Generalized log-sum inequalities

Supriyo Dutta
Centre for Theoretical Studies
Indian Institute of Technology Kharagpur
Kharagpur, India - 721302
Email: dosupriyo@gmail.com

Shigeru Furuichi∗
Department of Information Science
College of Humanities and Sciences, Nihon University
3-25-40, Sakurajyousui, Setagaya-Ku, Tokyo, 156-8550, Japan
Email: furuichi@chs.nihon-u.ac.jp

Abstract
In information theory, the so-called log-sum inequality is fundamental and a kind of generalization of the non-nагativity for the relative entropy. In this paper, we show the generalized log-sum inequality for two functions defined for scalars. We also give a new result for commutative matrices. In addition, we demonstrate further results for general non-commutative positive semi-definite matrices.

Keywords : log-sum inequality, operator convexity/concavity, operator monotone function and positive definite matrices

2010 Mathematics Subject Classification : 15A39, 15A45 and 47A63

1 Introduction
The log-sum inequality is a variant of the Jensen inequality of convex functions. It plays a crucial role in classical information theory for proving the Gibbs’ inequality or the convexity of Kullback-Leibler divergence [1].

Recall that, a function \( f : X \rightarrow \mathbb{R} \) 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

\[
\sum_{i=1}^{n} t_i f(x_i) \geq f \left( \sum_{i=1}^{n} t_i x_i \right), \text{ where } \sum_{i=1}^{n} t_i = 1 \text{ and } 0 \leq t_i \leq 1. \tag{1}
\]

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, considering it in the Jensen inequality we have

\[
\sum_{i=1}^{n} \frac{a_i}{b_j} \log \left( \frac{a_i}{b_j} \right) \geq \left( \sum_{i=1}^{n} \frac{a_i}{b_j} \right) \log \left( \sum_{j=1}^{n} \frac{a_i}{b_j} \right), \tag{2}
\]

where \( t_i = \frac{b_i}{\sum_{j=1}^{n} b_j} \) and \( x_i = \frac{a_i}{b_i} \). Simplifying this inequality by replacing \( a = \sum_{j=1}^{n} a_j \) and \( b = \sum_{j=1}^{n} b_j \) we obtain,

\[
\sum_{i=1}^{n} a_i \log \left( \frac{a_i}{b_i} \right) \geq a \log \left( \frac{a}{b} \right), \tag{3}
\]
which is the standard log-sum inequality. The inequality is valid for \( n = \infty \) provided \( a < \infty \) as well as \( b < \infty \). In \cite{4}, an analog of log-sum inequality is proved which is

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

when \( f \) is strictly convex at \( c = \frac{a}{b} \). The equality holds if and only if \( a_i = cb_i \), for all \( i \).

This article generalizes the log-sum inequality from multiple perspectives. First, we generalize log-sum inequality with two real-valued functions, which we mention in Theorem 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. But considering the concavity of a function, we may find analogous results. It is important as there are classes of functions whose convexity depends on specific values of a parameter, for instance, the \( q \)-deformed logarithm. Another objective for generalizing the log-sum inequality is investigating its matrix analogues. Here we observe that all the inequalities derived for the real functions can be easily extended as trace-from-inequality for commuting self-adjoint 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 a non-trivial problem, which we have solved under several conditions.

This article is distributed into five sections and an appendix. The appendix contains a number of important properties of \( q \)-deformed logarithm which we use in this article. Necessary preliminary ideas are discussed before using them in a mathematical derivation. In section 2, we consider generalized log-sum inequalities for real-valued functions. Here we discuss the log-sum inequality for deformed logarithms. Section 3 is dedicated to discuss log-sum inequality as a trace-form-inequality for commuting self-adjoint matrices. We observe that it has a number of immediate 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 have mentioned at the beginning of section 4. Then we conclude the article.

## 2 Generalized log-sum inequality for real functions

In this section, we discuss the log-sum inequality for real valued functions defined on real numbers.

**Theorem 1.** Let \( g \) be a real valued function whose domain contains the elements \( a_1, a_2, \ldots, a_n \) and \( b_1, b_2, \ldots, b_n \), such that \( g(b_i) > 0 \) for all \( i \). Consider another function \( f : [m_g, M_g] \to \mathbb{R} \) for which \( h(x) = xf(x) \) is convex. Where \( m_g := m(g, a_i, b_i) := \min \left\{ \frac{g(a_i)}{g(b_i)} \right\} \) and \( M_g := m(g, a_i, b_i) := \max \left\{ \frac{g(a_i)}{g(b_i)} \right\} \) are used throughout this paper. Then,

\[
\sum_{i=1}^{n} g(a_i) f \left( \frac{g(a_i)}{g(b_i)} \right) \geq \left( \sum_{i=1}^{n} g(a_i) \right) f \left( \frac{\sum_{i=1}^{n} g(a_i)}{\sum_{i=1}^{n} g(b_i)} \right). \tag{5}
\]

**Proof.** Consider two real numbers \( a \) and \( b \) such that \( a = \sum_{i=1}^{n} g(a_i) \) and \( b = \sum_{i=1}^{n} g(b_i) > 0 \). Note that, \( \frac{a}{b} \) is a convex combination of \( \frac{a_i}{b_i} \) for \( i = 1, 2, \ldots, n \) as we can express \( \frac{a}{b} = \sum_{i=1}^{n} \frac{g(b_i) g(a_i)}{b g(b_i)} \). Clearly, \( \frac{a}{b} \) belongs to the convex set \([m_g, M_g]\). Now,

\[
\sum_{i=1}^{n} g(a_i) f \left( \frac{g(a_i)}{g(b_i)} \right) = \sum_{i=1}^{n} g(b_i) g(a_i) f \left( \frac{g(a_i)}{g(b_i)} \right) = \sum_{i=1}^{n} b g(b_i) h \left( \frac{g(a_i)}{g(b_i)} \right). \tag{6}
\]

