An Identity for Expectations and Characteristic Function of Matrix Variate Skew-normal Distribution with Applications to Associated Stochastic Orderings

Tong Pu · Narayanaswamy Balakrishnan · Chuancun Yin

Received: 29 March 2021 / Revised: 27 May 2021 / Accepted: 6 October 2021 /
Published online: 27 September 2022
© School of Mathematical Sciences, University of Science and Technology of China and Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
We establish an identity for $Ef(Y) - Ef(X)$, when $X$ and $Y$ both have matrix variate skew-normal distributions and the function $f$ satisfies some weak conditions. The characteristic function of matrix variate skew normal distribution is then derived. We then make use of it to derive some necessary and sufficient conditions for the comparison of matrix variate skew normal distributions under six different orders, such as usual stochastic order, convex order, increasing convex order, upper orthant order, directionally convex order and supermodular order.

Keywords Characteristic function · Integral order · Matrix variate skew-normal distributions · Stochastic comparisons

Mathematics Subject Classification 60E10 · 60E15

1 Introduction and Motivation

Stochastic orders, which are partial orders on a set of random variables, provide methods to describe intuitively random variables being larger, riskier or more dependent.
Stochastic orders of variables have found key applications in such diverse fields as actuarial science, economics, comparison of experiments, reliability analysis and queueing theory. Different kinds of stochastic orders possess different properties, characterizations and applications, and interested readers may refer to the books by Denuit et al. [19], Müller and Stoyan [35] and Shaked and Shanthikumar [39] for elaborate details.

Given two random variables \( X \) and \( Y \) and a class of measurable functions \( F \), integral stochastic orders seek the order between \( X \) and \( Y \) by comparing \( Ef(Y) \) and \( Ef(X) \), where \( f \in F \). Integral stochastic orders include the most common stochastic orders like usual stochastic order and convex order. Müller [32], Müller and Stoyan [35] and Shaked and Shanthikumar [39] provide some important properties on various stochastic orders and their applications to various practical problems. Belzunce [13] showed relationships among some integral stochastic orders. Joe [28] studied concordance orders, which are special cases of integral stochastic orders, to obtain measures of multivariate concordance. Müller [33] provided a general treatment on integral stochastic orders, with the main tool being an identity for \( Ef(Y) - Ef(X) \), when \( X \) and \( Y \) are multivariate random variables. This identity was derived by using Fourier Inversion Theorem and then by employing integration by parts. Subsequently, necessary and sufficient conditions for some integral stochastic orders for multivariate normal distributions were obtained by Müller [33]. Ding and Zhang [20] extended the results in Müller [33] to Kotz-type distributions which form a special class of elliptical symmetric distributions. Landsman and Tsanakas [30] studied conditions under which bivariate elliptical distributions are ordered through the convex, increasing convex and concordance orderings. Davidov and Peddada [17] showed for elliptically distributed random vectors that the positive linear usual stochastic order coincides with the multivariate usual stochastic order. Recently, Pan et al. [37] studied convex and increasing convex orderings of multivariate elliptical random vectors. Yin [43] and Ansari and Rüschendorf [4] subsequently derived some other integral stochastic orderings of multivariate elliptical distributions.

The univariate skew-normal distribution was developed by Azzalini [6], which was subsequently extended to the multivariate case by Azzalini and Dalla Valle [9]. This distribution presents a mathematically tractable extension of the multivariate normal distribution and accommodates skewness in the model. The theory of skew-normal distribution provides a useful account of how to describe or fit asymmetric data. Matrix variate distributions have found key applications in such fields as analysis of multivariate longitudinal data (Anderlucci and Viroli [3]), time series (Chen et al. [15]) and actuarial science. As an example, consider \( p \) blood tests on some patients, where \( n \) variables are measured in each test. In this situation, a random matrix \( X \) of size \( n \times p \) is then observed for each patient. The introduction part of Caro-Lopera et al. [14] presents several detailed motivations for the study of some random matrices, and we refer the interested readers to this paper. Chen and Gupta [16] introduced matrix variate skew-normal distribution and discussed properties of this distribution. Domínguez-Molina et al. [21] introduced the matrix variate closed skew-normal distribution and presented two constructions for this distribution. Ning and Gupta [36] introduced the matrix variate extended skew-normal distribution and Rezaei et al. [38] introduced scale and shape mixture of matrix variate extended skew-normal distributions based on the result of Ning and Gupta [36]. The matrix variate extension of the skew-normal distribution was
presented in two ways, one by Chen and Gupta [16], and the other as a natural extension of the multivariate skew-normal distribution of Azzalini [9]. The matrix variate skew-normal distribution used in this paper is based on the latter. These two different definitions have been compared by Harrar and Gupta [24]. All these works primarily defined the distribution and derived the means, covariances, moments, moment generating functions and some other properties. In recent years, theoretical and practical works on skew-normal distribution have increased considerably. The characteristic functions of univariate and multivariate skew-normal distributions were derived by Kim and Genton [29]. Stochastic orderings of the univariate skew-normal distribution and the general skew-symmetric family of distributions were discussed by Azzalini and Regoli [10]. The characterizations of likelihood ratio order and usual stochastic order of the univariate skew-symmetric distribution were discussed by Hürlimann [25]. Recently, Jamali et al. [26] derived some integral stochastic orderings for multivariate skew-normal distribution, while Abdi et al. [1] studied properties and inferential methods for mean-mixtures of multivariate normal distributions. Jamali et al. [27] studied some integral stochastic orderings of mean-variance mixture of normal distributions and scale-shape mixture of skew distributions. Amiri et al. [2] studied some linear stochastic orderings in the scale mixtures of the multivariate skew-normal family of distributions and showed these linear orderings to be equivalent to multivariate orders.

