim_size:
        convex_counter = 0
        for i in range(its):
            A = GUE(N)
            B = GUE(N)
            if min(eigvals((f(A)+f(B))/2-f((A+B)/2))) >= 0:
                convex_counter += 1
        convex_plot.append(convex_counter)

    axs[first_axis, second_axis].bar(dim_size, convex_plot)
    axs[first_axis, second_axis].legend(["f(x)=x^"+str(exp)])
    axs[first_axis, second_axis].set(ylim=[0, its])
    axs[first_axis, second_axis].set(xlabel="Dimension")

    if second_axis <= first_axis:
        second_axis += 1
```

```
    else:
        second_axis -= 1
        first_axis += 1
```

# Bibliography


[AH11] Ando, T. and Hiai, F. "Operator log-convex functions and operator means". In: *Mathematische Annalen* vol. 350, no. 3 (2011), pp. 611–630.

[Bha07] Bhatia, R. *Positive definite matrices*. Princeton university press, 2007.

[Bha97] Bhatia, R. *Matrix analysis*. Vol. 169. Springer Science & Business Media, 1997.

[CFL16] Carlen, E. A., Frank, R. L., and Lieb, E. H. "Some operator and trace function convexity theorems". In: *Linear Algebra and its Applications* vol. 490 (2016), pp. 174–185.

[Hia16] Hiai, F. "Concavity of certain matrix trace and norm functions. II". In: *Linear Algebra and its Applications* vol. 496 (2016), pp. 193–220.

[KY20] Kirihata, M. and Yamashita, M. "Strengthened convexity of positive operator monotone decreasing functions". In: *MATHEMATICA SCANDINAVICA* vol. 126, no. 3 (2020), pp. 559–567.

[Lie73] Lieb, E. H. "Convex trace functions and the Wigner-Yanase-Dyson conjecture". In: *Les rencontres physiciens-mathématiciens de Strasbourg-RCP25* vol. 19 (1973), pp. 0–35.

[Rus02] Ruskai, M. B. "Inequalities for quantum entropy: A review with conditions for equality". In: *Journal of Mathematical Physics* vol. 43, no. 9 (2002), pp. 4358–4375.