On log-sum inequalities

Supriyo Dutta and Shigeru Furuichi

Article history
Received 1 June 2022
Accepted 30 December 2022

Communicated by R. Pereira

Keywords
Log-sum inequality; deformed logarithm; convex function; Löwner partial order; Hansen–Pedersen theory

Mathematics subject classifications
15A39; 15A42; 15A45

1. Introduction

In information theory, the so-called log-sum inequality is a generalization of Shannon inequality (often also called Klein inequality) which is used to prove the non-negativity of relative entropy. The essence of the non-negativity of the relative entropy is the simple inequality $\ln x \leq x - 1$ for $x > 0$. Therefore, log-sum inequality is important to study information theory. This is a variant of the Jensen inequality of convex functions, which plays a crucial role for proving the Gibbs’ inequality or the convexity of Kullback-Leibler divergence [1].

Recall that, a function $f : X \to \mathbb{R}$ defined on a convex set $X$ is said to be a convex function if for all $x_1, x_2 \in X$ and for all $t \in [0, 1]$ we have $tf(x_1) + (1 - t)f(x_2) \geq f(tx_1 + (1 - t)x_2)$. When $X$ is the set of real numbers, this inequality is mentioned as the Jensen inequality. In general, we represent the Jensen inequality [2,3] as

$$\sum_{i=1}^{n} t_i f(x_i) \geq f \left( \sum_{i=1}^{n} t_i x_i \right), \quad \text{where } \sum_{i=1}^{n} t_i = 1 \text{ and } 0 \leq t_i \leq 1.$$  \hspace{1cm} (1)

A twice-differentiable real valued function $f(x)$ on $\mathbb{R}$ is convex if and only if $f''(x) \geq 0$. Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ be non-negative numbers. As $f(x) = x \log(x)$ is a convex function.
function, the Jensen inequality (1) suggests
\[
\sum_{i=1}^{n} \frac{b_i}{\sum_{j=1}^{n} b_j} a_i \log \left( \frac{a_i}{b_i} \right) \geq \left( \sum_{i=1}^{n} \frac{b_i}{\sum_{j=1}^{n} b_j} a_i \right) \log \left( \sum_{i=1}^{n} \frac{b_i}{\sum_{j=1}^{n} b_j} a_i \right),
\]
where \( t_i = \frac{b_i}{\sum_{j=1}^{n} b_j} \) and \( x_i = \frac{a_i}{b_i} \). Now, replacing \( a = \sum_{j=1}^{n} a_j \) and \( b = \sum_{j=1}^{n} b_j \) we obtain,
\[
\sum_{i=1}^{n} a_i \log \left( \frac{a_i}{b_i} \right) \geq a \log \left( \frac{a}{b} \right),
\]
which is the standard log-sum inequality. The inequality is valid for \( n = \infty \) provided \( a < \infty \) as well as \( b < \infty \). In [4], an analog of log-sum inequality is derived as
\[
\sum_{i=1}^{n} b_i f \left( \frac{a_i}{b_i} \right) \geq b f \left( \frac{a}{b} \right),
\]
where \( f \) is strictly convex at \( \frac{a}{b} \). The equality holds if and only if \( a_i = c b_i \), for \( i = 1, 2, \ldots n \), and constant number \( c \).

Although log–sum inequality is important in classical information theory, a detailed study on this topic is not available in the literature, to the best of our knowledge. This article demonstrates a number of inequalities which are analogous to the log-sum inequality. First, we extend the log-sum inequality with two real-valued functions, which is illustrated in Theorem 2.1. This theorem expands the applicability of the log-sum inequality as it breaks the restriction that \( a_i \) and \( b_i \) are the real numbers. The convexity of a function is essential for proving log-sum inequality. Considering the concavity of a function, we also find analogous results. Investigating the variants of log-sum inequality for concave functions is crucial as there are classes of functions whose convexity depends on specific values of a parameter, for instance, the \( q \)-deformed logarithm. We also elaborate the matrix analogs of the log-sum inequality. Here we observe that all the inequalities derived for the real functions can be easily extended as trace-form-inequality for commuting Hermitian matrices. A recent trend in matrix analysis is constructing matrix theoretic counterparts of the known inequalities for real functions, where the matrix inequality is driven by the Löwner partial order relation. Proving the log-sum inequality in this context is non-trivial and difficult. We are able to derive it under several conditions. To the best of our knowledge, this is the first attempt in literature to extend the log-sum inequality for the non-commutative positive semidefinite matrices.

This article is distributed into five sections. Preliminary concepts are discussed before applying them in a mathematical derivation. In Section 2, we consider inequalities for real-valued functions. Here, we discuss the log-sum inequality for deformed logarithms, for instance the \( q \)-deformed logarithm. Section 3 is dedicated to the discussion of log-sum inequality as a trace-form-inequality for commuting Hermitian matrices. It has a number of consequences in quantum information theory. In the next section, we attempt to derive the log-sum inequality for Löwner partial order relation. This section extensively uses the idea of operator monotone, operator convex functions, and the operator Jensen inequality, which we mention at the beginning of Section 4. Then we conclude the article.
2. Log-sum inequalities for real functions

In this section, we extend the idea of the log-sum inequality for real-valued functions. Both concave and convex functions are considered for the investigation. Interestingly, the $q$-deformed logarithm is convex when $q < 2$, which assists us to illustrate two sets of log-sum inequalities. We begin our discussion with the following theorem.

**Theorem 2.1:** Let $g$ be a real valued function whose domain contains $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$, such that $g(b_i) > 0$ for all $i = 1, 2, \ldots, n$. Consider another function $f : [m_g, M_g] \to \mathbb{R}$ for which $h(x) = xf(x)$ is convex, where $m_g := \min_i \left\{ \frac{g(a_i)}{g(b_i)} \right\}$ and $M_g := \max_i \left\{ \frac{g(a_i)}{g(b_i)} \right\}$. Then,

$$\sum_{i=1}^{n} g(a_i)f\left(\frac{g(a_i)}{g(b_i)}\right) \geq \left(\sum_{i=1}^{n} g(a_i)\right)f\left(\frac{\sum_{i=1}^{n} g(a_i)}{\sum_{i=1}^{n} g(b_i)}\right).$$

**Proof:** Define $a = \sum_{i=1}^{n} g(a_i)$ and $b = \sum_{i=1}^{n} g(b_i) > 0$. Note that, $\frac{a}{b} = \sum_{i=1}^{n} \frac{g(b_i)}{g(a_i)}$, which is a convex combination of $\frac{g(a_i)}{g(b_i)}$ for $i = 1, 2, \ldots, n$. Clearly, the ratio $\frac{a}{b}$ belongs to the convex set $[m_g, M_g]$. Now,

$$\sum_{i=1}^{n} g(a_i)f\left(\frac{g(a_i)}{g(b_i)}\right) = \sum_{i=1}^{n} g(b_i)g(a_i)\frac{g(a_i)}{g(b_i)}f\left(\frac{g(a_i)}{g(b_i)}\right) = \sum_{i=1}^{n} b g(b_i) h\left(\frac{g(a_i)}{g(b_i)}\right).$$