As \( g(b_i) > 0 \) for all \( i \) 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 of equation (1) indicates

\[
\sum_{i=1}^{n} b \frac{g(b_i)}{b} h \left( \frac{g(a_i)}{g(b_i)} \right) \geq bh \left( \sum_{i=1}^{n} \frac{g(b_i) g(a_i)}{b g(b_i)} \right) = bh \left( \frac{1}{b} \sum_{i=1}^{n} g(a_i) \right) = bh \left( \frac{a}{b} \right). \tag{7}
\]
Expanding \( h \) we get
\[
\begin{align*}
\exp h \left( \frac{a}{b} \right) & = b \exp \left( \frac{a}{b} \right) = a f \left( \frac{a}{b} \right) = \left( \sum_{i=1}^{n} g(a_i) \right) f \left( \frac{\sum_{i=1}^{n} g(a_i)}{\sum_{i=1}^{n} g(b_i)} \right).
\end{align*}
\]  

Combining, we find the result.

Defining \( g(x) = x \) and \( f(x) = \log(x) \) for \( x \in \mathbb{R}_{\geq 0} \), we observe that \( h(x) = x \log(x) \) is a convex function. Then, the Theorem 1 leads us to the log-sum inequality mentioned in equation (3).

Theorem 1 is a generalization of log-sum inequality which expands the scope of its application. Note here that the domain of the function \( g \) need not be a convex set. In fact, we may consider \( a_i \) and \( b_i \) from arbitrary set such that \( g(b_i) > 0 \) for all \( i \). This condition is an essential for Theorem 1. In other words, the range of \( g \) must have an non-empty intersection with the set of positive reals.

It can be proved that, a twice-differentiable real valued function \( h(x) \) on \( \mathbb{R} \) is convex if and only if \( h''(x) \geq 0 \). Assuming \( f \) as a twice-differentiable function in \([m_g, M_g] \) we have
\[
\frac{d^2}{dx^2} (xf(x)) = xf''(x) + 2f'(x) \geq 0.
\]

If \( m_g \geq 0 \), that any monotone increasing and convex function \( f \) defined on \([m_g, M_g] \) fulfills equation (9). Now, we are in position to consider a few special cases of Theorem 1.

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

Thus we have
\[
\sum_{i=1}^{n} ra_i \{ \log(a_i) - \log(b_i) \} \geq \left( \sum_{i=1}^{n} a_i^r \right) \left\{ \log \left( \sum_{i=1}^{n} a_i^r \right) - \log \left( \sum_{i=1}^{n} b_i^r \right) \right\}.
\]  

Our next example consider the deformed logarithm which plays a crucial role in different branches of mathematics and mathematical physics. We refer the appendix for a number of characteristics of this function.

**Example 2.** The \( q \)-deformed logarithm is defined by
\[
f(x) = \ln_q(x) = \frac{x^{1-q} - 1}{1-q},
\]
which is also known as \( q \)-logarithm [5]. Therefore, \( h(x) = xf(x) = \frac{x^{2-q} - x}{1-q} \). Differentiating we get \( h'(x) = \frac{(2-q)x^{1-q} - 1}{1-q} \) and \( h''(x) = \frac{2-q}{x} \geq 0 \) when \( q < 2 \). Therefore \( h(x) \) is a convex function. Putting \( g(x) = x^r \) for \( x > 0 \) and real parameter \( r \) as well as \( f(x) = \ln_q(x) \) with \( q < 2 \) in Theorem 1 we obtain
\[
\sum_{i=1}^{n} a_i \ln_q \left( \frac{a_i^r}{b_i^r} \right) \geq \left( \sum_{i=1}^{n} a_i^r \right) \ln_q \left( \frac{\sum_{i=1}^{n} a_i^r}{\sum_{i=1}^{n} b_i^r} \right),
\]
where \( a_1, a_2, \ldots a_n \) and \( b_1, b_2, \ldots b_n \) are real numbers such that \( b_i^r > 0 \). Now applying the quotient rule of \( q \)-logarithm mentioned in equation (71) of appendix we find
\[
\ln_q \left( \frac{\sum_{i=1}^{n} a_i^r}{\sum_{i=1}^{n} b_i^r} \right) = \ln_q \left( \sum_{i=1}^{n} a_i^r \right) - \ln_q \left( \sum_{i=1}^{n} b_i^r \right)
\]
Combining we get
\[
\left( \sum_{i=1}^{n} b_i^r \right)^{1-q} \sum_{i=1}^{n} a_i \ln_q \left( \frac{a_i^r}{b_i^r} \right) \geq \left( \sum_{i=1}^{n} a_i \right) \left\{ \ln_q \left( \sum_{i=1}^{n} a_i \right) - \ln_q \left( \sum_{i=1}^{n} b_i \right) \right\}.
\]
For $r = 1$ the above inequality reduces to
\[
\left( \sum_{i=1}^{n} b_i \right)^{1-q} \sum_{i=1}^{n} a_i \ln_q \left( \frac{a_i}{b_i} \right) \geq \left( \sum_{i=1}^{n} a_i \right) \left\{ \ln_q \left( \sum_{i=1}^{n} a_i \right) - \ln_q \left( \sum_{i=1}^{n} b_i \right) \right\},
\]
for real numbers $a_1, a_2, \ldots, a_n$ and positive real numbers $b_1, b_2, \ldots, b_n$, as well as $q < 2$. Instead of $g(x) = x^r$ one may consider trigonometric, exponential, hyperbolic or any other functions to get new inequalities.

