Operator Jensen’s Inequality for Operator Superquadratic Functions

Mohammad W. Alomari 1, Christophe Chesneau 2,* and Ahmad Al-Khasawneh 3,4

1 Department of Mathematics, Faculty of Science and Information Technology, Irbid National University, Irbid 21110, Jordan
2 Department of Mathematics, Université de Caen Basse-Normandie, F-14032 Caen, France
3 Department of Information Technology, Faculty of Prince Al-Hussein Bin Abdullah II for Information Technology, The Hashemite University, Zanqa 13133, Jordan
4 Department of Cyber Security, Faculty of Science and Information Technology, Irbid National University, Irbid 21110, Jordan
* Correspondence: christophe.chesneau@unicaen.fr

Abstract: In this work, an operator superquadratic function (in the operator sense) for positive Hilbert space operators is defined. Several examples with some important properties together with some observations which are related to the operator convexity are pointed out. A general Bohr’s inequality for positive operators is thus deduced. A Jensen-type inequality is proved. Equivalent statements of a non-commutative version of Jensen’s inequality for operator superquadratic function are also established. Finally, several trace inequalities for superquadratic functions (in the ordinary sense) are provided as well.

Keywords: operator superquadratic; operator convex; self-adjoint; Jensen’s inequality; trace

MSC: 47A63; 47A56

1. Introduction

Let \( \mathcal{B}(\mathcal{H}) \) be the Banach algebra of all bounded linear operators defined on a complex Hilbert space \((\mathcal{H}, \langle \cdot, \cdot \rangle)\) with the identity operator \(1_\mathcal{H}\) in \(\mathcal{B}(\mathcal{H})\). A bounded linear operator \(A\) defined on \(\mathcal{H}\) is self-adjoint if \(\langle Ax, x \rangle \in \mathbb{R}\) for all \(x \in \mathcal{H}\). The spectrum of an operator \(A\) is the set of all \(\lambda \in \mathbb{C}\) for which the operator \(A_\mathcal{H} - A\) does not have a bounded linear inverse operator, and is denoted by \(\text{sp}(A)\). Consider the real vector space \(\mathcal{B}(\mathcal{H})_{\text{sa}}\) of self-adjoint operators on \(\mathcal{H}\) and its positive cone \(\mathcal{B}(\mathcal{H})^+\) of positive operators on \(\mathcal{H}\). Additionally, \(\mathcal{B}(\mathcal{H})^+_{\text{sa}}\) denotes the convex set of bounded self-adjoint operators on the Hilbert space \(\mathcal{H}\) with spectra in a real interval \(J\). A partial order is naturally equipped on \(\mathcal{B}(\mathcal{H})_{\text{sa}}\) by defining \(A \preceq B\) if and only if \(B - A \in \mathcal{B}(\mathcal{H})^+\). We write \(A \preceq B\) to mean that \(A\) is a strictly positive operator, or equivalently, \(A \succeq 0\) and \(A\) is invertible. When \(\mathcal{H} = \mathbb{C}^n\), we identify \(\mathcal{B}(\mathcal{H})\) with the algebra \(\mathfrak{M}_{n \times n}\) of \(n\)-by-\(n\) complex matrices. Then, \(\mathfrak{M}_{n \times n}^+\) is just the cone of \(n\)-by-\(n\) positive semidefinite matrices.

A linear map is defined to be \(\Phi: \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})\), which preserves additivity and homogeneity, that is, \(\Phi(\lambda_1 A + \lambda_2 B) = \lambda_1 \Phi(A) + \lambda_2 \Phi(B)\) for any \(\lambda_1, \lambda_2 \in \mathbb{C}\) and \(A, B \in \mathcal{B}(\mathcal{H})\). A linear map is positive \(\Phi: \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})\) if it preserves the order relation, that is, if \(A \in \mathcal{B}^+(\mathcal{H})\) then \(\Phi(A) \in \mathcal{B}^+(\mathcal{H})\), and in this case we write \(\Phi|_{\mathcal{B}(\mathcal{H})}\). Obviously, a positive linear map \(\Phi\) preserves the order relation, namely \(A \preceq B\) implies that \(\Phi(A) \preceq \Phi(B)\) and preserves the adjoint operation \(\Phi(A^*) = (\Phi(A))^*\). Moreover, \(\Phi\) is said to be unital if it preserves the identity operator, in this case, we write \(\Phi|_{\mathfrak{M}_{n \times n}(\mathcal{H})}\).

A linear map \(\Phi: \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})\) induces another map

\[
\id \otimes \Phi: C^{k \times k} \otimes \mathcal{B}(\mathcal{H}) \to C^{k \times k} \otimes \mathcal{B}(\mathcal{H}),
\]
in a natural way. If $\mathbb{C}^{k \times k} \otimes \mathcal{B}(\mathcal{H})$ is identified with the $C^*$-algebra $\mathcal{B}^{k \times k}(\mathcal{H})$ of $k \times k$-matrices with entries in $\mathcal{B}(\mathcal{H})$ then $\text{id} \otimes \Phi$ act as:

$$
\begin{pmatrix}
A_{11} & \cdots & A_{1k} \\
\vdots & \ddots & \vdots \\
A_{k1} & \cdots & A_{kk}
\end{pmatrix} \mapsto \begin{pmatrix}
\Phi(A_{11}) & \cdots & \Phi(A_{1k}) \\
\vdots & \ddots & \vdots \\
\Phi(A_{k1}) & \cdots & \Phi(A_{kk})
\end{pmatrix}.
$$

We say that $\Phi$ is $k$-positive if $\text{id} \otimes \Phi$ is a positive map, and $\Phi$ is called completely positive if $\Phi$ is $k$-positive for all $k$.

1.1. Superquadratic Functions

A function $f: I \to \mathbb{R}$ is called convex if

$$
f(ta + (1-t)\beta) \leq tf(a) + (1-t)f(\beta),
$$

for all points $a, \beta \in I$ and all $t \in [0,1]$. If $-f$ is convex then we say that $f$ is concave. Moreover, if $f$ is both convex and concave, then $f$ is said to be affine.

Geometrically, for two points $(x, f(x))$ and $(y, f(y))$ on the graph of $f$ are on or below the chord joining the endpoints for all $x, y \in I, x < y$. In symbols, we write

$$
f(t) \leq \frac{f(y) - f(x)}{y-x} (t-x) + f(x)
$$

for any $x \leq t \leq y$ and $x, y \in I$.

Equivalently, given a function $f: I \to \mathbb{R}$, we say that $f$ admits a support line at $x \in I$ if there exists a $\lambda \in \mathbb{R}$ such that

$$
f(t) \geq f(x) + \lambda(t-x)
$$

for all $t \in I$.

The set of all such $\lambda$ is called the subdifferential of $f$ at $x$, and it’s denoted by $\partial f$. Indeed, the subdifferential gives us the slopes of the supporting lines for the graph of $f$. So that if $f$ is convex then $\partial f(x) \neq \emptyset$ at all interior points of its domain.

From this point of view, Abramovich et al. [1] extend the above idea for what they called superquadratic functions. Namely, a function $f: [0, \infty) \to \mathbb{R}$ is called superquadratic provided that for all $x \geq 0$ there exists a constant $C_x \in \mathbb{R}$ such that

$$
f(t) \geq f(x) + C_x(t-x) + f(|t-x|)
$$

for all $t \geq 0$. We say that $f$ is subquadratic if $-f$ is superquadratic. Thus, for a superquadratic function, we require that $f$ lie above its tangent line plus a translation of $f$ itself. If $f$ is differentiable and satisfies $f(0) = f'(0) = 0$, then we know easily that the $C_x$ appearing in the definition is necessarily $f'(x)$ (see [2]).

At first glance, the superquadratic function looks to be stronger than the convex function itself, but if $f$ takes negative values then it may be considered a weaker function. Therefore, if $f$ is superquadratic and non-negative, then $f$ is convex and increasing [1] (see also [3]).

Moreover, the following result holds for superquadratic functions.

**Lemma 1** ([1]). Let $f$ be a superquadratic function. Then

1. $f(0) \leq 0$.
2. if $f$ is differentiable and $f(0) = f'(0) = 0$, then $C_x = f'(x)$ for all $x \geq 0$.
3. if $f(x) \geq 0$ for all $x \geq 0$, then $f$ is convex and $f(0) = f'(0) = 0$. 
The next result gives a sufficient condition when convexity (concavity) implies super(sub)quadraticity.