As $g(b_i) > 0$ for all $i = 1, 2, \ldots, n$ and $b = \sum_{i=1}^{n} g(b_i)$ we have $0 \leq \frac{g(b_i)}{b} \leq 1$ and $\sum_{i=1}^{n} \frac{g(b_i)}{b} = 1$. Now, the Jensen inequality indicates

$$\sum_{i=1}^{n} b g(b_i) h\left(\frac{g(a_i)}{g(b_i)}\right) \geq bh\left(\sum_{i=1}^{n} g(b_i) g(a_i) \frac{g(a_i)}{b g(b_i)}\right) = bh\left(\frac{1}{b} \sum_{i=1}^{n} g(a_i)\right) = bh\left(\frac{a}{b}\right).$$

Expanding $h$ we get

$$bh\left(\frac{a}{b}\right) = b f\left(\frac{a}{b}\right) = af\left(\frac{a}{b}\right) = \left(\sum_{i=1}^{n} g(a_i)\right)f\left(\frac{\sum_{i=1}^{n} g(a_i)}{\sum_{i=1}^{n} g(b_i)}\right).$$

Combining, we obtain the result. 

From now on, we denote $m_g := \min_i \left\{ \frac{g(a_i)}{g(b_i)} \right\}$ and $M_g := \max_i \left\{ \frac{g(a_i)}{g(b_i)} \right\}$ throughout the article, where $g$ is a real valued function whose domain of definition contains $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$, such that $g(b_i) > 0$ for all $i = 1, 2, \ldots, n$.

Defining $g(x) = x$ and $f(x) = \log(x)$ for $x \in \mathbb{R}_{>0}$, we observe that $h(x) = x \log(x)$ is a convex function. Then, Theorem 2.1 leads us to the log-sum inequality mentioned in Equation (2).
Theorem 2.1 expands the scope of applications of the log-sum inequality. The domain of \( g \) may not be a convex set. We can consider \( a_i \) and \( b_i \) arbitrarily such that \( g(b_i) > 0 \) for all \( i = 1, 2, \ldots n \). In other words, the range of \( g \) must have a non-empty intersection with the set of positive reals.

Let \( f \) be a twice-differentiable function in \([m_g, M_g]\). The function \( xf(x) \) is convex if and only if

\[
\frac{d^2}{dx^2}(xf(x)) = xf''(x) + 2f'(x) \geq 0. \tag{5}
\]

If \( m_g \geq 0 \), then any monotone increasing and convex function \( f \) defined on \([m_g, M_g]\) fulfill Equation (5). Now, we are in a position to consider a few special cases of Theorem 2.1.

**Example 2.2:** Define \( g(x) = x^r \) for some real parameter \( r \), and \( f(x) = \log(x) \) for \( x > 0 \). As \( h(x) = x \log(x) \) is a convex function, applying Theorem 2.1 we observe

\[
\sum_{i=1}^{n} a_i^r \log \left( \frac{a_i^r}{b_i^r} \right) \geq \left( \sum_{i=1}^{n} a_i^r \right) \log \left( \frac{\sum_{i=1}^{n} a_i^r}{\sum_{i=1}^{n} b_i^r} \right).
\]

Thus we have

\[
\sum_{i=1}^{n} r a_i^r \left\{ \log(a_i) - \log(b_i) \right\} \geq \left( \sum_{i=1}^{n} a_i^r \right) \left\{ \log \left( \sum_{i=1}^{n} a_i^r \right) - \log \left( \sum_{i=1}^{n} b_i^r \right) \right\}. \tag{6}
\]

**Example 2.3:** The \( q \)-deformed logarithm or \( q \)-logarithm [5] is defined by

\[
f(x) = \ln_q(x) = \frac{x^{1-q} - 1}{1 - q}, \tag{7}
\]

for \( x > 0 \) and \( q \neq 1 \). Note that \( xf(x) \) is a convex function for \( q \leq 2 \) as \( \frac{d^2}{dx^2}(xf(x)) = \frac{2 - q}{x^q} \geq 0 \). The division rule of \( q \)-deformed logarithm can be expressed as

\[
\ln_q \left( \frac{x}{y} \right) = \ln_q(x) - \ln_q(y) = \ln_q(x) - \ln_q(y) = \frac{\ln_q(x) - \ln_q(y)}{y^{1-q}}. \tag{8}
\]

Putting \( g(x) = x^r \) for \( x > 0 \) and real parameter \( r \) as well as \( f(x) = \ln_q(x) \) with \( q \leq 2 \) in Theorem 2.1 we obtain

\[
\sum_{i=1}^{n} a_i^r \ln_q \left( \frac{a_i^r}{b_i^r} \right) \geq \left( \sum_{i=1}^{n} a_i^r \right) \ln_q \left( \frac{\sum_{i=1}^{n} a_i^r}{\sum_{i=1}^{n} b_i^r} \right),
\]

where \( a_1, a_2, \ldots a_n \) and \( b_1, b_2, \ldots b_n \) are positive real numbers. Now applying Equation (8), we find

\[
\ln_q \left( \frac{\sum_{i=1}^{n} a_i^r}{\sum_{i=1}^{n} b_i^r} \right) = \frac{\ln_q \left( \sum_{i=1}^{n} a_i^r \right) - \ln_q \left( \sum_{i=1}^{n} b_i^r \right)}{\left( \sum_{i=1}^{n} b_i^r \right)^{1-q}}.
\]
Combining we get
\[
\left( \sum_{i=1}^{n} b_i^r \right)^{1-q} \sum_{i=1}^{n} a_i^r \ln_q \left( \frac{a_i^r}{b_i^r} \right) \geq \left( \sum_{i=1}^{n} a_i^r \right) \left( \ln_q \left( \sum_{i=1}^{n} a_i^r \right) - \ln_q \left( \sum_{i=1}^{n} b_i^r \right) \right). \quad (9)
\]

For \( r = 1 \) the above inequality reduces to
\[
\left( \sum_{i=1}^{n} b_i \right)^{1-q} \sum_{i=1}^{n} a_i \ln_q \left( \frac{a_i}{b_i} \right) \geq \left( \sum_{i=1}^{n} a_i \right) \left( \ln_q \left( \sum_{i=1}^{n} a_i \right) - \ln_q \left( \sum_{i=1}^{n} b_i \right) \right). \quad (10)
\]

Instead of \( g(x) = x^r \), one may consider trigonometric, exponential, hyperbolic, or any other function to get new inequalities.