Recall that, a real valued function $f$ defined on a convex set $X$ is said to be concave if $-f(x)$ is convex. Applying it in equation (1) observe that for a concave function $f$ we have
\[
\sum_{i=1}^{n} t_i f(x_i) \leq f \left( \sum_{i=1}^{n} t_i x_i \right), \text{ where } \sum_{i=1}^{n} t_i = 1 \text{ and } 0 \leq t_i \leq 1.
\]
(16)

Considering $h(x) = xf(x)$ as a concave function in equation (7), we obtain
\[
\sum_{i=1}^{n} g(a_i) f \left( \frac{g(a_i)}{g(b_i)} \right) \leq \left( \sum_{i=1}^{n} g(a_i) \right) f \left( \frac{\sum_{i=1}^{n} g(a_i)}{\sum_{i=1}^{n} g(b_i)} \right),
\]
under the equivalent conditions on $a_i$ and $b_i$ as well as the real valued functions $f$ and $g$ as mentioned in theorem 1.

When $q > 2$ in equation (11) we observe that $h(x) = xf(x)$ is a concave function. Therefore the inequalities in equation (14) and (15) becomes
\[
\left( \sum_{i=1}^{n} b_i^{q} \right)^{1-q} \sum_{i=1}^{n} a_i^{q} \ln_q \left( \frac{a_i^{q}}{b_i^{q}} \right) \leq \left( \sum_{i=1}^{n} a_i^{q} \right) \left\{ \ln_q \left( \sum_{i=1}^{n} a_i^{q} \right) - \ln_q \left( \sum_{i=1}^{n} b_i^{q} \right) \right\},
\]
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 $-xf(x) = xf(\frac{1}{x})$ is a concave function. The equation (3) 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
\[
\sum_{i=1}^{n} a_i \log \left( \frac{b_i}{a_i} \right) \leq a \log \left( \frac{b}{a} \right).
\]
(20)
In general, $-xf(x) = xf(\frac{1}{x})$ does not hold. For instance, consider $f(x) = \ln_q(x)$, which refers $-x \ln_q(x) = x^q \ln_q \left( \frac{1}{x} \right)$. This fact leads us to a new inequality which we consider in the following theorem.

**Theorem 2.** Let $g$ be a real valued function whose domain contains the elements $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$, such that $g(a_i) > 0$ for all $i$. Let $f : [m_g, M_g] \to \mathbb{R}$ be a function for which $h(x) = xf(\frac{1}{x})$ is a concave function. 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,
\[
\sum_{i=1}^{n} g(a_i) f \left( \frac{g(b_i)}{g(a_i)} \right) = \sum_{i=1}^{n} g(b_i) \cdot g(a_i) \cdot f \left( \frac{g(b_i)}{g(a_i)} \right) = \sum_{i=1}^{n} \frac{g(b_i)}{b} h \left( \frac{g(a_i)}{g(b_i)} \right) \leq bh \left( \sum_{i=1}^{n} \frac{g(b_i)}{b} g(a_i) \right) \text{ (as } h(x) = xf \left( \frac{1}{x} \right) \text{ is concave)}
\]
\[
= bh \left( \frac{1}{b} \sum_{i=1}^{n} g(a_i) \right) = bh \left( \frac{a}{b} \right) = \frac{b a}{f} \left( \frac{b}{a} \right) = af \left( \frac{b}{a} \right)
\]
\[
\leq \left( \sum_{i=1}^{n} g(a_i) \right) f \left( \frac{\sum_{i=1}^{n} g(b_i) \cdot g(a_i)}{\sum_{i=1}^{n} g(a_i)} \right).
\]
\[
\square.
\]

We know that a twice-differentiable function \( h(x) \) is concave if and only if \( h''(x) < 0 \), which indicates
\[
\frac{d^2(xf(1/x))}{dx^2} = \frac{1}{x^3} f''(1/x) \leq 0.
\]
Note that, the equation (22) need not be equivalent to the equation (9). For example, consider the function \( f(x) = \frac{x}{x^2 + 2} \), where \( x > 0 \). We can calculate
\[
f'(x) = \frac{2 - x^2}{(x^2 + 2)^2}, \text{ and } f''(x) = \frac{2x(x^2 - 6)}{(x^2 + 2)^3}.
\]
Now, equation (22) suggests that
\[
\frac{1}{x^3} f''(1/x) = \frac{2 - 12x^2}{(2x^2 + 1)^3} \leq 0 \text{ when } x > \frac{1}{\sqrt{6}}.
\]
But, putting the values of \( f' \) and \( f'' \) in equation (9) we observe that
\[
x f''(x) + 2 f'(x) = \frac{8 - 12x^2}{(x^2 + 2)^3} < 0 \text{ when } x > \frac{\sqrt{2}}{3}.
\]
Now, considering \( f(x) = \frac{x}{x^2 + 2} \) and \( g(x) = x \) in theorem 2 we have
\[
\sum_{i=1}^{n} a_i^2 b_i \leq \frac{(\sum_{i=1}^{n} a_i)^2 (\sum_{i=1}^{n} b_i)}{2(\sum_{i=1}^{n} a_i)^2 + (\sum_{i=1}^{n} b_i)^2}.
\]
where \( a_i \) and \( b_i \) are greater than \( \sqrt{2}/3 \).