**Lemma 2 ([1]).** If \( f' \) is convex (concave) and \( f(0) = f'(0) = 0 \), then \( f \) is super(sub)quadratic. The converse is not true.

**Remark 1.** In general, non-negative subquadratic functions do not imply concavity. In other words, there exists a subquadratic function that is convex. For example, \( f(x) = x^p, x \geq 0 \) and \( 1 \leq p \leq 2 \) is subquadratic and convex.

Among others, Abramovich et al. [1] proved that the inequality

\[
\int \varphi \, d\mu \leq \int \varphi(s) - f \left( \frac{\varphi(s) - \int \varphi \, d\mu}{s - \int \varphi \, d\mu} \right) \, d\mu(s)
\]

holds for all probability measures \( \mu \) and all non-negative, \( \mu \)-integrable functions \( \varphi \) if and only if \( f \) is superquadratic. For more details the reader may refer to [3–6].

### 1.2. Operator Convexity and Jensen’s Inequality

Let \( f \) be a real-valued function defined on \( J \). A \( k \)-th order divided difference of \( f \) at distinct points \( x_0, \ldots, x_k \) in \( J \) may be defined recursively by

\[
[x_0]f = f(x_0) \quad \text{if } \ k = 0,
\]

\[
[x_0, x_1]f = \frac{[x_1]f - [x_0]f}{x_1 - x_0} \quad \text{if } \ k = 1,
\]

\[
[x_0, x_1, x_2]f = \frac{[x_2]f - [x_1]f}{x_2 - x_0} \quad \text{if } \ k = 2.
\]

For instance, the first three divided differences are given as follows:

\[
[x_0]f = f(x_0)
\]

\[
[x_0, x_1]f = \frac{[x_1]f - [x_0]f}{x_1 - x_0}
\]

A function \( f : J \to \mathbb{R} \) is said to be matrix monotone of degree \( n \) or \( n \)-monotone, if for every \( A, B \in \mathcal{M}_{n \times n} \), it is true that \( A \leq B \) if and only if \( f(A) \leq f(B) \). Similarly, \( f \) is said to be operator monotone if \( f \) is \( n \)-monotone for all \( n \in \mathbb{N} \). Additionally, \( f \) is called operator convex if it is matrix convex (\( n \)-convex for all \( n \)); that is, if for every pair of self-adjoint operators \( A, B \in \mathcal{M}_{n \times n} \), we have

\[
f(\lambda A + (1 - \lambda)B) \leq \lambda f(A) + (1 - \lambda)f(B)
\]

for all \( \lambda \in [0, 1] \). If the inequality is reversed then \( f \) is called operator concave. In case we have general Hilbert space \( \mathcal{H} \), the above definition holds for every pair of bounded self-adjoint operators \( A \) and \( B \) in \( \mathcal{B}(\mathcal{H}) \), whose spectra obtained in \( f \). For more details, see [7] and the recent survey [8].

In 1955, Bendat and Sherman [9] have shown that \( f \) is operator convex on the open interval \((-1, 1)\) if and only if it has the following (unique) representation:

\[
f(t) = \beta_0 + \beta_1 t + \frac{1}{2} \beta_2 \int_{-1}^{1} \frac{t^2}{1 - \lambda t} \, d\mu(\lambda)
\]

for \( \beta_2 \geq 0 \) and some probability measure \( \mu \) on \([-1, 1]\) (it could be the Borel measure). In particular, \( f \) must be analytic with \( f(0) = \beta_0, f'(0) = \beta_1 \) and \( f''(0) = \beta_2 \).

We recall that the celebrated Löwner–Heinz inequality reads that:
Lemma 3. Let $A, B \in \mathcal{B}(\mathcal{H})^+$ such that $A \geq B$. Then $A^r \geq B^r$ for all $r \in [0, 1]$.

On the other hand the mapping $t \mapsto t^p$ ($p > 1$) is not operator monotone, for more details, see [10–12].

The classical Jensen’s inequality can be formulated as

$$f\left(\sum_{j=1}^{n} \lambda_j x_j\right) \leq \sum_{j=1}^{n} \lambda_j f(x_j)$$

(5)

is valid for all real-valued convex functions $f$ defined on $[m, M]$, for every $x_1, \cdots, x_n \in [m, M]$ and every positive real number $\lambda_j$ ($1 \leq j \leq n$) such that $\sum_{j=1}^{n} \lambda_j = 1$.

The inequality (5) would be rephrased under the matrix situation by putting

$$A = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & x_n \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} \sqrt{\lambda_1} \\ \vdots \\ \sqrt{\lambda_n} \end{pmatrix}$$

then the classical Jensen’s inequality (5) is expressed as

$$f(\langle Ax, x \rangle) \leq (f(A) x, x),$$

(6)

which is one of the operator versions of the classical Jensen’s inequality, see [11,13], the recent monograph [14], as well as [15–18].

Kadison [19] established his famous non-commutative version of the previous inequality, where he proved that, for every self-adjoint matrix $A$, the inequality

$$\Phi^2(A) \leq \Phi\left(A^2\right)$$

(7)

for every positive unital linear map $\Phi : \mathcal{M}_{m \times n}(\mathbb{C}) \to \mathcal{M}_{k \times k}(\mathbb{C})$.

This inequality was generalized later by Davis in [20], where he obtained that this is true when $f$ is a matrix convex function and $\Phi$ is completely positive, that is,

$$f(\Phi(A)) \leq \Phi(f(A)).$$

(8)

The latter restriction about complete positivity of $\Phi$ was removed by Choi [21] who proved that (7) remains valid for all positive unital linear maps provided $f$ is matrix convex.

Another non-commutative operator version of the classical Jensen’s inequality under the situation that

$$A = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & x_n \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} \sqrt{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix},$$

the classic Jensen’s inequality is expressed as

$$f(V^*AV) \leq V^*f(A)V.$$
(2) \( f(C^*AC) \leq C^*f(A)C \), for every \( A \in \mathcal{B}(\mathcal{H}) \) and contraction \( C \in \mathcal{B}(\mathcal{H}) \), that is, \( C^*C \leq 1_{\mathcal{H}} \).

(3) \( f\left( \sum_{j=1}^{n} C_j^*A_jC_j \right) \leq \sum_{j=1}^{n} C_j^*f(A_j)C_j \), for all \( A_j \in \mathcal{B}(\mathcal{H}) \) and \( C_j \in \mathcal{B}(\mathcal{H}) \) with \( \sum_{j=1}^{n} C_j^*C_j \leq 1_{\mathcal{H}} \) (\( j = 1, 2, \ldots, k \)).

(4) \( f(PAP) \leq Pf(A)P \), for every \( A \in \mathcal{B}(\mathcal{H}) \) and projection \( P \).

Here, we give some popular examples of operator convex and concave function [6].

(1) For each \( p \in [0, 1] \), \( t^p \) is operator concave on \([0, \infty)\).

(2) The function \( t \log t \) is operator convex on \([0, \infty)\).

This work is organized as follows: after this introduction; in Section 2, the operator superquadratic functions for positive Hilbert space operators are introduced and elaborated. Several examples with some important properties together with some observations related to operator convexity are pointed out. A general Bohr’s inequality for positive operators is thus deduced. A Jensen’s type inequality is proved in Section 3. Equivalent statements to operator convexity are pointed out. A general Bohr’s inequality for positive operators are also established. Finally, several trace inequalities for superquadratic functions (in the ordinary sense) are provided as well in Section 4.

2. Operator Superquadratic Function

Definition 1. Let \( J = [0, M] \subseteq [0, \infty) \). A real-valued continuous function \( f(t) \) on an interval \( J \) is said to be an operator superquadratic function if

\[
    f(\alpha A + (1 - \alpha)B) \leq \alpha f(A) + f((1 - \alpha)|A - B|) + (1 - \alpha)f(B) - f(\alpha|A - B|) \tag{10}
\]

holds for all \( \alpha \in [0, 1] \) and for every positive operators \( A \) and \( B \) on a Hilbert space \( \mathcal{H} \) whose spectra are contained in \( J \subset [0, \infty) \). We say that \( f \) is an operator subquadratic function if \(-f\) is an operator superquadratic function. Moreover, if the equality holds in (10), we say that \( f \) is an operator quadratic function.