Recall that, a real-valued function \( f \) defined on a convex set \( X \) is said to be concave if \( -f \) is convex. Thus, if \( f \) is concave on \( X \) and \( x_i \in X \), we have
\[
\sum_{i=1}^{n} t_i f(x_i) \leq f \left( \sum_{i=1}^{n} t_i x_i \right), \quad \text{where} \quad \sum_{i=1}^{n} t_i = 1 \quad \text{and} \quad 0 \leq t_i \leq 1.
\]

Considering \( h(x) = xf(x) \) as a concave function in Equation (4), we obtain
\[
\sum_{i=1}^{n} g(a_i) f \left( \frac{g(a_i)}{g(b_i)} \right) \leq f \left( \sum_{i=1}^{n} g(a_i) \right) f \left( \frac{\sum_{i=1}^{n} g(a_i)}{\sum_{i=1}^{n} g(b_i)} \right),
\]
under the equivalent conditions on \( a_i \) and \( b_i \) as well as the real valued functions \( f \) and \( g \) as mentioned in Theorem 2.1.

When \( q > 2 \) in Equation (7) we observe that \( h(x) = xf(x) \) is a concave function. Therefore the inequalities in Equations (9) and (10) become
\[
\left( \sum_{i=1}^{n} b_i^r \right)^{1-q} \sum_{i=1}^{n} a_i^r \ln_q \left( \frac{a_i^r}{b_i^r} \right) \leq \left( \sum_{i=1}^{n} a_i^r \right) \left( \ln_q \left( \sum_{i=1}^{n} a_i^r \right) - \ln_q \left( \sum_{i=1}^{n} b_i^r \right) \right). \quad (11)
\]
and
\[
\left( \sum_{i=1}^{n} b_i \right)^{1-q} \sum_{i=1}^{n} a_i \ln_q \left( \frac{a_i}{b_i} \right) \leq \left( \sum_{i=1}^{n} a_i \right) \left( \ln_q \left( \sum_{i=1}^{n} a_i \right) - \ln_q \left( \sum_{i=1}^{n} b_i \right) \right),
\]
respectively, under the similar conditions on \( a_i \) and \( b_i \).

If \( f(x) = \log(x) \) we find that \( -xf(x) = xf \left( \frac{1}{x} \right) \) is a concave function. Equation (2) suggests that
\[
- \sum_{i=1}^{n} a_i \log \left( \frac{a_i}{b_i} \right) \leq - a \log \left( \frac{a}{b} \right)
\]
which implies
\[ \sum_{i=1}^{n} a_i \log \left( \frac{b_i}{a_i} \right) \leq a \log \left( \frac{b}{a} \right). \]

In general, \(-xf(x) = xf \left( \frac{1}{x} \right)\) does not hold. For instance, consider \(f(x) = \ln q(x)\), which refers \(-x \ln q(x) = x^d \ln q \left( \frac{1}{x} \right)\). This fact leads us to a new inequality which we consider in the following theorem.

**Theorem 2.4:** Let \(g\) be a real-valued function whose domain contains \(a_1, a_2, \ldots, a_n\) and \(b_1, b_2, \ldots, b_n\), such that \(g(a_i) > 0\) for \(i = 1, 2, \ldots, n\); and \(f : [\tilde{m}_g, \tilde{M}_g] \to \mathbb{R}\) be a function for which \(h(x) = xf \left( \frac{1}{x} \right)\) is a concave function, where \(\tilde{m}_g = \min_{1 \leq i \leq n} \left\{ \frac{g(b_i)}{g(a_i)} \right\}\), and \(\tilde{M}_g = \max_{1 \leq i \leq n} \left\{ \frac{g(b_i)}{g(a_i)} \right\}\). Then,
\[ \sum_{i=1}^{n} g(a_i) f \left( \frac{g(b_i)}{g(a_i)} \right) \leq \left( \sum_{i=1}^{n} g(a_i) \right) f \left( \frac{\sum_{i=1}^{n} g(b_i)}{\sum_{i=1}^{n} g(a_i)} \right). \]

**Proof:** It is easy to find that,
\[
\begin{align*}
\sum_{i=1}^{n} g(a_i) & f \left( \frac{g(b_i)}{g(a_i)} \right) = \sum_{i=1}^{n} g(b_i) \frac{g(a_i)}{g(b_i)} \left( \frac{g(b_i)}{g(a_i)} \right) = \sum_{i=1}^{n} b g(b_i) \left( \frac{g(a_i)}{g(b_i)} \right) \\
& \leq bh \left( \sum_{i=1}^{n} \frac{g(b_i) g(a_i)}{b g(b_i)} \right) \quad \text{[as } h(x) = xf \left( \frac{1}{x} \right) \text{ is concave]} \\
& = bh \left( \frac{1}{b} \sum_{i=1}^{n} g(a_i) \right) = bh \left( \frac{a}{b} \right) = b a f \left( \frac{b}{a} \right) = af \left( \frac{b}{a} \right) \\
& \leq \left( \sum_{i=1}^{n} g(a_i) \right) f \left( \frac{\sum_{i=1}^{n} g(b_i)}{\sum_{i=1}^{n} g(a_i)} \right). \quad \blacksquare
\end{align*}
\]

From now on, we use \(\tilde{m}_g = \min_{1 \leq i \leq n} \left\{ \frac{g(b_i)}{g(a_i)} \right\}\), and \(\tilde{M}_g = \max_{1 \leq i \leq n} \left\{ \frac{g(b_i)}{g(a_i)} \right\}\) throughout this article, where \(g\) is a real valued function whose domain of definition contains \(a_1, a_2, \ldots, a_n\) and \(b_1, b_2, \ldots, b_n\), such that \(g(a_i) > 0\) for all \(i = 1, 2, \ldots, n\).

We know that a twice-differentiable function \(h(x)\) is concave if and only if \(h''(x) \leq 0\), which indicates
\[
\frac{d^2}{dx^2} \left( xf \left( \frac{1}{x} \right) \right) = \frac{1}{x^3} f'' \left( \frac{1}{x} \right) \leq 0.
\]  
(12)

Note that, Equations (12) and (5) are not equivalent, in general. For example, consider the function \(f(x) = \frac{x}{x^2 + 2}\), where \(x > 0\). We can calculate
\[
f'(x) = \frac{2 - x^2}{(x^2 + 2)^2}, \quad \text{and} \quad f''(x) = \frac{2x (x^2 - 6)}{(x^2 + 2)^3}.
\]
Now, Equation (12) suggests that $xf\left(\frac{1}{x}\right)$ is concave when

$$\frac{1}{x^3}f''\left(\frac{1}{x}\right) = \frac{2x^2(1 - 6x^2)}{(2x^2 + 1)^3} \leq 0 \quad \text{that is} \quad x \geq \frac{1}{\sqrt{6}}.$$ 

But, putting the values of $f'$ and $f''$ in Equation (5) we observe that $xf(x)$ is concave when

$$xf''(x) + 2f'(x) = \frac{8 - 12x^2}{(x^2 + 2)^3} \leq 0 \quad \text{when} \quad x \geq \sqrt[3]{\frac{2}{3}}.$$ 

Now, replacing $f(x) = \frac{x}{x^2 + 2}$ and $g(x) = x$ in Theorem 2.4 we observe

$$\sum_{i=1}^{\infty} \frac{a_i^2 b_i}{2a_i^2 + b_i^2} \leq \frac{\left(\sum_{i=1}^{n} a_i\right)^2 \left(\sum_{i=1}^{n} b_i\right)}{2 \left(\sum_{i=1}^{n} a_i\right)^2 + \left(\sum_{i=1}^{n} b_i\right)^2},$$

where $a_i$ and $b_i$ are greater than $\frac{1}{\sqrt{6}}$.