3 Generalized log-sum inequality for commuting matrices

We begin this section with a number of basic concepts of matrix analysis. A self-adjoint matrix \( A \) can be represented by \( A = U \Lambda(A) U^\dagger \), where \( U \) is a unitary matrix and \( \Lambda(A) \) is a diagonal matrix, such that \( \Lambda(A) = \text{diag}\{a_1, a_2, \ldots, a_n\} \), where \( a_i \) are the eigenvalues of \( A \). If two 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 \( B = U \Lambda(B) U^\dagger \), holds 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 self-adjoint matrix with eigenvalues in \( J \), then
\[
f(A) = U \text{ diag}\{f(a_i) : i = 1, 2, \ldots, n\} U^\dagger.
\]

**Theorem 3.** Let \( A \) and \( B \) be two commuting self-adjoint 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 \( i \). In addition, \( f : [m_\alpha, M_\beta] \to \mathbb{R} \) 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 $B$ be two commuting self-adjoint matrices there exists an unitary matrix $U$, such that,

$$g(A) = U \text{diag}\{g(a_i) : i = 1, 2, \ldots n\} U^\dagger$$
$$g(B) = U \text{diag}\{g(b_i) : i = 1, 2, \ldots n\} U^\dagger$$

$$(g(B))^{-1} = U \left\{ \frac{1}{g(b_i)} : i = 1, 2, \ldots n \right\} U^\dagger, \quad (28)$$

as $g(b_i) > 0$ for all $i$. Therefore,

$$f(g(A)(g(B))^{-1}) = U \text{diag}\left\{ f\left(\frac{g(a_i)}{g(b_i)}\right) : i = 1, 2, \ldots n \right\} U^\dagger,$$

which implies

$$g(A)f(g(A)(g(B))^{-1}) = U \text{diag}\left\{ g(a_i)f\left(\frac{g(a_i)}{g(b_i)}\right) : i = 1, 2, \ldots n \right\} U^\dagger. \quad (29)$$

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, we have

$$\text{trace}(g(A)f(g(A)(g(B))^{-1})) = \sum_{i=1}^{n} g(a_i)f\left(\frac{g(a_i)}{g(b_i)}\right). \quad (30)$$

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 1 we obtain

$$\sum_{i=1}^{n} g(a_i)f\left(\frac{g(a_i)}{g(b_i)}\right) = \left(\sum_{i=1}^{n} g(a_i)\right) f\left(\frac{\sum_{i=1}^{n} g(a_i)}{\sum_{i=1}^{n} g(b_i)}\right) = \text{trace}(g(A)) f\left(\frac{\text{trace}(g(A))}{\text{trace}(g(B))}\right). \quad (31)$$

Combining we get the result. \qed

The following corollary holds trivially from the above theorem.

**Corollary 1.** Given two positive definite, commuting matrices $A$ and $B$ we have

$$\text{trace}(\exp(A \log(A))) - \text{trace}(\exp(A \log(B))) \geq \text{trace}(A) \log\left(\frac{\text{trace}(A)}{\text{trace}(B)}\right).$$

**Proof.** Consider $f(x) = \log(x)$ and $g(x) = x$ in Theorem 3. As $A$ and $B$ are positive definite we have $a_i > 0$ and $b_i > 0$ for all $i$. Note that,

$$\sum_{i=1}^{n} a_i \log\left(\frac{a_i}{b_i}\right) = \sum_{i=1}^{n} \log(a_i^{a_i}) - \sum_{i=1}^{n} \log(b_i^{b_i}) = \text{trace}(\exp(A \log(A))) - \text{trace}(\exp(A \log(B))). \quad (32)$$

Also, applying $\sum_{i=1}^{n} a_i = \text{trace}(A)$ and $\sum_{i=1}^{n} b_i = \text{trace}(B)$ we observe that $\text{trace}(g(A)) f\left(\frac{\text{trace}(g(A))}{\text{trace}(g(B))}\right) = \text{trace}(A) \log\left(\frac{\text{trace}(A)}{\text{trace}(B)}\right)$. Combining we get the result. \qed

Theorem 3 may be considered as a counterpart of Theorem 1 for commuting self-adjoint matrices. Immediately, we find the matrix counterparts of a number of inequalities which we have derived at the last section. This matrix inequalities have immediate consequences in quantum information theory.

Recall that, in quantum information theory [6] 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

$$D(\rho||\sigma) = \text{trace}[\rho (\log(\rho) - \log(\sigma))].$$

**Example 3.** The matrix counterpart of equation (10) is represented as

$$\text{trace}[A^r \log(A^r B^{-r})] \geq \text{trace}(A^r) [\log(\text{trace}(A^r)) - \log(\text{trace}(B^r))], \quad (33)$$

where $A$ and $B$ are positive definite matrices.
where $A$ and $B$ are self-adjoint commuting matrices as well as $B$ is positive definite. If $r = 1$ the above inequality takes the following form:

$$\text{trace} \left[ A \log(AB^{-1}) \right] \geq \text{trace}(A) [\log(\text{trace}(A)) - \log(\text{trace}(B))].$$  \hspace{1cm} (34)

Considering commutativity of $A$ and $B$, that is $AB = BA$, we can prove that $\log(AB^{-1}) = \log(A) - \log(B)$. Replacing it at the left hand side we find

$$\text{trace} \left[ A(\log(A) - \log(B)) \right] \geq \text{trace}(A) [\log(\text{trace}(A)) - \log(\text{trace}(B))].$$  \hspace{1cm} (35)

This inequality is already derived in [7, 8, Theorem 3.3]. In addition, considering $\text{trace}(A) = \text{trace}(B) = 1$, we find that $A$ and $B$ are two density matrices. Then, the quantum relative entropy $D(A||B) = \text{trace}[A(\log(A) - \log(B))] \geq 0$.

Considering $B = I$, the identity matrix of order $n$, in equation (33) we observe that

$$\text{trace}[A^r \log A^r] \geq \text{trace}(A^{r'}) [\log(\text{trace}(A^{r'}))] - (n).$$  \hspace{1cm} (36)

Putting $r = 1$ in the above inequality we have $\text{trace}[A \log A] \geq \text{trace}(A) [\log(\text{trace}(A)) - \log(n)]$. In addition, if $\text{trace}(A) = 1$ we have $\text{trace}[A \log A] \geq -\log(n)$. Now, $A$ is a density matrix. The von-Neumann entropy of $A$ is $-\text{trace}[A \log A] \leq \log(n)$, which is the maximum of von-Neumann entropy.