It is convenient to note that if \( f \) satisfies (10), then with \( A = x \) and \( B = y \) (two positive scalars), we can obtain the Jensen’s inequality for superquadratic functions and if \( f \) is continuous (which is necessary to define an operator functions), then (4) would imply that \( f \) is a superquadratic function. Thus, we observe that:

Corollary 1. If \( f \) is an operator superquadratic function then \( f \) is a real superquadratic function.

Let \( f(t) = \alpha t + \beta \). Then \( f \) is operator subquadratic on every bounded interval for all \( \alpha, \beta \geq 0 \). Indeed, we have

\[
    f\left( \frac{A + B}{2} \right) + f\left( \frac{|A - B|}{2} \right) - \frac{f(A) + f(B)}{2} = \frac{\alpha A + B}{2} + \beta + \left[ \frac{\alpha A}{2} + \frac{\alpha B}{2} + \beta \right] - \frac{\alpha A + \beta + \alpha B + \beta}{2} = \alpha \frac{|A - B|}{2} + \beta \geq 0.
\]

Moreover, \( g(t) = -f(t) \) is operator superquadratic.

We can easily show that the function \( t \mapsto t^3 \) is not an operator superquadratic nor an operator subquadratic function. Simply, assume \( f(t) = t^3, t \in [0, \infty) \), and let

\[
    A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.
\]
we have

\[ \frac{A^3 + B^3}{2} - \left( \frac{A + B}{2} \right)^3 - \left( \frac{|A - B|}{2} \right)^3 = \frac{1}{4} \left( 9 \ 7 \ 5 \right) \leq 0. \]

However, the map \( t \mapsto t^2 \) is a non-negative operator convex on \((0, \infty)\) and it is also operator superquadratic on \((0, \infty)\). Indeed, by (10), we have

\[
(aA + (1 - a)B)^2 \leq aA^2 + (1 - a)B^2 - a(1 - a)^2 |A - B|^2 \\
\iff a^2 A^2 + (1 - a)^2 B^2 + a(1 - a)(AB + BA) \leq aA^2 + (1 - a)B^2 \\
- \left[a(1 - a)^2 + (1 - a)\right](A - B)^2 \\
\iff a(a - 1)A^2 + a(a - 1)B^2 + a(a - 1)(AB + BA) \leq a(a - 1)(A - B)^2 \\
\iff (A + B)^2 \geq (A - B)^2 \quad \text{for } a(a - 1) < 0 \\
\Rightarrow |A + B| \geq |A - B| \quad g(t) = \sqrt{t} \text{ is operator monotone,}
\]

which is true since \( A, B > 0 \), and this proves that \( t^2 \) is an operator superquadratic function.

From the definition of the operator superquadratic function, we have

\[
f(aA + (1 - a)B) \leq a[f(A) - f((1 - a)|A - B|)] + (1 - a)[f(B) - f(a|A - B|)] \quad (11)
\]

for any arbitrary positive operators \( A, B \in \mathcal{B}(\mathcal{H}) \) and each \( a \in [0, 1] \).

In particular, by setting \( B = \langle Ax, x \rangle 1_{\mathcal{P}} \) in (10), we have

\[
f(aA + (1 - a)\langle Ax, x \rangle 1_{\mathcal{P}}) \leq a[f(A) - f((1 - a)|A - \langle Ax, x \rangle 1_{\mathcal{P}}|)] \\
+ (1 - a)[f(\langle Ax, x \rangle) - f(a|A - \langle Ax, x \rangle 1_{\mathcal{P}}|)] \quad (12)
\]

for each positive operator \( A \in \mathcal{B}(\mathcal{H}) \) and all \( a \in [0, 1] \).

From this point of view (12), Kian early in [22] and then jointly with Dragomir in [23] proved a finite dimensional operator version of Jensen’s inequality for superquadratic functions (in the ordinary sense) under the interpretation that for \( A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \) and \( x = \begin{pmatrix} \sqrt{\lambda} \\ \sqrt{1 - \lambda} \end{pmatrix} \), then we have \( \langle Ax, x \rangle = \lambda a + (1 - \lambda)b \) and it follows that

\[
|A - \langle Ax, x \rangle| = \begin{pmatrix} (1 - \lambda)|a - b| & 0 \\ 0 & \lambda |a - b| \end{pmatrix}.
\]

Therefore, as a matrix Jensen’s inequality for a superquadratic function \( f : [0, \infty) \to \mathbb{R} \) we have

\[
f(\langle Ax, x \rangle) \leq f(A)x, x) + f(|A - \langle Ax, x \rangle|)x, x).
\]

This result was generalized for positive unital linear maps, as follows:

**Theorem 2** ([23]). Let \( A \in \mathcal{B}(\mathcal{H}) \) be a positive operator and \( \Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}) \) be a positive unital linear map. If \( f : [0, \infty) \to \mathbb{R} \) is a super(sub)quadratic function, then we have

\[
\langle \Phi(f(A))x, x \rangle \geq \langle f(|\Phi(A)x, x \rangle) + \langle \Phi(f(|A - \Phi(A)x, x \rangle 1_{\mathcal{P}})|)x, x \rangle.
\]

for every \( x \in \mathcal{H} \) with \( \|x\| = 1 \).
The above inequality and other consequences were proved later by the first author of this paper in [24], where a different approach is used.

**Proposition 1.** Let $f$ be an operator superquadratic function on $J$. Then

1. $f(0) \leq 0$.
2. If $f$ is non-negative, then $f$ is operator convex and $f(0) = 0$.

**Proof.**

1. Setting $A = B = 0$ in (1), we get that $f(0) \leq 0$.
2. Since $f$ is continuous and non-negative, from (1), we have

$$f(\alpha A + (1-\alpha)B) \leq \alpha f(A) - f((1-\alpha)|A-B|) + (1-\alpha)[f(B) - f(\alpha|A-B|)]$$

which means that $f$ is operator convex. To show that $f(0) = 0$, we have by the part (1) $f(0) \leq 0$ and by the assumption $f(x)$ is non-negative, that is, $f(x) \geq 0$ for all $x \in J$. In particular, we have $f(0) \geq 0$. Thus, $f(0) = 0$. \[Q.E.D.\]

**Example 1.** Let $f(t) = t^{-1}$. Then $f$ is non-negative operator convex on $(0, \infty)$. However, $f$ is not an operator superquadratic function on $(0, \infty)$. For instance, let

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$ 

Applying (13) for $f(t) = t^{-1}$, we get

$$\begin{align*}
\frac{A^{-1} + B^{-1}}{2} - \left( \frac{A + B}{2} \right)^{-1} - \left( \frac{|A-B|}{2} \right)^{-1} &= \frac{1}{12} \begin{pmatrix} 8 & 0 \\ 0 & 9 \end{pmatrix} - \frac{2}{12} \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix} - \frac{6}{12} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \\
&= \frac{1}{12} \begin{pmatrix} -4 & 0 \\ 0 & -11 \end{pmatrix} < 0.
\end{align*}$$

**Proposition 2.** Let $f$ be a real-valued continuous function defined on an interval $[0, \infty)$. If $f$ is operator convex and non-positive, then $f$ is an operator superquadratic function.

**Proof.** Since $f$ is operator convex,

$$\frac{f(A) + f(B)}{2} - f\left( \frac{A + B}{2} \right) \geq 0.$$

However, $f$ is also non-positive, so that

$$\frac{f(A) + f(B)}{2} - f\left( \frac{A + B}{2} \right) - f\left( \frac{|A-B|}{2} \right) \geq -f\left( \frac{|A-B|}{2} \right) \geq 0,$$

which means that $f$ is an operator superquadratic function. \[Q.E.D.\]

**Example 2.** Let $f(t) = t \log(t)$, $t \in [0, \infty)$. Then it is well known that $f$ is operator convex. Clearly, $f$ is negative for all $t \in (0, 1) \subseteq [0, \infty)$. Therefore, $f(t) = t \log(t)$ is an operator superquadratic function for all $t \in (0, 1)$.