### 3. Log-sum inequalities for commuting Hermitian matrices

We begin this section with a number of basic concepts of matrix analysis. Given any Hermitian matrix $A$, there exists a unitary matrix $U$, such that, $A = U\Lambda(A)U^\dagger$, where $\Lambda(A) = \text{diag}\{a_1, a_2, \ldots, a_n\}$ is a diagonal matrix whose diagonal entries, $a_i$ for $i = 1, 2, \ldots, n$, are the eigenvalues of $A$. Throughout this article $U^\dagger$ denotes the conjugate transpose of the matrix $U$. If the Hermitian matrices $A$ and $B$ commute, that is $AB = BA$, then there exists a unitary matrix $U$, such that, $A = U\Lambda(A)U^\dagger$ and $B = U\Lambda(B)U^\dagger$, hold simultaneously. We also denote $a = \text{trace}(A) = \sum_{i=1}^{n} a_i$. Let $f$ be a continuous real-valued function defined on an interval $J$ and $A$ be a Hermitian matrix with eigenvalues in $J$, then

$$f(A) = U\text{diag}(f(a_1), \ldots, f(a_n))U^\dagger. \quad (13)$$

Now, we demonstrate the results derived in Section 2 for square Hermitian matrices of same size.

**Theorem 3.1:** Let $A$ and $B$ be commuting Hermitian matrices, with the sets of eigenvalues $\Lambda(A) = \{a_1, a_2, \ldots, a_n\}$ and $\Lambda(B) = \{b_1, b_2, \ldots, b_n\}$, respectively. Also, let $g : \Lambda(A) \cup \Lambda(B) \to \mathbb{R}$ be a function, such that, $g(b_i) > 0$ for all $i = 1, 2, \ldots, n$. In addition, $f : [m_g, M_g] \to \mathbb{R}$ is a function for which $xf(x)$ is convex. Then,

$$\text{trace}[g(A)f(g(A)(g(B)^{-1}))] \geq \text{trace}[g(A)]f\left(\frac{\text{trace}(g(A))}{\text{trace}(g(B))}\right).$$

**Proof:** As $A$ and $B$ are commuting Hermitian matrices there exists a unitary matrix $U$, such that

$$g(A) = U\text{diag}(g(a_1), \ldots, g(a_n))U^\dagger,$$
\[ g(B) = U \text{diag}\{g(b_1), \ldots, g(b_n)\} U^\dagger, \quad \text{and} \]
\[
(g(B))^{-1} = U \text{diag}\left\{ \frac{1}{g(b_1)}, \ldots, \frac{1}{g(b_n)} \right\} U^\dagger,
\]
as \(g(b_i) > 0\) for all \(i = 1, 2, \ldots n\). Therefore,
\[
f(g(A)(g(B))^{-1}) = U \text{diag}\left\{ f\left( \frac{g(a_1)}{g(b_1)} \right), \ldots, f\left( \frac{g(a_n)}{g(b_n)} \right) \right\} U^\dagger,
\]
which indicates
\[
g(A)f(g(A)(g(B))^{-1}) = U \text{diag}\left\{ g(a_1)f\left( \frac{g(a_1)}{g(b_1)} \right), \ldots, g(a_n)f\left( \frac{g(a_n)}{g(b_n)} \right) \right\} U^\dagger.
\]
Note that, \(g(a_i)f\left( \frac{g(a_i)}{g(b_i)} \right)\) are the eigenvalues of the matrix \(g(A)f(g(A)(g(B))^{-1})\) for \(i = 1, 2, \ldots n\). Hence,
\[
\text{trace}(g(A)f(g(A)(g(B))^{-1})) = \sum_{i=1}^{n} g(a_i)f\left( \frac{g(a_i)}{g(b_i)} \right).
\]
Also, \(\sum_{i=1}^{n} g(a_i) = \text{trace}(g(A))\) and \(\sum_{i=1}^{n} g(b_i) = \text{trace}(g(B))\). Applying Theorem 2.1 we obtain
\[
\sum_{i=1}^{n} g(a_i)f\left( \frac{g(a_i)}{g(b_i)} \right) = \left( \sum_{i=1}^{n} g(a_i) \right) f\left( \frac{\sum_{i=1}^{n} g(a_i)}{\sum_{i=1}^{n} g(b_i)} \right) = \text{trace}[g(A)]f\left( \frac{\text{trace}(g(A))}{\text{trace}(g(B))} \right).
\]
Combining we get the result.

A Hermitian matrix \(A\) is called positive semidefinite if \(x^*Ax \geq 0\) for all \(x \in \mathbb{C}^n\). A Hermitian matrix \(A\) is also called positive definite if \(x^*Ax > 0\) for all \(0 \neq x \in \mathbb{C}^n\), that is, a positive definite matrix is positive semi-definite and invertible.

Considering \(f(x) = \log(x)\) and \(g(x) = x\) for \(x > 0\) in Theorem 3.1, the following corollary holds trivially since \(\frac{d^2}{dx^2}(x \log(x)) = \frac{1}{x} > 0\).

**Corollary 3.2:** Given positive definite, commuting matrices \(A\) and \(B\),
\[
\text{trace}(A \log(A)) - \text{trace}(A \log(B)) \geq \text{trace}(A) \log\left( \frac{\text{trace}(A)}{\text{trace}(B)} \right).
\]
(14)

The inequality in the above corollary is often called the Peierls–Bogoliubov inequality [6, Theorem 2] which gives a lower bound of the quantum relative entropy for positive definite matrices \(A\) and \(B\). Theorem 3.1 has a general form of the Peierls–Bogoliubov inequality under the condition \(AB = BA\).