**Example 4.** The matrix counterpart of equation (14) will be given by

$$(\text{trace}(B^r))^{1-q} \text{trace} \left[ A^r \ln_q (A^r B^{-r}) \right] \geq (\text{trace}(A^r)) \ln_q (\text{trace}(A^r)) - \ln_q (\text{trace}(B^r)), \hspace{1cm} (37)$$

where $A$ and $B$ are self-adjoint commuting matrices as well as $B$ is positive definite. For $r = 1$ we have

$$(\text{trace}(B))^{1-q} \text{trace} \left[ A \ln_q (AB^{-1}) \right] \geq (\text{trace}(A)) \ln_q (\text{trace}(A)) - \ln_q (\text{trace}(B)). \hspace{1cm} (38)$$

Note that in the above two inequalities we have $q < 2$. We have already mentioned that for $q > 2$ equation (14) is not valid. The appropriate inequality is given by equation (18). Its matrix counterpart is given by

$$(\text{trace}(B^r))^{1-q} \text{trace} \left[ A^r \ln_q (A^r B^{-r}) \right] \leq (\text{trace}(A^r)) \ln_q (\text{trace}(A^r)) - \ln_q (\text{trace}(B^r)), \hspace{1cm} (39)$$

for self-adjoint commuting matrices $A$ and $B$ as well as positive definite matrix $B$.

Theorem 2 can also generates a trace form inequality, which we mention below without a proof.

**Theorem 4.** Let $A$ and $B$ be two commuting self-adjoint 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 $g : \Lambda(A) \cup \Lambda(B) \rightarrow \mathbb{R}$ be a function, such that, $g(b_i) > 0$ for all $i$. In addition, $f : [m_g, M_g] \rightarrow \mathbb{R}$ is a function which $h(x) = xf \left( \frac{1}{x} \right)$ is a concave function. 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 Generalized inequalities with Löwner partial order**

Recall that given self-adjoint 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. It is induced in the real space of Hermitian matrices by the cone of positive semidefinite matrices. If $A$ is positive definite we write $A > 0$.

We have already mentioned in equation (27) that a real valued function $f$ defined on an interval $J$ can also be defined for a self-adjoint matrix whose eigenvalues are in $J$. Now, a real-valued continuous function $f(t)$ is said to be operator monotone if $A \leq B$ implies $f(A) \leq f(B)$. A continuous function $f : J \rightarrow \mathbb{R}$ defined on an interval $J$ is said to be an operator convex function if

$$f(\lambda A + (1 - \lambda)B) \leq \lambda f(A) + (1 - \lambda) f(B), \hspace{1cm} (40)$$

for all $\lambda \in [0, 1]$ and every pair of self-adjoint operators $A$, and $B$ on an infinite dimensional Hilbert space $\mathcal{H}$ with spectra in $\mathcal{I}$. A function $f$ is operator concave if $-f$ is operator convex. The function $f$ is called matrix convex of order $n$ if the dimension of $\mathcal{H}$ is $n$. 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 $1 \leq r \leq 2$ [9, 10].

The operator Jensen inequality [11, 12] plays a crucial role for further development of this article. Below we mention it:
Lemma 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. \(f(A^1XA) \leq A^1f(X)A\) for all \(A\) with \(|A| \leq 1\) and every self-adjoint \(X\) with spectrum in \([0, \alpha]\).
3. \(f(A^1XA + B^1YB) \leq A^1f(X)A + B^1f(Y)B\) for all \(A, B\) with \(A^1A + B^1B \leq I\) and all \(X, Y\) with spectrum in \([0, \alpha]\).
4. \(f(PXP) \leq Pf(X)P\) for every projection \(P\) and every self-adjoint \(X\) with spectrum in \([0, \alpha]\).

We utilize Dirac bra-ket notation \([14]\) for simplifying the notations in its proof. Recall that \(|u\rangle\) denotes a column vector or 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 self-adjoint matrix of order \(n\). Now the spectral decomposition of any self-adjoint 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 \(\lambda_i\).

Theorem 5. Let \(A_1, A_2, \ldots, A_m\) and \(B_1, B_2, \ldots, B_m\) be a 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

\[
\sum_{i=1}^m A_i^{1/2} f\left(A_i^{-1/2} B_i A_i^{-1/2}\right) A_i^{1/2} \leq A^{1/2} f\left(\sum_{i=1}^m A_i^{-1/2} B_i A_i^{-1/2}\right) A_i^{1/2}.
\]

Proof. As \(B_i\) is a self-adjoint operator, 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

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

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

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

Combining we get

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

\[
= \sum_{k=1}^n f\left(\lambda_k^{(i)}\right) |w_k^{(i)}\rangle \langle w_k^{(i)}| = f\left(B_i\right).
\]

Summing over \(i\) we find

\[
\sum_{i=1}^m A_i^{1/2} f\left(A_i^{-1/2} B_i A_i^{-1/2}\right) A_i^{1/2} = \sum_{i=1}^m f\left(B_i\right).
\]

If \(A \geq mI\) we have \(A^2 \geq m^2I\), that is \(\sum_{i=1}^m A_i^{-1/2} A_i^{-1/2} \leq I\). Now, applying Lemma 1 we find that

\[
A^{-1/2} \sum_{i=1}^m f\left(B_i\right) A^{-1/2} = \sum_{i=1}^m A^{-1/2} f\left(B_i\right) A^{-1/2} \leq f\left(\sum_{i=1}^m A_i^{-1/2} B_i A_i^{-1/2}\right)
\]

\[
= f\left(A^{-1/2} \sum_{i=1}^m B_i A^{-1/2}\right) = f\left(A^{-1/2} B A^{-1/2}\right).
\]
Multiplying $A^{\frac{1}{2}}$ in both side of the above inequality we get
\[
\sum_{i=1}^{m} f(B_i) \leq A^{\frac{1}{2}} f \left( (A^{\frac{-1}{2}})^{\dagger} BA^{\frac{-1}{2}} \right) A^{\frac{1}{2}}.
\] (47)

Putting the value of $f(B_i)$ from equation (44) we find the result.