**Proposition 3.** Let $f$ be a real-valued continuous function defined on an interval $[0, \infty)$. If $f$ is operator concave and non-negative, then $f$ is operator subquadratic.
1. If \( f \) is a positive unital linear map, then a general operator Bohr inequality can be given in the form

\[
X = \| \Phi \|_{\mathcal{B}}(X^*)X
\]

which means that \( f \) is operator subquadratic.

2. If \( f \) is an operator superquadratic function, then we have

\[
f\left( \frac{A + B}{2} \right) - f\left( \frac{A - B}{2} \right) + f\left( \frac{|A - B|}{2} \right) \geq 0,
\]

which means that \( f \) is operator subquadratic.

\[\Box\]

Example 3. Let \( f : (0, \infty) \to (0, \infty) \), given by \( f(t) = t^r \), \( r \in [0,1] \). Then \( f \) is operator subquadratic on \( (0, \infty) \). However, \( f \) is also operator concave, so that

\[
\frac{A^r + B^r}{2} \leq \left( \frac{A + B}{2} \right)^r \leq \left( \frac{|A - B|}{2} \right)^r.
\]

On Bohr’s Inequality

The classical Bohr inequality for scalars reads that: if \( a \) and \( b \) are complex numbers and \( p, q > 1 \) with \( \frac{1}{p} + \frac{1}{q} = 1 \), then

\[
|a - b|^2 \leq p|a|^2 + q|b|^2.
\]

An operator version of this inequality was treated by Hirzallah [25] and the latter by many authors. See, for example, [13,26]. Namely, in [25], we find that

\[
\]

is valid for all \( A, B \in \mathcal{B}(\mathcal{H}) \) and \( p, q > 1 \) with \( \frac{1}{p} + \frac{1}{q} = 1 \) and \( p \leq q \), where \( |X| = (X^*X)^{1/2} \) is the absolute value of the operator \( X \).

Recently, it is shown in [24] that, for a positive self-adjoint operator \( A \in \mathcal{B}(\mathcal{H}) \) and a positive unital linear map \( \Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}) \), the following inequalities hold:

1. If \( f : [0, \infty) \to \mathbb{R} \) is a real superquadratic function, then we have

\[
\| \Phi(f(|A - \|\Phi(A)\|_{1,\mathcal{H}}|)) \| \leq \| \Phi(f(A)) \| - f(\|\Phi(A)\|) - f(0).
\]

In particular, for \( f(t) = t^r \), \( r \geq 2 \), \( t \geq 0 \), we have

\[
\| \Phi(|A - \|\Phi(A)\|_{1,\mathcal{H}}^r) \| \leq \| \Phi(A^r) \| - \|\Phi(A)\|^r.
\]

2. If \( f : [0, \infty) \to \mathbb{R} \) is a real subquadratic function, then we have

\[
\| \Phi(f(|A - \|\Phi(A)\|_{1,\mathcal{H}}|)) \| \geq \| \Phi(f(A)) \| - f(\|\Phi(A)\|) - f(0).
\]

In particular, for \( f(t) = t^r \), \( 0 < r \leq 2 \), \( t \geq 0 \), we have

\[
\| \Phi(|A - \|\Phi(A)\|_{1,\mathcal{H}}^r) \| \geq \| \Phi(A^r) \| - \|\Phi(A)\|^r.
\]

Now, set \( \alpha = \frac{1}{2} \), \( p > 1 \) so that \( 1 - \alpha = \frac{1}{2} \) with \( \frac{1}{p} + \frac{1}{q} = 1 \) in (10). If \( f \) is an operator superquadratic function, then a general operator Bohr inequality can be given in the form

\[
pq f\left( \frac{A + B}{p} \right) + pf\left( \frac{|A - B|}{p} \right) + qf\left( \frac{|A - B|}{q} \right) \leq qf(A) + pf(B)
\]

(13)
for all \( p, q > 1 \) with \( \frac{1}{p} + \frac{1}{q} = 1 \). In particular, for \( p = q = 2 \), we have

\[
f\left( \frac{A + B}{2} \right) + f\left( \frac{|A - B|}{2} \right) \leq f(A) + f(B).
\]

(14)

If \( f \) is subquadratic then the inequalities (13) and (14) are reversed.

As a direct example, let \( f(t) = t^r, t \geq 0, r \in [0,1] \). Then \( f \) is operator subquadratic.

Hence, by (13), we have

\[
pq\left( \frac{A}{p} + \frac{B}{q} \right)^r + \left( p^{1-r} + q^{1-r} \right)|A - B|^r \geq qA^r + pB^r,
\]

which is equivalent to writing

\[
(p + q)^{1-r}(qA + pB)^r + \left( p^{1-r} + q^{1-r} \right)|A - B|^r \geq qA^r + pB^r,
\]

since \( pq = p + q \) for all \( p, q > 1 \) with \( \frac{1}{p} + \frac{1}{q} = 1 \), for all positive operators \( A, B \in \mathcal{B}(\mathcal{H})^+ \).

In particular, for \( p = q = 2 \), we have

\[
(A + B)^r + |A - B|^r \geq 2^{r-1}(A^r + B^r).
\]

3. Operator Jensen’s Inequality

In order to prove our main results, we need the following lemmas:

**Lemma 4** ([11]). If \( A \in \mathcal{B}(\mathcal{H}) \) is self-adjoint and \( U \) is unitary, that is, \( U^*U = UU^* = 1_{\mathcal{H}} \), then \( f(U^*AU) = U^*f(A)U \) for every continuous function \( f \) on \( \text{sp}(A) \).

**Lemma 5** ([27]). Define a unitary matrix \( E_n = \text{diag}(\xi, \xi^2, \ldots, \xi^{n-1}, 1) \) in \( M_{n \times n}(\mathbb{C}) \subset \mathcal{B}(\mathcal{H}^n) \), where \( \xi = \exp\left( \frac{2\pi i}{n} \right) \). Then, for each element \( A = (a_{ij}) \in \mathcal{B}(\mathcal{H}^n) \), we have

\[
\frac{1}{n} \sum_{k=1}^{n} E^{-k}AE^{k} = \text{diag}(a_{11}, a_{22}, \ldots, a_{nn}).
\]

**Lemma 6** ([27]). Let \( P \) denote the projection in \( M_{n \times n}(\mathbb{C}) \) given by \( P_{ij} = n^{-1} \) for all \( i \) and \( j \), so that \( P \) is the projection of rank one on the subspace spanned by the vector \( \xi + \xi^2 + \cdots + \xi^n \) in \( \mathbb{C}^n \), where \( \xi, \xi^2, \ldots, \xi^n \) are the standard basis vectors. Then with \( E \) as in Lemma 5 we obtain the pairwise orthogonal projections \( P_k = E^{-k}PE^k \), for \( 1 \leq k \leq n \), with \( \sum_{k=1}^{n} P_k = 1_{\mathcal{H}} \).

In order to establish our main first result, we need the following primary result:

**Lemma 7.** Let \( w_1, \ldots, w_n \) be positive real numbers such that \( W_n = \sum_{k=1}^{n} w_k \) and let \( A_1, \ldots, A_n \) be positive operators on a Hilbert space \( \mathcal{B}(\mathcal{H}) \) with spectra contained in a real interval \( I \). If \( f \) is an operator superquadratic function on \( I \), then

\[
f\left( \frac{1}{W_n} \sum_{k=1}^{n} w_k A_k \right) \leq \sum_{k=1}^{n} \frac{w_k}{W_n} f(A_k) - \sum_{k=1}^{n} \frac{w_k}{W_n} \left( \left| A_k - \sum_{j=1}^{n} \frac{w_j}{W_n} A_j \right| \right).
\]

(15)

In particular, as a useful case, for \( w_k = 1 \) for all \( 1 \leq k \leq n \), we have

\[
f\left( \frac{1}{n} \sum_{k=1}^{n} A_k \right) \leq \frac{1}{n} \sum_{k=1}^{n} f(A_k) - \frac{1}{n} \sum_{k=1}^{n} \left( \left| A_k - \frac{1}{n} \sum_{j=1}^{n} A_j \right| \right).
\]