Theorem 3.1 is a matrix-theoretic counterpart of Theorem 2.1, which assists us to establish a number of matrix inequalities, which are derived in Section 2 for real functions. Some of these matrix inequalities have immediate consequences in quantum information theory. Consider the following examples.
Example 3.3: The matrix counterpart of Equation (6) is represented as
\[
\text{trace} [A^r \log(A^r B^{-r})] \geq \text{trace}(A^r) [\log(\text{trace}(A^r)) - \log(\text{trace}(B^r))],
\]
where \( A \) and \( B \) are Hermitian commuting matrices as well as \( B \) is positive definite. The inequality (15) can be obtained by putting \( g(x) := x^r \) and \( f(x) := \log x \) for \( x > 0 \) in Theorem 3.1. Putting \( r = 1 \) in the above inequality we have
\[
\text{trace} [A \log(AB^{-1})] \geq \text{trace}(A) [\log(\text{trace}(A)) - \log(\text{trace}(B))].
\]
Considering \( AB = BA \), we can prove that \( \log(AB^{-1}) = \log(A) - \log(B) \). Replacing it at the left hand side, we have the inequality (14) in Corollary 3.2. See [6, Theorem 2] and [7, Theorem 3.3] for the Peierls–Bogoliubov inequality.

Recall that, in quantum information theory [8] a quantum state is represented by a density matrix \( \rho \) which is a positive semidefinite, Hermitian matrix with \( \text{trace}(\rho) = 1 \). The von-Neumann entropy is a well-known measure of quantum information which is given by \( -\text{trace}(\rho \log(\rho)) \) for a density matrix \( \rho \). Given two density matrices \( \rho \) and \( \sigma \) the quantum relative entropy is determined by \( D(\rho || \sigma) = \text{trace}[\rho (\log(\rho) - \log(\sigma))] \).

Considering \( \text{trace}(A) = \text{trace}(B) = 1 \) in Equation (14), we find that \( A \) and \( B \) are density matrices. Then, the quantum relative entropy given two density matrices \( A \) and \( B \),
\[
D(\rho || \sigma) = \text{trace} [A(\log(A) - \log(B))] \geq 0.
\]
Considering \( B = I \), the identity matrix of order \( n \), in Equation (15) we observe that
\[
\text{trace}[A^r \log A^r] \geq \text{trace}(A^r) [\log(\text{trace}(A^r)) - \log(n)].
\]
Putting \( r = 1 \) we have \( \text{trace}[A \log A] \geq \text{trace}(A) [\log(\text{trace}(A)) - \log(n)] \). In addition, if \( A \) is a density matrix, that is, \( \text{trace}(A) = 1 \) then the above inequality indicates \( \text{trace}[A \log A] \geq -\log(n) \). The von-Neumann entropy of \( A \) is \( -\text{trace}[A \log A] \leq \log(n) \), which is its maximum value.

Example 3.4: The matrix counterpart of Equation (9) will be given by
\[
(\text{trace}(B^r))^{1 - q} \text{trace} [A^r \ln_q (A^r B^{-r})] \geq (\text{trace}(A^r)) [\ln_q (\text{trace}(A^r)) - \ln_q (\text{trace}(B^r))],
\]
where \( A \) and \( B \) are Hermitian commuting matrices as well as \( B \) is positive definite. The above inequality can be obtained by putting \( g(x) := x^r \) and \( f(x) := \ln_q x \) for \( x > 0 \) and \( q \leq 2 \) in Theorem 3.1 with simple calculations. For \( r = 1 \) we have
\[
(\text{trace}(B))^{1 - q} \text{trace} [A \ln_q (AB^{-1})] \geq (\text{trace}(A)) [\ln_q (\text{trace}(A)) - \ln_q (\text{trace}(B))].
\]
Note that in the above two inequalities we consider \( q \leq 2 \) to ensure convexity of \( x \ln_q(x) \). For \( q \geq 2 \), Equation (11) suggests
\[
(\text{trace}(B^r))^{1 - q} \text{trace} [A^r \ln_q (A^r B^{-r})] \leq (\text{trace}(A^r)) [\ln_q (\text{trace}(A^r)) - \ln_q (\text{trace}(B^r))],
\]
for Hermitian commuting matrices \( A \) and \( B \) as well as positive definite matrix \( B \).
Theorem 2.4 also indicates a trace inequality, which we mention below without a proof.

**Theorem 3.5:** Let $A$ and $B$ be commuting Hermitian matrices, with the sets of eigenvalues $\Lambda(A) = \{a_1, a_2, \ldots, a_n\}$ and $\Lambda(B) = \{b_1, b_2, \ldots, b_n\}$, respectively. Also, let $g : \Lambda(A) \cup \Lambda(B) \to \mathbb{R}$ be a function, such that, $g(b_i) > 0$ for all $i = 1, 2, \ldots, n$. In addition, $f : [\tilde{m}_g, \tilde{M}_g] \to \mathbb{R}$ is a function for which $h(x) = xf \left( \frac{1}{x} \right)$ is concave. Then,

$$\text{trace}[g(A)f(g(A)(g(B))^{-1})] \leq \text{trace}[g(A)f \left( \frac{\text{trace}(g(A))}{\text{trace}(g(B))} \right)].$$

**4. Log-sum inequalities with Löwner partial order**

Recall that given Hermitian matrices $A$ and $B$ we write $A \geq B$ or $B \leq A$ if the matrix $A - B$ is positive semidefinite. This ordering is called the Löwner partial order. If $A$ is positive definite we write $A > 0$.

Let $M_n(J)$ be the set of all complex matrices of order $n$ with eigenvalues in $J \subset \mathbb{R}$. Equation (13) indicates that we can redefine a function $f : J \to \mathbb{R}$ as a matrix function $f : M_n(J) \to M_n(\mathbb{C})$, the set of all complex matrices of order $n$. Let $A$ and $B$ be two Hermitian matrices of order $n$ with eigenvalues in $J$. Now, a continuous function $f : J \to \mathbb{R}$ is said to be a matrix monotone of order $n$ if $A \leq B$ (Hermitian) implies $f(A) \leq f(B)$. Also, $f$ is said to be a matrix convex function of order $n$ if

$$f(\lambda A + (1 - \lambda)B) \leq \lambda f(A) + (1 - \lambda)f(B),$$

for all $\lambda \in [0, 1]$ and every pair of Hermitian matrices $A$ and $B$ of order $n$. An operator monotone function is a matrix monotone function for any natural order $n$. Similarly, an operator convex function is a matrix convex function for any natural order $n$. A function $f$ is operator concave if $-f$ is operator convex. Well-known operator monotone functions are $f_1(t) = t^r$ for $0 < r \leq 1$ for $t \in [0, \infty)$, as well as $f_2(t) = \log(t)$ for $t \in (0, \infty)$. In addition, $g(t) = t^r$ is operator convex on $(0, \infty)$ for $-1 \leq r \leq 0$ and $1 \leq r \leq 2$, see [9,10]. The operator Jensen inequality [11,12] plays a crucial role for further development, which we mention below:

**Lemma 4.1 ([11]):** If $f$ is a continuous, real function defined on an interval $[0, \alpha)$ with $\alpha \leq \infty$, the following conditions are equivalent.