\textbf{Remark 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
\[
\sum_{i=1}^{m} \frac{a_i}{b_i} \left( a_i^\frac{1}{2} b_i a_i^{-\frac{1}{2}} \right) \leq \frac{a}{b} f \left( a^\frac{-1}{2} b a^{-\frac{1}{2}} \right) a^\frac{1}{2},
\] (48)
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
\[
\sum_{i=1}^{m} \frac{a_i}{b_i} \left( \frac{b_i}{a_i} \right) \leq a f \left( \frac{b}{a} \right).
\] (49)

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

\textbf{Corollary 2.} Let $A_1, A_2, \ldots, A_m$ and $B_1, B_2, \ldots, B_m$ be two sets of positive definite self-adjoint 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
\[
\sum_{i=1}^{m} A_i^{\frac{1}{2}} f (A_i^{\frac{1}{2}} B_i A_i^{\frac{1}{2}}) A_i^{\frac{1}{2}} \leq 0.
\]

\textbf{Proof.} The proof follows trivially from Theorem 5. □

\textbf{Remark 2.} The operator Shannon inequality was given in [13]
\[
\sum_{i=1}^{m} A_i^{\frac{1}{2}} \log \left( A_i^{\frac{1}{2}} B_i A_i^{\frac{1}{2}} \right) A_i^{\frac{1}{2}} \leq 0
\]
under the assumption $\sum_{i=1}^{m} A_i = \sum_{i=1}^{n} B_i = I$. Our condition in Corollary 2 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 $A \geq mI$ is satisfied in Theorem 5. However, we have not obtained a proper result for contractive condition such as $A_i \leq I$. Closing this section, we give a result which does not impose an additional condition for the matrices $A_i$. We need the following known facts for proving the next theorem:

\textbf{Lemma 2.} ([15]) Let $X$ and $A$ be bounded linear operators on a Hilbert space $\mathcal{H}$. Suppose that $X \geq 0$ and $\|A\| \leq 1$. If $f$ is an operator monotone function defined on $[0, \infty)$, then $A^f(X)A \leq f(A^fXA)$.

\textbf{Lemma 3.} ([9, p.14]) For any square matrix $X_i$ and positive definite matrices $A_i$ we have
\[
\sum_{i=1}^{m} X_i^{\dagger} A_i^{-1} X_i \geq \left( \sum_{i=1}^{m} X_i \right)^{\dagger} \left( \sum_{i=1}^{m} A_i \right)^{-1} \left( \sum_{i=1}^{m} X_i \right).
\] (50)

\textbf{Theorem 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, $f$ is an operator monotone function defined on $[0, \infty)$. Then we have
\[
\frac{1}{m} \left( \sum_{i=1}^{m} B_i \right)^{\dagger} \sum_{i=1}^{m} f \left( B_i^{\frac{1}{2}} A_i^{-1} B_i^{\frac{1}{2}} \right) \sum_{i=1}^{m} B_i \right)^{\frac{1}{2}} \leq B^{\frac{1}{2}} \left[ f \left( B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}} \right) \right]^{-1} B^{\frac{1}{2}},
\] (51)
and
\[
\sum_{i=1}^{m} A_i^{\frac{1}{2}} B_i \left[ f \left( B_i^{\frac{1}{2}} A_i^{-1} B_i^{\frac{1}{2}} \right) \right]^{-1} B_i^{\frac{1}{2}} A_i^{\frac{1}{2}} \leq \frac{1}{m} \left( \sum_{i=1}^{m} A_i \right)^{\dagger} B^{\frac{1}{2}} \left[ f \left( B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}} \right) \right]^{-1} B^{\frac{1}{2}} \left( \sum_{i=1}^{m} A_i \right)^{\frac{1}{2}}.
\] (52)
Proof. We can write

\[ B_t^\frac{1}{2} A_i^{-1} B_t^\frac{1}{2} = B_t^\frac{1}{2} \left( B_t^\frac{-1}{2} B_t^\frac{1}{2} \right) A_i^{-1} \left( B_t^\frac{-1}{2} B_t^\frac{1}{2} \right) B_t^\frac{1}{2} = B_t^\frac{1}{2} B_t^\frac{-1}{2} \left( B_t^\frac{-1}{2} A_i^{-1} B_t^\frac{1}{2} \right) B_t^\frac{-1}{2} B_t^\frac{1}{2}. \]  

(53)

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

\[ B_t^\frac{1}{2} A_i^{-1} B_t^\frac{1}{2} \geq B_t^\frac{1}{2} B_t^\frac{-1}{2} \left( B_t^\frac{-1}{2} A^{-1} B_t^\frac{1}{2} \right) B_t^\frac{-1}{2} B_t^\frac{1}{2}. \]  

(54)

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

\[ f \left( B_t^\frac{1}{2} A_i^{-1} B_t^\frac{1}{2} \right) \geq f \left( B_t^\frac{1}{2} B_t^\frac{-1}{2} \left( B_t^\frac{-1}{2} A^{-1} B_t^\frac{1}{2} \right) B_t^\frac{-1}{2} B_t^\frac{1}{2} \right). \]  

(55)

From \( B > B_i \) for all \( i \), we can see \( I > B^{-1/2} B_i B^{-1/2} = \left( B_i^{-1/2} B^{-1/2} \right)^\dagger B_i^{1/2} B^{-1/2} \) which implies \( 1 > \| \left( B_i^{1/2} B^{-1/2} \right)^\dagger B_i^{-1/2} B^{-1/2} \| = \| B_i^{-1/2} B^{-1/2} \|^2 \), since \( \| A \| = \| A^\dagger A \|^1/2 \) for every operator \( A \) in general. Thus we have \( \| B_i^{1/2} B^{-1/2} \| < 1 \). We also have \( \| B_t^\frac{-1}{2} B_t^\frac{1}{2} \| \leq 1 \) so that we have the following inequality by Lemma 2,

\[ f \left( B_t^\frac{1}{2} B_t^\frac{-1}{2} \left( B_t^\frac{-1}{2} A^{-1} B_t^\frac{1}{2} \right) B_t^\frac{-1}{2} B_t^\frac{1}{2} \right) \geq \left( B_t^\frac{1}{2} B_t^\frac{-1}{2} \right) f \left( B_t^\frac{1}{2} A^{-1} B_t^\frac{1}{2} \right) \left( B_t^\frac{1}{2} B_t^\frac{-1}{2} \right). \]