(16)
Proof. Assume \( f \) is operator superquadratic. If \( n = 2 \), then the inequality (15) reduces to (10) with \( \alpha = \frac{w_1}{W_2} \) and \( 1 - \alpha = \frac{w_2}{W_2} \). Let us suppose that inequality (15) holds for \( n - 1 \). Then, for \( n \)-tuples \((A_1, \ldots, A_n)\) and \((w_1, \ldots, w_n)\), we have

\[
f \left( \frac{1}{W_n} \sum_{k=1}^{n} w_k A_k \right) = f \left( \frac{w_n}{W_n} A_n + \sum_{k=1}^{n-1} \frac{w_k}{W_n} A_k \right)
\]

\[
= f \left( \frac{w_n}{W_n} A_n + \frac{W_{n-1}}{W_n} \sum_{k=1}^{n-1} \frac{w_k}{W_{n-1}} A_k \right)
\]

\[
\leq \frac{w_n}{W_n} f(A_n) + \frac{W_{n-1}}{W_n} f \left( \sum_{k=1}^{n-1} \frac{w_k}{W_{n-1}} A_k \right)
\]

\[
= \frac{w_n}{W_n} f(A_n) + \frac{W_{n-1}}{W_n} f \left( \sum_{k=1}^{n-1} \frac{w_k}{W_{n-1}} A_k \right) - \frac{w_n}{W_n} f \left( \frac{W_{n-1}}{W_n} A_n - \frac{1}{W_{n-1}} \sum_{k=1}^{n-1} \frac{W_{n-1} w_k A_k}{k} \right)
\]

\[
- \frac{W_{n-1}}{W_n} f \left( \frac{w_n}{W_n} A_n - \frac{1}{W_{n-1}} \sum_{k=1}^{n-1} \frac{w_k A_k}{k} \right),
\]

and this is exactly equivalent to writing, for any \( 1 \leq m \leq n \),

\[
f \left( \frac{1}{W_m} \sum_{k=1}^{m} w_k A_k \right) \leq \sum_{k=1}^{m} \frac{w_k}{W_m} f(A_k) - \sum_{k=1}^{m} \frac{w_k}{W_m} f \left( A_k - \sum_{j=1}^{m} \frac{w_j}{W_m} A_j \right),
\]

which proves the desired result in (15). The particular case follows by setting \( w_k = 1 \) for all \( k = 1, \ldots, n \) so that \( W_n = n \). \( \square \)

Remark 2. The result in Lemma 7 was proved by Mond and Pečarić in [28] for all operator convex functions and all bounded self-adjoint operators whose spectra are contained in \( J \). Therefore, in case \( f \) is positive, the inequality (15) might be considered as a respective extension and a new refinement of that result proved in [28].

Theorem 3. Let \( f : J \to \mathbb{R} \) be a real-valued continuous function. Let \((A_1, \ldots, A_n)\) be an \( n \)-tuple of positive operators on a Hilbert space \( \mathcal{H} \) with spectra contained in \( J \). Then the following conditions are equivalent:

1. \( f \) is an operator superquadratic function.

2. The inequality

\[
f \left( \sum_{k=1}^{n} C_k A_k C_k \right) \leq \sum_{k=1}^{n} C_k f(A_k) C_k - \sum_{k=1}^{n} C_k f \left( A_k - \sum_{j=1}^{n} C_j A_j C_j \right) C_k
\]

holds for every \( n \)-tuple \((C_1, \ldots, C_n)\) of operators on \( \mathcal{H} \) that satisfy the condition \( \sum_{k=1}^{n} C_k = 1_{\mathcal{H}} \).

3. The inequality

\[
f \left( \sum_{k=1}^{n} P_k A_k P_k \right) \leq \sum_{k=1}^{n} P_k f(A_k) P_k - \sum_{k=1}^{n} P_k f \left( A_k - \sum_{j=1}^{n} P_j A_j P_j \right) P_k
\]

(18)
holds for every $n$-tuple $(P_1, \cdots, P_n)$ of projections on $\mathcal{H}$ with $\sum_{k=1}^n P_k = 1_{\mathcal{H}}$.

**Proof.** (1) $\Rightarrow$ (2). We say that $C = (C_1, \cdots, C_n)$ is a unitary column if there is a unitary $n \times n$ operator matrix $U = (u_{ij})$, one of whose columns is $(C_1, \cdots, C_n)$. Thus, $u_{ij} = C_i$ for some $j$ and all $i$. Assume that we are given a unitary $n$-column $(C_1, \cdots, C_n)$, and choose a unitary $U_n = (u_{ij})$ in $\mathcal{B}(\mathcal{H}^n)$ such that $u_{kn} = C_k$. Let $E = \text{diag}(\xi, \xi^2, \cdots, \xi^{n-1}, 1)$ as in Lemma 4 and put $X = \text{diag}(A_1, \cdots, A_n)$, both regarded as elements, in $\mathcal{B}(\mathcal{H}^n)$. Thus, using the spectral decomposition theorem, we have

$$f\left(\sum_{k=1}^n C_k^* A_k C_k\right) = f((U_n^* X U_n)_{nn}) = f\left(\left(\frac{1}{n} \sum_{k=1}^n E^{-k} U_k^* X U_k E^k\right)_{nn}\right).$$

We note that, since $f(\text{diag}(y_1, \cdots, y_n)) = \text{diag}(f(y_1), \cdots, f(y_n))$,

$$f(y_n) = f(\text{diag}(y_1, \cdots, y_n))_{nn}.$$

Using the above facts taking into account Lemmas 4–7 together with the inequality (16), the operator superquadracity of $f$ implies that

$$f\left(\sum_{k=1}^n C_k^* A_k C_k\right) = f\left(\left(\frac{1}{n} \sum_{k=1}^n E^{-k} U_k^* X U_k E^k\right)_{nn}\right)$$

$$= \left(f\left(\frac{1}{n} \sum_{k=1}^n E^{-k} U_k^* X U_k E^k\right)\right)_{nn}$$

$$\leq \frac{1}{n} \sum_{k=1}^n f\left(E^{-k} U_k^* X U_k E^k\right)_{nn}$$

$$= \left(\frac{1}{n} \sum_{k=1}^n E^{-k} U_k^* f(X) U_k E^k\right)_{nn}$$

$$= \left(\frac{1}{n} \sum_{k=1}^n E^{-k} U_k^* \left(X - \frac{1}{n} \sum_{j=1}^n E^{-j} U_j^* X U_j E^j\right) U_k E^k\right)_{nn}$$

$$= \left(\frac{1}{n} \sum_{k=1}^n E^{-k} U_k^* f(X) U_k E^k\right)_{nn}$$

$$= \left(\frac{1}{n} \sum_{k=1}^n E^{-k} U_k^* \left(X - \frac{1}{n} \sum_{j=1}^n E^{-j} U_j^* X U_j E^j\right) U_k E^k\right)_{nn}$$

$$= \left(\frac{1}{n} \sum_{k=1}^n E^{-k} U_k^* f(X) U_k E^k\right)_{nn}$$

$$= \left(\frac{1}{n} \sum_{k=1}^n E^{-k} U_k^* \left(X - \frac{1}{n} \sum_{j=1}^n E^{-j} U_j^* X U_j E^j\right) U_k E^k\right)_{nn}$$

$$= \left(\frac{1}{n} \sum_{k=1}^n E^{-k} U_k^* f(X) U_k E^k\right)_{nn}$$

$$= \left(\frac{1}{n} \sum_{k=1}^n E^{-k} U_k^* \left(X - \frac{1}{n} \sum_{j=1}^n E^{-j} U_j^* X U_j E^j\right) U_k E^k\right)_{nn}$$

It remains to mention that, when the column is just unital, we extend it to the unitary $(n+1)$-column $(C_1, \cdots, C_n, 0)$ and choose $A_{n+1}$ arbitrarily, but with spectrum in $f$ (see [29]). By the first part of the proof, we therefore have
\[
\begin{align*}
&f\left(\sum_{k=1}^{n} C_k^* A_k C_k\right) = f\left(\sum_{k=1}^{n+1} C_k^* A_k C_k\right) \\
&\leq \sum_{k=1}^{n+1} C_k^* f(A_k) C_k - \sum_{k=1}^{n+1} C_k^* \left| A_k - \sum_{j=1}^{n+1} C_j^* A_j C_j \right| C_k \\
&= \sum_{k=1}^{n} C_k^* f(A_k) C_k - \sum_{k=1}^{n} C_k^* \left| A_k - \sum_{j=1}^{n} C_j^* A_j C_j \right| C_k
\end{align*}
\]

and thus the proof of statement (2) is completely established.