1. $f$ is operator convex and $f(0) \leq 0$.
2. $f(A^\dagger XA) \leq A^\dagger f(X)A$ for all $A$ with $||A|| \leq 1$ and every Hermitian $X$ with spectrum in $[0, \alpha)$.
3. $f(A^\dagger XA + B^\dagger YB) \leq A^\dagger f(X)A + B^\dagger f(Y)B$ for all $A, B$ with $A^\dagger A + B^\dagger B \leq I$ and all $X, Y$ with spectra in $[0, \alpha)$.
4. $f(PXP) \leq Pf(X)P$ for every projection $P$ and every Hermitian $X$ with spectrum in $[0, \alpha)$.

We utilize Dirac’s bra-ket notation [13] for simplifying our notations. Recall that $|u\rangle$ denotes a column vector of length $n$, or equivalently an $n \times 1$ matrix. The conjugate transpose of $|u\rangle$ is given by $\langle u|$, which is a row vector of length $n$. Note that, $|u\rangle \langle u|$ is a Hermitian
matrix of order $n$. Now the spectral decomposition of any Hermitian matrix $A$ can be written as $A = \sum_{i=1}^{n} \lambda_i |u_i\rangle \langle u_i|$ where $|u_i\rangle$ is a normalized eigenvector corresponding to the eigenvalue $\lambda_i$.

The superadditivity of a mean has been shown in [14] and the corresponding advanced results for a mean and a perspective have been studied in [15,16]. See also [17,18] for recent results. In the sequel, we use the standard notation (see [14, Equation (3.8)]) $A_{\sigma_f}B := A^{\frac{1}{2}}f(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{-\frac{1}{2}}$ for positive definite matrices $A$ and $B$. The following proposition is proven by the elementary calculations under a certain condition.

**Proposition 4.2:** Let $A_1, A_2, \ldots, A_m$ and $B_1, B_2, \ldots, B_m$ be two sets of positive definite matrices of order $n$, and $A \geq mI$, where $\sum_{i=1}^{m} A_i =: A$. Also, let $f$ be an operator concave function on an interval $[0, \alpha)$ containing the eigenvalues of $B_i$ and $B = \sum_{i=1}^{m} B_i$. Then

$$\sum_{i=1}^{m} A_i \sigma_f B_i \leq A \sigma_f B.$$

**Proof:** As each $B_i$ is a Hermitian matrix, the spectral decomposition of $B_i$ can be written as

$$B_i = \sum_{k=1}^{n} \lambda_k^{(i)} |u_k^{(i)}\rangle \langle u_k^{(i)}|,$$

where $|u_k^{(i)}\rangle$ is an eigenvector associated to the eigenvalue $\lambda_k^{(i)}$ of $B_i$. Then we have

$$A_i^{-\frac{1}{2}}B_i A_i^{-\frac{1}{2}} = \sum_{k=1}^{n} \lambda_k^{(i)} |w_k^{(i)}\rangle \langle w_k^{(i)}|,$$

where $|w_k^{(i)}\rangle := A_i^{-\frac{1}{2}}|u_k^{(i)}\rangle$. Now for a function $f$ we have

$$f \left( A_i^{-\frac{1}{2}}B_i A_i^{-\frac{1}{2}} \right) = \sum_{k=1}^{n} f \left( \lambda_k^{(i)} \right) |w_k^{(i)}\rangle \langle w_k^{(i)}|.$$

Combining we get

$$A_i \sigma_f B_i = \sum_{k=1}^{n} f \left( \lambda_k^{(i)} \right) A_i^{\frac{1}{2}} |w_k^{(i)}\rangle \langle w_k^{(i)}| A_i^{\frac{1}{2}} = \sum_{k=1}^{n} f \left( \lambda_k^{(i)} \right) |u_k^{(i)}\rangle \langle u_k^{(i)}| = f(B_i).$$

Summing over $i$ we find

$$\sum_{i=1}^{m} A_i \sigma_f B_i = \sum_{i=1}^{m} f(B_i). \quad (16)$$

If $A \geq mI$, then we have $\sum_{i=1}^{m} A^{-\frac{1}{2}}A^{-\frac{1}{2}} = mA^{-1} \leq I$. Now, applying Lemma 4.1 we find that

$$A^{-\frac{1}{2}} \sum_{i=1}^{m} f(B_i) A^{-\frac{1}{2}} = \sum_{i=1}^{m} A^{-\frac{1}{2}} f(B_i) A^{-\frac{1}{2}} \leq f \left( \sum_{i=1}^{m} A^{-\frac{1}{2}} B_i A^{-\frac{1}{2}} \right)$$

$$= f \left( A^{-\frac{1}{2}} \sum_{i=1}^{m} B_i A^{-\frac{1}{2}} \right) = f \left( A^{-\frac{1}{2}} BA^{-\frac{1}{2}} \right).$$
Multiplying $A^{\frac{1}{2}}$ in both side of the above inequality we get

$$\sum_{i=1}^{m} f(B_i) \leq A \sigma f B.$$ 

Putting the value of $f(B_i)$ from Equation (16) we find the result. 

**Remark 4.1:** If $A_i$ and $B_i$ are replaced by positive real numbers $a_i$ and $b_i$ respectively in the above theorem, we get

$$\sum_{i=1}^{m} a_i \sigma f b_i \leq a \sigma f b,$$

where $a = \sum_{i=1}^{m} a_i$ and $b = \sum_{i=1}^{m} b_i$ as well as $f$ is a convex function. Simplifying we get

$$\sum_{i=1}^{m} a_if\left(\frac{b_i}{a_i}\right) \leq a f\left(\frac{b}{a}\right). \quad (17)$$

Considering $f(x) = \log(x)$ we get the usual log-sum inequality.

**Corollary 4.3:** Let $A_1, A_2, \ldots, A_m$ and $B_1, B_2, \ldots, B_m$ be two sets of positive definite Hermitian operators with $A = \sum_{i=1}^{m} A_i$ and $B = \sum_{i=1}^{m} B_i$, such that, $A = B$. Also, let $f$ be an operator concave function on an interval $[0, \alpha)$ containing the eigenvalues of $B_i$ and $B$ as well as $f(1) = 0$. Then

$$\sum_{i=1}^{m} A_i \sigma f B_i \leq 0.$$

**Proof:** The proof follows trivially from Proposition 4.2. 

**Remark 4.2:** The operator Shannon inequality was given in [19]

$$\sum_{i=1}^{m} A_i \sigma \log B_i \leq 0$$

under the assumption $\sum_{i=1}^{m} A_i = \sum_{i=1}^{m} B_i = I$. Our condition in Corollary 4.3 is slightly weaker than this assumption. We also observe that the operator Shannon inequality holds for any operator concave function $f$.