That is,

\[ f \left( B_t^\frac{1}{2} B_t^\frac{-1}{2} \right) \geq \left( B_t^\frac{1}{2} B_t^\frac{-1}{2} \right) f \left( B_t^\frac{1}{2} A^{-1} B_t^\frac{1}{2} \right) \left( B_t^\frac{1}{2} B_t^\frac{-1}{2} \right). \]  

(56)

Multiplying \( B_t^\frac{-1}{2} \) to the both sides, we have

\[ B_t^\frac{-1}{2} f \left( B_t^\frac{1}{2} A_i^{-1} B_t^\frac{1}{2} \right) B_t^\frac{-1}{2} \geq B_t^\frac{-1}{2} f \left( B_t^\frac{1}{2} A^{-1} B_t^\frac{1}{2} \right) B_t^\frac{-1}{2}. \]  

(57)

Taking an inverse of the both sides, we have

\[ B_t^\frac{1}{2} \left[ f \left( B_t^\frac{1}{2} A_i^{-1} B_t^\frac{1}{2} \right) \right]^{-1} B_t^\frac{1}{2} \leq B_t^\frac{1}{2} \left[ f \left( B_t^\frac{1}{2} A^{-1} B_t^\frac{1}{2} \right) \right]^{-1} B_t^\frac{1}{2}. \]  

(58)

Thus we have

\[ \frac{1}{m} \sum_{i=1}^{m} B_t^\frac{1}{2} \left[ f \left( B_t^\frac{1}{2} A_i^{-1} B_t^\frac{1}{2} \right) \right]^{-1} B_t^\frac{1}{2} \leq B_t^\frac{1}{2} \left[ f \left( B_t^\frac{1}{2} A_i^{-1} B_t^\frac{1}{2} \right) \right]^{-1} B_t^\frac{1}{2}. \]  

(59)

Applying Lemma 3 to the above, we have

\[ \frac{1}{m} \sum_{i=1}^{m} B_t^\frac{1}{2} \left[ f \left( B_t^\frac{1}{2} A_i^{-1} B_t^\frac{1}{2} \right) \right]^{-1} B_t^\frac{1}{2} \geq \frac{1}{m} \left( \sum_{i=1}^{m} B_t^\frac{1}{2} \right) \left[ \sum_{i=1}^{m} f \left( B_t^\frac{1}{2} A_i^{-1} B_t^\frac{1}{2} \right) \right]^{-1} \left( \sum_{i=1}^{m} B_t^\frac{1}{2} \right). \]  

(60)

Combining (59) and (60), we get the inequality (51).

To prove the inequality (52), we start from (58). By multiplying \( A_i^\frac{1}{2} \) to the both sides in (58), we have

\[ A_i^\frac{1}{2} B_t^\frac{1}{2} \left[ f \left( B_t^\frac{1}{2} A_i^{-1} B_t^\frac{1}{2} \right) \right]^{-1} B_t^\frac{1}{2} A_i^\frac{1}{2} \leq A_i^\frac{1}{2} B_t^\frac{1}{2} \left[ f \left( B_t^\frac{1}{2} A^{-1} B_t^\frac{1}{2} \right) \right]^{-1} B_t^\frac{1}{2} A_i^\frac{1}{2}. \]

Taking a summation on \( i \) from 1 to \( m \) for the both sides in the above inequality, we obtain

\[ \sum_{i=1}^{m} A_i^\frac{1}{2} B_t^\frac{1}{2} \left[ f \left( B_t^\frac{1}{2} A_i^{-1} B_t^\frac{1}{2} \right) \right]^{-1} B_t^\frac{1}{2} A_i^\frac{1}{2} \leq \sum_{i=1}^{m} A_i^\frac{1}{2} B_t^\frac{1}{2} \left[ f \left( B_t^\frac{1}{2} A^{-1} B_t^\frac{1}{2} \right) \right]^{-1} B_t^\frac{1}{2} A_i^\frac{1}{2}. \]

That is, we have

\[ \sum_{i=1}^{m} A_i^\frac{1}{2} B_t^\frac{1}{2} \left[ (-1) f \left( B_t^\frac{1}{2} A_i^{-1} B_t^\frac{1}{2} \right) \right]^{-1} B_t^\frac{1}{2} A_i^\frac{1}{2} \geq \sum_{i=1}^{m} A_i^\frac{1}{2} B_t^\frac{1}{2} \left[ (-1) f \left( B_t^\frac{1}{2} A^{-1} B_t^\frac{1}{2} \right) \right]^{-1} B_t^\frac{1}{2} A_i^\frac{1}{2}. \]  