(2) \(\Rightarrow\) (3). It obviously holds.

(3) \(\Rightarrow\) (1). Let \(A\) and \(B\) be positive and bounded linear operators with spectra in \(J\) and \(0 \leq \lambda \leq 1\).

Consider

\[
X = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \quad P = \begin{pmatrix} 1_{\mathcal{H}} & 0 \\ 0 & 0 \end{pmatrix}, \quad Q = 1_{\mathcal{H} \oplus \mathcal{H}} - P,
\]

\[
C = \begin{pmatrix} \sqrt{1 - \lambda} & -\sqrt{1 - \lambda} \\ \sqrt{1 - \lambda} & \sqrt{1 - \lambda} \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} \sqrt{1 - \lambda} & -\sqrt{1 - \lambda} \\ \sqrt{1 - \lambda} & \sqrt{1 - \lambda} \end{pmatrix}.
\]

Then, \(C\) and \(D\) are unitary operators on \(\mathcal{H} \oplus \mathcal{H}\). We have

\[
C^* X C = \begin{pmatrix} \lambda A + (1 - \lambda) B & 0 \\ 0 & (1 - \lambda) A + \lambda B \end{pmatrix},
\]

\[
D^* X D = \begin{pmatrix} (1 - \lambda) A + \lambda B & 0 \\ 0 & \lambda A + (1 - \lambda) B \end{pmatrix},
\]

\[
PC^* X CP + QD^* X DQ = \begin{pmatrix} 0 & 0 \\ 0 & (1 - \lambda) A + \lambda B \end{pmatrix}.
\]

Thus, we have

\[
\begin{align*}
&f\left(\lambda A + (1 - \lambda) B \\ 0 & (1 - \lambda) A + \lambda B \right) \\
&= \begin{pmatrix} f(\lambda A + (1 - \lambda) B) & 0 \\ 0 & f((1 - \lambda) A + \lambda B) \end{pmatrix} \\
&= f(\lambda f(A) + (1 - \lambda) f(B) \\ 0 & (1 - \lambda) f(A) + \lambda f(B) \right) \\
&\leq Pf(\lambda f(A) + (1 - \lambda) f(B) \\ 0 & (1 - \lambda) f(A) + \lambda f(B) \right) \\
&+ Qf((1 - \lambda)|A - B| + (1 - \lambda)f(|A - B|) \end{align*}
\]

Hence, \(f\) is operator superquadratic on \(J\) by seeing the \((1,1)\)-components. \(\square\)
Remark 3. An operator convex version of Theorem 3 was proved by Hansen and Pedersen in [27]. Therefore, in the case of \( f \) being positive, the inequalities (17) and (18) could be considered as new refinements of the result proven in [27]; for example the function \( f(t) = t^2, \ t > 0 \), is a nontrivial example that refines the Hansen-Pedersen inequalities in [27].

A refinement of the classical Jensen’s inequality (9) could be elaborated as follows:

**Corollary 2.** Let \( f : J \rightarrow \mathbb{R} \) be a real-valued continuous function. Let \( A \) be a positive operator on a Hilbert space \( \mathcal{H} \) with spectra contained in \( J \). If \( f \) is an operator superquadratic function, then the inequality

\[
f(C^*AC) \leq C^*f(A)C - C^*f(|A - C^*AC|)C
\]

holds for every operator \( C \) on \( \mathcal{H} \) that satisfies the condition \( C^*C = 1_{\mathcal{H}} \).

**Proof.** It follows from Theorem 3 by setting \( n = 1 \).

**Remark 4.** Let \( f : J \rightarrow \mathbb{R} \) be a real-valued continuous function. Let \( A \) be a positive operator on a Hilbert space \( \mathcal{H} \) with spectra contained in \( J \). If \( f \) is an operator subquadratic function, then the inequality

\[
f(C^*AC) \geq C^*f(A)C - C^*f(|A - C^*AC|)C
\]

holds for every operator \( C \) on \( \mathcal{H} \) that satisfies the condition \( C^*C = 1_{\mathcal{H}} \). Furthermore, by applying the subquadratic function \( f(t) = t^r, \ t > 0 \ (r \in [0, 1]) \), then we have

\[
(C^*AC)^r \geq C^*A^rC - C^*|A - C^*AC|^rC
\]

for all \( r \in [0, 1] \).

A generalization of the famous inequality of Davis–Choi (8) and thus (19) to any positive unital linear map.

**Theorem 4.** Let \( \mathcal{H} \) be a Hilbert space. Let \( f : J \rightarrow \mathbb{R} \) be a real-valued continuous function. Let \( A \) be a positive operator on a Hilbert space \( \mathcal{H} \) with spectra contained in \( J \) and \( \Phi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H}) \) be a positive unital linear map. If \( f \) is an operator superquadratic function, then the inequality

\[
f(\Phi(A)) \leq \Phi(f(A)) - \Phi(f(|A - \Phi(A)|))
\]

holds. If \( f \) is operator subquadratic, then the inequality (20) is reversed. Thus, the following refinement of the celebrated Kadison inequality (7) is valid:

\[
\Phi^2(A) \leq \Phi(A^2) - \Phi(|A - \Phi(A)|^2).
\]

**Proof.** Let \( A \in \mathcal{B}(\mathcal{H}) \) be positive. Assume that \( \mathcal{A} \) is the \( C^* \)-subalgebra of \( \mathcal{B}(\mathcal{H}) \) generated by \( A \) and \( 1_{\mathcal{H}} \). Without loss of generality, we may assume that \( \Phi \) is defined on \( \mathcal{A} \). Since every unital positive linear map on a commutative \( C^* \)-algebra is completely positive. It follows that \( \Phi \) is completely positive. So there exists (by Stinespring’s theorem [30]), some isometry \( V : \mathcal{H} \rightarrow \mathcal{H} \); and a unital \(*\)-homomorphism \( \rho \) from \( \mathcal{A} \) into the \( C^* \)-algebra \( \mathcal{B}(\mathcal{H}) \) such that \( \Phi(A) = V^*\rho(A)V \). Clearly, \( f(\rho(A)) = \rho(f(A)) \), for all continuous functions \( f \). Thus,
\[
f(\Phi(A)) = f(V^*\rho(A)V) \\
\leq V^*f(\rho(\Phi(A))V) - V^*f(|\rho(A - \Phi(A)||V) \quad \text{(by (17) with } n = 1) \\
= V^*\rho(f(A))V - V^*\rho(f(|A - \Phi(A)||V) \\
= \Phi(f(A)) - \Phi(f(|A - \Phi(A)||V)
\]

which proves the required inequality. The last inequality holds by applying (20) to the superquadratic function \( f(t) = t^2, t > 0 \).

**Corollary 3.** Let \( \mathcal{H} \) be a Hilbert space. Let \( f: [0, \infty) \to \mathbb{R} \) be a real-valued continuous function and \( \Phi_k: B(\mathcal{H}) \to B(\mathcal{H}) \ (k = 1, \ldots, n) \) be a positive linear mappings with \( \sum_{k=1}^{n} \Phi_k(1_{\mathcal{H}}) = 1_{\mathcal{H}} \). Then, \( f \) is an operator superquadratic function if and only if

\[
f\left(\sum_{k=1}^{n} \Phi_k(A_k)\right) \leq \sum_{k=1}^{n} \Phi_k(f(A_k)) - \sum_{k=1}^{n} \Phi_k\left(f\left(|A_k - \sum_{j=1}^{n} \Phi_j(A_j)|\right)\right)
\]

for all positive operators \( A_1 \cdots A_n \) in \( B(\mathcal{H}) \).

**Proof.** The proof is obvious, and thus omitted.

**4. Jensen’s Trace Inequality**

Let \( A \in M_{n \times n}(\mathbb{C}) \), and recall that the trace of a square matrix equals the sum of the eigenvalues counted with multiplicities. Moreover, the trace of a Hermitian matrix is real. If \( A \) is a linear operator represented by a square matrix with real or complex entries and if \( \lambda_1, \ldots, \lambda_n \) are the eigenvalues of \( A \), then \( \text{Tr}(A) = \sum_j \lambda_j \). This follows from the fact that \( A \) is always similar to its Jordan form, an upper triangular matrix having \( \lambda_1, \ldots, \lambda_n \) on the main diagonal.