If every $A_i$ is expansive (i.e. $A_i \geq I$), then the condition $A \geq mI$ in Proposition 4.2 is satisfied. However, we have not obtained a proper result for contractive condition such as $A_i \leq I$. Closing this section, we present a result which does not impose an additional condition on the matrices $A_i$. We need the following known facts for proving the next theorem:

**Lemma 4.4 ([20]):** Let $X$ and $A$ be bounded linear operators on a Hilbert space $H$. Suppose that $X \geq 0$ and $||A|| \leq 1$. If $f : [0, \infty) \to [0, \infty)$ is an operator monotone function defined on $[0, \infty)$, then $A^\dagger f(X)A \leq f(A^\dagger X A)$.
Lemma 4.5 ([9, p.14]): For any square matrix $X_i$ and positive definite matrices $A_i$ we have

$$\sum_{i=1}^{m} X_i^\dagger A_i^{-1} X_i \geq \left( \sum_{i=1}^{m} X_i \right)^\dagger \left( \sum_{i=1}^{m} A_i \right)^{-1} \left( \sum_{i=1}^{m} X_i \right).$$

Theorem 4.6: Let $A_1, A_2, \ldots, A_m$ and $B_1, B_2, \ldots, B_m$ be two sets of positive definite matrices, as well as $A = \sum_{i=1}^{m} A_i$ and $B = \sum_{i=1}^{m} B_i$. Also, $f : [0, \infty) \to [0, \infty)$ is an operator monotone function. Then

$$\frac{1}{m} \left( \sum_{i=1}^{m} B_i^\frac{1}{2} \right) \left[ \sum_{i=1}^{m} f \left( B_i^\frac{1}{2} A_i^{-1} B_i^\frac{1}{2} \right) \right]^{-1} \left( \sum_{i=1}^{m} B_i^\frac{1}{2} \right) \leq \left( B^{-1} \sigma_f A^{-1} \right)^{-1},$$

and

$$\sum_{i=1}^{m} A_i^\frac{1}{2} (B_i^{-1} \sigma_f A_i^{-1}) A_i^\frac{1}{2} \geq \frac{1}{m} \left( \sum_{i=1}^{m} A_i^\frac{1}{2} \right) (B^{-1} \sigma_f A^{-1}) \left( \sum_{i=1}^{m} A_i^\frac{1}{2} \right).$$

Proof: We can write

$$B_i^\frac{1}{2} A_i^{-1} B_i^\frac{1}{2} = B_i^\frac{1}{2} B^{-\frac{1}{2} B_i^\frac{1}{2}} A_i^{-1} \left( B_i^\frac{1}{2} B^{-\frac{1}{2} B_i^\frac{1}{2}} \right) B_i^\frac{1}{2} = B_i^\frac{1}{2} B^{-\frac{1}{2} \left( B_i^\frac{1}{2} A_i^{-1} B_i^\frac{1}{2} \right)} B_i^\frac{1}{2} B_i^\frac{1}{2}.$$

We have $A \geq A_i$, that is $A^{-1} \leq A_i^{-1}$. Also for any complex square matrix $X$ we have $X^\dagger A^{-1} X \leq X^\dagger A_i^{-1} X$. Applying these together we obtain

$$B_i^\frac{1}{2} A_i^{-1} B_i^\frac{1}{2} \geq B_i^\frac{1}{2} B^{-\frac{1}{2} \left( B_i^\frac{1}{2} A_i^{-1} B_i^\frac{1}{2} \right)} B_i^\frac{1}{2} B_i^\frac{1}{2}.$$

As $f(t)$ is a matrix monotone function, we can write

$$f \left( B_i^\frac{1}{2} A_i^{-1} B_i^\frac{1}{2} \right) \geq f \left( B_i^\frac{1}{2} B^{-\frac{1}{2} \left( B_i^\frac{1}{2} A_i^{-1} B_i^\frac{1}{2} \right)} B_i^\frac{1}{2} B_i^\frac{1}{2} \right).$$

From $B \geq B_i$ for all $i = 1, 2, \ldots, n$, we can see

$$I \geq B^{-\frac{1}{2} B_i B^{-\frac{1}{2} B_i B^{-\frac{1}{2}}}} = \left( B_i^\frac{1}{2} B^{-\frac{1}{2} B_i B^{-\frac{1}{2}}} \right)^\dagger B_i^\frac{1}{2} B^{-\frac{1}{2} B_i B^{-\frac{1}{2}}}$$

which implies $1 \geq ||(B_i^\frac{1}{2} B^{-\frac{1}{2}})^\dagger B_i^\frac{1}{2} B^{-\frac{1}{2}}|| = ||B_i^\frac{1}{2} B^{-\frac{1}{2}}||^2$, since $||A|| = ||A^\dagger A||^{\frac{1}{2}}$ for every operator $A$ in general. Thus we have $||B_i^\frac{1}{2} B^{-\frac{1}{2}}|| \leq 1$. We also have $||B^{-\frac{1}{2}} B_i^\frac{1}{2}|| \leq 1$ so that we have the following inequality by Lemma 4.4,

$$f \left( B_i^\frac{1}{2} B^{-\frac{1}{2} \left( B_i^\frac{1}{2} A_i^{-1} B_i^\frac{1}{2} \right)} B_i^\frac{1}{2} B_i^\frac{1}{2} \right) \geq B_i^\frac{1}{2} \left( B^{-1} \sigma_f A^{-1} \right) B_i^\frac{1}{2}.$$

That is,

$$f \left( B_i^\frac{1}{2} A_i^{-1} B_i^\frac{1}{2} \right) \geq B_i^\frac{1}{2} \left( B^{-1} \sigma_f A^{-1} \right) B_i^\frac{1}{2}.$$
Multiplying $B_i^{-\frac{1}{2}}$ to the both sides, we have
\[ B_i^{-1}\sigma_i A_i^{-1} \geq B^{-1}\sigma A^{-1}. \] (20)

Taking an inverse of the both sides and a summation over $i = 1, \ldots, m$, we have
\[ \frac{1}{m} \sum_{i=1}^{m} (B_i^{-1}\sigma_i A_i^{-1})^{-1} \leq (B^{-1}\sigma A^{-1})^{-1}. \] (21)

Applying Lemma 4.5 to the left hand side in (21), we have
\[ \frac{1}{m} \sum_{i=1}^{m} (B_i^{-1}\sigma_i A_i^{-1})^{-1} \geq \frac{1}{m} \left( \sum_{i=1}^{m} B_i^{\frac{1}{2}} \right) \left( \sum_{i=1}^{m} f\left( B_i^{\frac{1}{2}} A_i^{-\frac{1}{2}} B_i^{\frac{1}{2}} \right) \right)^{-1} \left( \sum_{i=1}^{m} B_i^{\frac{1}{2}} \right). \] (22)

Combining (21) and (22), we get the inequality (18).