(61)
Applying Lemma 3 to the right hand side in (61), we have
\[
\sum_{i=1}^{m} A_i^\frac{1}{2} B_i^{\frac{1}{2}} \left[ (-1) f \left( B_i^{\frac{1}{2}} A_i^{-1} B_i^{\frac{1}{2}} \right) \right]^{-1} B_i^{\frac{1}{2}} A_i^{\frac{1}{2}} \geq \left( \sum_{i=1}^{m} A_i^\frac{1}{2} \right) \left( (-m) B^{\frac{1}{2}} f \left( B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}} \right) B^{\frac{1}{2}} \right)^{-1} \left( \sum_{i=1}^{m} A_i^\frac{1}{2} \right),
\]
that is,
\[
- \sum_{i=1}^{m} A_i^\frac{1}{2} B_i^{\frac{1}{2}} \left[ f \left( B_i^{\frac{1}{2}} A_i^{-1} B_i^{\frac{1}{2}} \right) \right]^{-1} B_i^{\frac{1}{2}} A_i^{\frac{1}{2}} \geq (-1) \left( \sum_{i=1}^{m} A_i^\frac{1}{2} \right) \frac{1}{m} B^{\frac{1}{2}} \left[ f \left( B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}} \right) \right]^{-1} B^{\frac{1}{2}} \left( \sum_{i=1}^{m} A_i^\frac{1}{2} \right),
\]
which implies
\[
\sum_{i=1}^{m} A_i^\frac{1}{2} B_i^{\frac{1}{2}} \left[ f \left( B_i^{\frac{1}{2}} A_i^{-1} B_i^{\frac{1}{2}} \right) \right]^{-1} B_i^{\frac{1}{2}} A_i^{\frac{1}{2}} \leq \frac{1}{m} \left( \sum_{i=1}^{m} A_i^\frac{1}{2} \right) B^{\frac{1}{2}} \left[ f \left( B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}} \right) \right]^{-1} B^{\frac{1}{2}} \left( \sum_{i=1}^{m} A_i^\frac{1}{2} \right), \tag{62}
\]
Combining (61) and (62), we get the inequality (52).

**Remark 3.** Considering \(a_i\) and \(b_i\) are positive real numbers and replacing \(A_i \equiv a_i\) and \(B_i \equiv b_i\) as well as \(f\) is a monotone increasing function we have from equation (51)
\[
\frac{1}{m} \left( \sum_{i=1}^{m} b_i \right)^2 \left[ \sum_{i=1}^{m} f \left( \frac{b_i}{a_i} \right) \right]^{-1} \leq b \left[ f \left( \frac{b}{a} \right) \right]^{-1}, \tag{63}
\]
Also, from equation (52) we can write
\[
\sum_{i=1}^{m} a_i b_i \left[ f \left( \frac{b_i}{a_i} \right) \right]^{-1} \leq \frac{1}{m} \left( \sum_{i=1}^{m} a_i \right)^2 b \left[ f \left( \frac{b}{a} \right) \right]^{-1}. \tag{64}
\]

Both inequalities in (63) and (64) do not have same form of log-sum inequality (49). Therefore we have to conclude that the restricted condition (so-called expansivity of matrices \(A_i\)) given in Theorem 5 leads us to obtain the generalized log-sum inequality for non-commutative matrices. However, we could not obtain such an inequality without restricted condition in Theorem 6. In the further studies on this topic, we would like to obtain the log-sum type inequality for non-commutative \textit{contractive} matrices.

5 Conclusion

The log-sum inequality which is mentioned in equation (3) plays a crucial role in classical and quantum information theory. In this article, we present a number of inequalities which can be considered as its generalization. Inequality (3) depends on the convexity of the function \(x \log(x)\) which was relaxed utilizing the function \(x f(x)\), in [4]. We illustrate that the concavity of \(x f(\frac{1}{x})\) is also useful for deriving similar inequalities. We discuss these generalizations 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 are also discussed. These results can be generalized for the commuting self-adjoint matrices as trace-form inequalities. Later we discuss the generalized log-sum inequality for the operator monotone and convex functions in the context of Löwner partial order relation of the positive semi-definite matrices. These generalizations are the applications of Hansen operator inequality.

Acknowledgments

SD is thankful to IIT Kharagpur for a post-doctoral fellowship during this work. SF was partially supported by JSPS KAKENHI Grant Number 16K05257.
Appendix: Properties of deformed logarithms

The equation (11) defines the $q$-deformed logarithm. Below, we mention a number of its properties which can be established easily:

1. $\ln_q \left( \frac{x}{x} \right) = \ln_q(1) = 0$.  
(65)

2. As $\ln_q(y) = \frac{x^{1-q} - 1}{1-q}$ we have $y^{1-q} = 1 + (1 - q) \ln_q(y)$.

3. The product rules of logarithm:

$$\ln_q(xy) = \ln_q(x) + \ln_q(y) + (1 - q) \ln_q(x) \ln_q(y).$$  
(66)

$$\ln_q(xy) = x^{1-q} \ln_q(y) + \ln_q(x).$$  
(67)

4. Putting $y \equiv \frac{1}{y}$ in the expression of the product rule we find

$$0 = \ln_q \left( \frac{y}{y} \right) = \ln_q(1) + \ln_q \left( \frac{1}{y} \right) + (1 - q) \ln_q(y) \ln_q \left( \frac{1}{y} \right).$$  
(68)
Simplifying, we get
\[ \ln_q \left( \frac{1}{y} \right) = -\frac{\ln_q(y)}{1 + (1 - q) \ln_q(y)} = -\frac{\ln_q(y)}{y^{1-q}}. \] (69)

Using another expression of the product rule we find
\[ \ln_q \left( \frac{y}{y} \right) = y^{1-q} \ln_q(y) + \ln_q(y) = 0, \]
we thus have,
\[ \ln_q \left( \frac{1}{y} \right) = -\frac{1}{y^{1-q}} \ln_q(y). \] (70)

5. Therefore the division rule of the logarithm can be expressed as
\[
\ln_q \left( \frac{x}{y} \right) = \ln_q(x) + \ln_q \left( \frac{1}{y} \right) + (1 - q) \ln_q(x) \ln_q \left( \frac{1}{y} \right)
= \frac{\ln_q(x) - \ln_q(y)}{1 + (1 - q) \ln_q(y)} = \frac{\ln_q(x) - \ln_q(y)}{y^{1-q}},
\] (71)
which is the quotient rule for deformed logarithm.