The inner product

\[
\langle A, B \rangle = \text{Tr}(A^*B),
\]

which is defined on the space of all complex (or real) \( m \times n \) matrices, is called the Frobenius norm, which satisfies the submultiplicative property as a matrix norm.

If \( A \) and \( B \) are real positive semi-definite matrices of the same size, using the Cauchy–Schwarz inequality, we have

\[
0 \leq \text{Tr}^2(AB) \leq \text{Tr}\left(A^2\right)\text{Tr}\left(B^2\right) \leq \text{Tr}^2(A)\text{Tr}^2(B).
\]

The concept of trace of a matrix is generalized to the trace class of compact operators on Hilbert spaces, and the analog of the Frobenius norm is called the Hilbert–Schmidt norm, which is can be defined as

\[
\langle A, B \rangle = \text{Tr}(A^*B) = \sum_i \langle Ae_i, Be_i \rangle
\]

over all orthonormal basis of \( \mathcal{H}, \{e_i : i \in I\} \) (see [31]).

A Hilbert–Schmidt operator is a bounded operator \( A \) on a Hilbert space \( \mathcal{H} \) with finite Hilbert–Schmidt norm

\[
\|A\|_{HS}^2 = \text{Tr}(A^*A) = \sum_{i \in I} \|Ae_i\|^2,
\]
Let $f$ be a real-valued continuous function defined on an interval $J$ and let $m$ and $n$ be natural numbers. If $f$ is a superquadratic function (in the ordinary sense), then the inequality

$$
\|A\|_1 = \operatorname{Tr}|A| := |\sum_i \langle |A|e_i, e_i \rangle| = \sum_i \langle (A^*A)^{1/2}e_i, e_i \rangle
$$

is finite. In this case, the trace of $A$, which is given by the sum

$$
\operatorname{Tr}(A) = \sum_i \langle Ae_i, e_i \rangle,
$$

is absolutely convergent and is independent of the choice of the orthonormal basis. When $\mathcal{H}$ is finite-dimensional, every operator is a trace class and this definition of trace of $A$ coincides with the definition of the trace of a matrix.

Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and $n$ be any integer. It is well known that if $t \mapsto f(t)$ is convex (monotone increasing), then the trace function $A \mapsto \operatorname{Tr}(f(A))$ is convex (monotone increasing); see [32,33], and the recent work in [34].

In 2003, Hansen and Pedersen [27] proved the following version of Jensen’s inequality:

$$
\operatorname{Tr}\left( f\left( \sum_{k=1}^n C_k^* A_k C_k \right) \right) \leq \operatorname{Tr}\left( \sum_{k=1}^n C_k^* f(A_k) C_k \right)
$$

for every $n$-tuple $(A_1, \ldots, A_n)$ of positive $m \times m$ matrices with spectra contained in $J$ and every $n$-tuple $(C_1, \ldots, C_n)$ of $m \times m$ matrices with $\sum_{k=1}^n C_k^* C_k = I_m$, where $f$ is assumed to be convex on $J$.

Using the concept of superquadratic functions, we could give the following refinement of Hansen–Pedersen trace inequality:

**Theorem 5.** Let $f$ be a real-valued continuous function defined on an interval $J$ and let $m$ and $n$ be natural numbers. If $f$ is a superquadratic function (in the ordinary sense), then the inequality

$$
\operatorname{Tr}\left( f\left( \sum_{k=1}^n C_k^* A_k C_k \right) \right) \leq \operatorname{Tr}\left( \sum_{k=1}^n C_k^* f(A_k) C_k \right) - \operatorname{Tr}\left( \sum_{k=1}^n C_k^* \left( A_k - \operatorname{Tr}\left( \sum_{j=1}^n C_j^* A_j C_j \right) I_m \right) C_k \right)
$$

(21)

holds for every $n$-tuple $(A_1, \ldots, A_n)$ of positive $m \times m$ matrices with spectra contained in $J$ and every $n$-tuple $(C_1, \ldots, C_n)$ of $m \times m$ matrices with $\sum_{k=1}^n C_k^* C_k = I_m$. Conversely, if the inequality (21) is satisfied for some $n$ and $m$, where $n > 1$, then $f$ is a superquadratic function. If $f$ is subquadratic, then the inequality (21) is reversed.

**Proof.** Our proof is motivated by [27]. Let $A_k = \sum_{\lambda \in \sigma(A_k)} \lambda E_k(\lambda)$ denote the spectral resolution of $A_k$ for $1 \leq k \leq n$. Then, $E_k(\lambda)$ is the spectral projection of $A_k$ on the eigenspace corresponding to $\lambda$ if $\lambda$ is an eigenvalue for $A_k$, otherwise $E_k(\lambda) = 0$. For each unit vector $\xi$ in $C^n$, let us define the probability measure

$$
\mu_\xi(S) = \left( \sum_{k=1}^n C_k^* E_k(S) C_k \xi, \xi \right) = \sum_{k=1}^n \langle E_k(S) C_k \xi, C_k \xi \rangle
$$
for any (Borel) set S in \(\mathbb{R}\). Note that if \(y = \sum_{k=1}^{n} C_k A_k C_k\), then

\[
(y \xi, \xi) = \left\langle \sum_{k=1}^{n} C_k^* A_k C_k \xi, \xi \right\rangle = \sum_{k=1}^{n} \sum_{1 \in \text{sp}(A_k)} \lambda E_k(\lambda) C_k \xi, C_k \xi \right\rangle
= \int \lambda d \mu_\xi(\lambda).
\]

If a unit vector \(\xi\) is an eigenvector for \(y\), then the corresponding eigenvalue is \((y \xi, \xi)\), and \(\xi\) is also an eigenvector for \(f(y)\) with the corresponding eigenvalue \((f(y) \xi, \xi) = f((y \xi, \xi))\). In this case, we have

\[
\left\langle f \left( \sum_{k=1}^{n} C_k^* A_k C_k \right) \xi, \xi \right\rangle = \left\langle (f(y)) \xi, \xi \right\rangle
= f((y \xi, \xi))
\leq \int \left[ f(\lambda) - f \left( \left| \lambda - \int \lambda d \mu_\xi(\lambda) \right| \right) \right] d \mu_\xi(\lambda) \quad \text{(by 4)}
= \sum_{k=1}^{n} \left\langle \sum_{1 \in \text{sp}(A_k)} f(\lambda) - f \left( \left| \lambda - \int \lambda d \mu_\xi(\lambda) \right| \right) E_k(\lambda) C_k \xi, C_k \xi \right\rangle
= \sum_{k=1}^{n} \left[ C_k^* f(A_k) C_k - C_k^* f(A_k - \langle y \xi, \xi \rangle) C_k \xi, C_k \xi \right]\]

Summing over an orthonormal basis of eigenvectors for \(y\), we get the desired result in (21). \(\Box\)

**Corollary 4.** Let \(f\) be a real-valued continuous function defined on an interval \(J\) and let \(m\) and \(n\) be natural numbers. If \(f\) is a superquadratic function (in the ordinary sense), then the inequality

\[
\text{Tr}(f(C^* AC)) \leq \text{Tr}(C^* f(A) C) - \text{Tr}(C^* f(A - \text{Tr}(C^* AC) I_m) C) \quad \text{(22)}
\]

holds for every positive \(m \times m\) matrix \(A\) with spectrum contained in \(J\) and every \(m \times m\) matrix \(C\) with \(C^* C = I_m\). If \(f\) is subquadratic, then the inequality (22) is reversed. Furthermore, we have

\[
\text{Tr}(\langle C^* AC \rangle^p) \leq \text{Tr}(C^* A^p C) - \text{Tr}(C^* A - \text{Tr}(C^* AC) I_m|C^p) \quad \text{(23)}
\]

for every \(p \geq 2\), and

\[
\text{Tr}(\langle C^* AC \rangle^p) \geq \text{Tr}(C^* A^p C) - \text{Tr}(C^* A - \text{Tr}(C^* AC) I_m|C^p) \quad \text{(24)}
\]

for every \(p \in (0, 1]\).