To prove the inequality (19), we multiply $A_i^{\frac{1}{2}}$ to the both sides in (20) and take a summation over $i = 1, \ldots, m$, we have
\[ \sum_{i=1}^{m} A_i^{\frac{1}{2}} (B_i^{-1}\sigma_i A_i^{-1}) A_i^{\frac{1}{2}} \geq \sum_{i=1}^{m} A_i^{\frac{1}{2}} (B^{-1}\sigma A^{-1}) A_i^{\frac{1}{2}}. \] (23)

Applying Lemma 4.5 to the right hand side in (23) and taking an account for $B^{-1}\sigma A^{-1} = ((B^{-1}\sigma A^{-1})^{-1})^{-1}$, we have
\[ \sum_{i=1}^{m} A_i^{\frac{1}{2}} (B^{-1}\sigma A^{-1}) A_i^{\frac{1}{2}} \geq \left( \sum_{i=1}^{m} A_i^{\frac{1}{2}} \right) \left( \sum_{i=1}^{m} (B^{-1}\sigma A^{-1})^{-1} \right)^{-1} \left( \sum_{i=1}^{m} A_i^{\frac{1}{2}} \right). \] (24)

Combining (23) and (24), we get the inequality (19).

**Remark 4.3:** Considering $a_i$ and $b_i$ are positive real numbers and replacing $A_i =: a_i$ and $B_i =: b_i$ as well as $f : [0, \infty) \to [0, \infty)$ is a monotone increasing function we have from Equations (18) and (19) respectively,
\[ \sum_{i=1}^{m} f\left( \frac{b_i}{a_i} \right) = \frac{1}{m} \left( \sum_{i=1}^{m} b_i^{\frac{1}{2}} \right)^2 \frac{1}{b} f\left( \frac{b}{a} \right) \] (25)
and
\[ \sum_{i=1}^{m} a_i f\left( \frac{b_i}{a_i} \right) = \frac{1}{m} \left( \sum_{i=1}^{m} a_i^{\frac{1}{2}} \right)^2 \frac{1}{b} f\left( \frac{b}{a} \right). \] (26)

If we take a special case $a_i := \frac{a}{m}$ and $b_i := \frac{b}{m}$ for $i = 1, \ldots, m$ in both (25) and (26), then we reach to the trivial inequality $(m - 1)f\left( \frac{b}{a} \right) \geq 0$. 


5. Conclusion

The log-sum inequality which is mentioned in Equation (2) plays a crucial role in classical and quantum information theory. In this article, we present a number of analogous inequalities. Inequality (2) depends on the convexity of the function \( x \log(x) \) which was relaxed utilizing the function \( xf(x) \), in [4]. We illustrate that the concavity of \( xf\left(\frac{1}{x}\right) \) is also efficient for deriving similar inequalities. We discuss these inequalities with two functions instead of the single function \( f \), which increase their scope of applications. In this context, the log-sum inequality for \( q \)-deformed logarithm is illustrated. We elaborate these results for the commuting Hermitian matrices as trace-form inequalities. Later we discuss the log-sum inequality for the operator monotone and convex functions in the context of Löwner partial order relation of the positive semi-definite matrices, which is an application of the Hansen operator inequality. The following problem may be attempted in future.

Both inequalities in (25) and (26) do not have same form of log-sum inequality (17). Therefore we have to conclude that the restricted condition (so-called expansivity of matrices \( A_i \)) given in Proposition 4.2 leads us to obtain the log-sum inequality for non-commutative matrices. However, we could not obtain such an inequality without a restricted condition in Theorem 4.6. In the further studies on this topic, we would like to obtain the log-sum inequality for non-commutative contractive matrices. We hope that the obtained results will be useful to study the parametric information theory in the future.

Acknowledgments

The authors would like to thank the referees for their careful and insightful comments to improve our manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Supriyo Dutta was a post-doctoral fellow at the Indian Institute of Technology Kharagpur during this work. Shigeru Furuichi was partially supported by JSPS KAKENHI [grant number 21K03341].

ORCID

Supriyo Dutta http://orcid.org/0000-0003-2369-7915
Shigeru Furuichi http://orcid.org/0000-0002-9929-0954

References

[1] Cover TM, Thomas JA. Elements of information theory. New Delhi: John Wiley & Sons; 1999.
[2] Jensen JLWV, et al. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Math. 1906;30:175–193.
[3] Niculescu C, Persson LE. Convex functions and their applications. New York: Springer; 2006.
[4] Csiszár I, Shields PC, et al. Information theory and statistics: a tutorial. Found Trends Commun Inf Theory. 2004;1(4):417–528.
[5] Kac V, Cheung P. Quantum calculus. Berlin: Springer Science & Business Media; 2001.
[6] Bebiano N, Providência Jr J, Lemos R. Matrix inequalities in statistical mechanics. Linear Algebra Appl. 2004;376(1):265–273.
[7] Furuichi S, Yanagi K, Kuriyama K. Fundamental properties of Tsallis relative entropy. J Math Phys. 2004;45(12):4868–4877.
[8] Wilde MM. Quantum information theory. Cambridge: Cambridge University Press; 2013.
[9] Zhan X. Matrix inequalities. Berlin: Springer; 2002.
[10] Bhatia R. Matrix analysis. Vol. 169. Berlin: Springer Science & Business Media; 2013.
[11] Hansen F, Pedersen GK. Jensen’s inequality for operators and Löwner’s theorem. Math Ann. 1982;258(3):229–241.
[12] Hansen F, Pedersen GK. Jensen’s operator inequality. Bull London Math Soc. 2003;35(4):553–564.
[13] Dirac PAM. A new notation for quantum mechanics. Math Proc Cambridge Philos Soc. 1939;35(5):416–418.
[14] Kubo F, Ando T. Means of positive linear operators. Math Ann. 1980;246:205–224.
[15] Dragomir SS. Operator quasilinearity of some functionals associated with Davis–Choi–Jensen’s inequality for positive maps. Bull Aust Math Soc. 2017;95:322–332.
[16] Nikoufar I. Some properties of the Tsallis relative operator $\phi$–entropy. J Math Phys Anal Geom. 2021;17(2):216–232.
[17] Gümüş IH, Moradi HR, Sababheh M. More accurate operator means inequalities. J Math Anal Appl. 2018;465(1):267–280.
[18] Furuichi S, Moradi HR, Sababheh M. New sharp inequalities for operator means. Linear Multilinear A. 2019;67(8):1567–1578.
[19] Furuta T. Parametric extensions of Shannon inequality and its reverse one in Hilbert space operators. Linear Algebra Appl. 2004;381:219–235.
[20] Hansen F. An operator inequality. Math Ann. 1980;246(3):249–250.