**Proof.** The result follows by setting \(n = 1\) in Theorem 5. The inequality (23) follows by applying the superquadratic function \(f(t) = t^p, p \geq 2\). Similarly, the inequality (24) follows by applying the subquadratic function \(f(t) = t^p, p \in (0, 1]\). \(\Box\)

The inequality (21) could be extended for general positive Hilbert space operators mapped under a positive unital linear map, as follows:
Theorem 6. Let $f$ be a real-valued continuous function defined on $[0, \infty)$. Let $A_j \in M_{m \times m}(\mathbb{C})$ $(j = 1, \cdots, n)$ be positive operators. Let $\Phi_j : M_{m \times m}(\mathbb{C}) \to M_{m \times m}(\mathbb{C})$ $(j = 1, \cdots, n)$ be a positive linear map, such that $\sum_{j=1}^{n} \Phi_j(I_m) = I_m$, where $I_m$ is the identity matrix of $m \times m$. If $f$ is a superquadratic function, then

$$\text{Tr} \left( \sum_{j=1}^{n} \Phi_j(f(A_j)) \right) \geq \text{Tr} \left( f \left( \sum_{j=1}^{n} \Phi_j(A_j) \right) \right) + \text{Tr} \left( \sum_{j=1}^{n} \Phi_j \left( f \left( \sum_{j=1}^{n} \Phi_j(A_j) \right) I_m \right) \right)$$

(25)

holds for every $n$-tuple $(A_1, \cdots, A_n)$ of positive $m \times m$ matrices with spectra contained in $I$. Conversely, if the inequality (25) is satisfied for some $n$ and $m$, where $n > 1$, then $f$ is a superquadratic function. If $f$ is subquadratic, then the inequality (25) is reversed.

Proof. Firstly, let $A_j \in M_{m \times m}(\mathbb{C})$ $(j = 1, \cdots, n)$ be positive. Assume that $B$ is the $C^*$-subalgebra of $M_{m \times m}(\mathbb{C})$ generated by $A_j$ and identity $I_m$. Let $\Phi_j : M_{m \times m}(\mathbb{C}) \to M_{m \times m}(\mathbb{C})$ $(j = 1, \cdots, n)$ be a positive linear map, such that $\sum_{j=1}^{n} \Phi_j(I_m) = I_m$. Let $(V_1, \cdots, V_n)$ be $n$-tuple of $m \times m$ matrices with $\sum_{j=1}^{n} V_j^*V_j = I_m$. Without loss of generality, assume that $\Phi_j$ is defined on $B$. Since every unital positive linear map on a commutative $C^*$-algebra is completely positive. It follows that $\Phi_j$ is completely positive. So, by the Stinespring’s theorem [30], there exists, an isometry $V_j : M_{m \times m}(\mathbb{C}) \to M_{m \times m}(\mathbb{C})$ (such isometry $V_j$ is valid for all $j$ since each such $B$ is $C^*$-subalgebra can be generated by different $A_j$ for all $j = 1, \cdots, n$), and a unital *-homomorphism $\rho$ from $B$ into the $C^*$-algebra $M_{m \times m}(\mathbb{C})$, such that

$$\Phi_j(A_j) = V_j^* \rho(A_j) V_j, \quad j = 1, \cdots, n.$$  

(26)

Clearly, $f(\rho(A_j)) = \rho(f(A_j))$, for all continuous functions $f$. Thus,

$$f(\Phi_j(A_j)) = f \left( V_j^* \rho(A_j) V_j \right)$$

$$\leq V_j^* f(\rho(A_j)) V_j - V_j^* f \left( \rho \left( A_j - V_j^* \rho(A_j) V_j \right) \right) V_j$$

(by (17) with $n = 1$)

$$= V_j^* \rho(f(A_j)) V_j - V_j^* \rho\left( f \left( |A_j - \Phi_j(A_j)| \right) \right) V_j$$

$$= \Phi_j(f(A_j)) - \Phi_j(f(|A_j - \Phi_j(A_j)|)).$$

Let $\Phi_j(A_j) = \sum_{\lambda \in \text{sp}(\Phi_j(A_j))} \lambda E_j(\lambda)$ denote the spectral resolution of $\Phi_j(A_j)$ for $j = 1, \cdots, n$.

Then, $E_j(\lambda)$ is the spectral projection of $\Phi_j(A_j)$ on the eigenspace corresponding to $\lambda$ if $\lambda$ is an eigenvalue for $\Phi_j(A_j)$, otherwise $E_j(\lambda) = 0$. For each unit vector $\zeta \in \mathbb{C}^m$, let us define the probability measure

$$\mu_{\zeta}(S) = \sum_{j=1}^{n} \langle V_j^* E_j(S) V_j \zeta, \zeta \rangle = \sum_{j=1}^{n} \langle E_j(S) V_j \zeta, V_j \zeta \rangle.$$
for any (Borel) set $S$ in $\mathbb{R}$. Note that, if $w = \sum_{j=1}^{n} V_j^* A_j V_j$, then
\[
\langle w \zeta, \zeta \rangle = \left\langle \sum_{j=1}^{n} V_j^* A_j V_j \zeta, \zeta \right\rangle = \left\langle \sum_{j=1}^{n} V_j^* \sum_{j \in \text{sp}(\Phi_j(A_j))} \lambda E_j(\lambda) V_j \zeta, V_j \zeta \right\rangle \\
= \int \lambda d\mu_{\zeta}(\lambda).
\]

If a unit vector $\zeta$ is an eigenvector for $w$, then the corresponding eigenvalue is $\langle w \zeta, \zeta \rangle$, and $\zeta$ is also an eigenvector for $f(w)$ with corresponding eigenvalue $\langle f(w) \zeta, \zeta \rangle = f(\langle w \zeta, \zeta \rangle)$. In this case, and taking into account the representation (26), we have
\[
\left\langle f \left( \sum_{j=1}^{n} \Phi_j(A_j) \right) \zeta, \zeta \right\rangle = \left\langle f \left( \sum_{j=1}^{n} V_j^* A_j V_j \right) \zeta, \zeta \right\rangle \\
= \left\langle f(w) \zeta, \zeta \right\rangle = f(\langle w \zeta, \zeta \rangle) \\
= f \left( \int \lambda d\mu_{\zeta}(\lambda) \right) \\
\leq \int \left[ f(\lambda) - f \left( \left| \lambda - \int \lambda d\mu_{\zeta}(\lambda) \right| \right) \right] d\mu_{\zeta}(\lambda) \quad \text{(by (4))} \\
= \sum_{j=1}^{n} \left\langle \sum_{j \in \text{sp}(\Phi_j(A_j))} f(\lambda) - f \left( \left| \lambda - \int \lambda d\mu_{\zeta}(\lambda) \right| \right) E_j(\lambda) V_j \zeta, V_j \zeta \right\rangle \\
= \sum_{j=1}^{n} \left\langle \left[ V_j^* f(A_j) V_j - V_j^* f(\langle A_j - \langle w \zeta, \zeta \rangle \rangle V_j \zeta, V_j \zeta \right\rangle \\
= \sum_{j=1}^{n} \left\langle \left[ \Phi_j(f(A_j)) \right] - \Phi_j \left( f \left( \left| A_j - \sum_{j=1}^{n} \Phi_j(A_j) \zeta, \zeta \right\rangle \right) \right) \zeta, \zeta \right\rangle.
\]

Summing over an orthonormal basis of eigenvectors for $w$, we get the desired result in (25). □

As a special case, we can deduce the following result:

**Corollary 5.** Let $f$ be a real-valued continuous function defined on $[0, \infty)$. Let $A \in \mathcal{M}_{m \times m}(\mathbb{C})$ be a positive operator and $\Phi : \mathcal{M}_{m \times m}(\mathbb{C}) \to \mathcal{M}_{m \times m}(\mathbb{C})$ be a positive unital linear map with $\Phi(I_m) = I_m$. If $f$ is superquadratic, then we have
\[
\text{Tr}(\Phi(f(A))) \geq \text{Tr}(f(\Phi(A))) + \text{Tr}(\Phi(f(|A - \text{Tr}(\Phi(A))I_m|))). \tag{27}
\]

A particular case is the choice $\Phi(A) = C^* AC$, where $C \in \mathcal{M}_{m \times m}(\mathbb{C})$, is such that $C^* C = I_m$. Indeed, this reduces to the inequality (22).

**Proof.** The proof follows by setting $n = 1$ in (25). □

and editing, M.W.A., C.C. and A.A.-K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The referees’ insightful criticism and careful reading of this work’s original manuscript helped the authors make the final version better, so they sincerely thank them.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