Our work here follows those of Müller [32] and Jamali et al. [26]. Specifically, we extend the results of Jamali et al. [26] for the multivariate skew-normal distribution to the matrix variate case. Necessary and sufficient conditions are derived for some important integral stochastic orders for matrix variate skew-normal distributions. The main tool used is an identity for $E f (Y) - E f (X)$, when $X$ and $Y$ are both skew-normally distributed matrices. We need the characteristic function of skew-normally distributed matrices for establishing this identity. For this purpose, we also extend the result of Kim and Genton [29] on the characteristic function of multivariate skew-normal to the matrix variate case. This study makes a significant contribution to research on matrix variate skew-normal distribution by deriving its characteristic function and some conditions of various stochastic orderings.

The rest of this paper is organized as follows. In Sect. 2, we review skew-normal distributions in multivariate and matrix variate cases and state some of their key properties. We also present a brief review of integral stochastic orderings. In Sect. 3, we derive the characteristic function of the matrix variate skew-normal distribution. In Sect. 4, we establish an identity for $E f (Y) - E f (X)$, when $X$ and $Y$ are both matrix variate skew-normally distributed random variables. In Sect. 5, the ordering results for matrix variate skew-normal distribution, based on the results of Sect. 4, are presented. Finally, Sect. 6 presents some concluding remarks.

2 Preliminaries

The following notation will be used throughout this paper. We will use lowercase letters, bold lowercase letters and bold capital letters to denote numbers, vectors and matrices, respectively; $\Phi (\cdot)$ and $\phi (\cdot)$ to denote the cumulative distribution function and probability density function of the univariate standard normal distribution, respec-
tively; and $\Phi_n (\cdot; \mu, \Sigma)$ and $\phi_n (\cdot; \mu, \Sigma)$ to denote the cumulative distribution function and probability density function of the multivariate $n$-dimensional normal distribution with mean vector $\mu$ and covariance matrix $\Sigma$, $N_n (\mu, \Sigma)$.

For any $B \in \mathbb{R}^{n \times p}$, $B^T$ denotes the transpose of $B$. If $B$ is a square matrix, then $B^{-1}$ denotes the inverse of $B$. For $B = (b_1, b_2, \ldots, b_p)$, we will use $\text{vec} (B) = \left( b_1^T, b_2^T, \ldots, b_p^T \right)^T$ to denote the matrix vectorization, and $\text{tr} (B)$ and $\text{etr} (B)$ to denote the trace and exponential trace of matrix $B$, respectively. For $A \in \mathbb{R}^{l \times m}$, $B \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{n \times p}$ and $D \in \mathbb{R}^{p \times q}$, we will use $B \otimes C$ to denote the Kronecker product of $B$ and $C$. It is known that $\text{vec} (BCD) = (D^T \otimes B) \text{vec} (C)$, $(A \otimes B)^T = A^T \otimes B^T$, and $(A \otimes C)(B \otimes D) = (AB \otimes CD)$; see Magnus and Neudecker [31]. Throughout this paper, the inequality between vectors or matrices denotes componentwise inequalities.

### 2.1 Skew-Normal Distributions

We first recall the definition of multivariate skew-normal distribution given by Azzalini and Capitanio [7]. Consider a full rank $n \times n$ covariance matrix $\Omega = (\omega_{ij})$, and let $\overline{\Omega} = \omega^{-1} \Omega \omega^{-1}$ be the associated correlation matrix, where $\omega = \text{diag} \{ \sqrt{\omega_{11}}, \ldots, \sqrt{\omega_{nn}} \}$.

**Definition 2.1** (Azzalini and Capitanio [7]) An $n$-dimensional random vector $Z$ is said to have a multivariate skew-normal distribution with location parameter $\mu$, scale parameter $\Omega$ and skewness parameter $\alpha$, denoted by $SN_n (\mu, \Omega, \alpha)$, if its probability density function is

\[
f_Z (z) = 2\phi_n (z; \mu, \Omega) \Phi \left( \alpha^T \omega^{-1} (z - \mu) \right), \quad z \in \mathbb{R}^n. \tag{2.1}\]

The characteristic function of an $n$-dimensional random vector $X$ is given by $\Psi_X (t) = E \left[ \exp \left( it^T X \right) \right]$, where $t$ is an $n$-dimensional vector and $i = \sqrt{-1}$ is the imaginary number, while the characteristic function of an $n \times p$-dimensional random matrix $Y$ is defined as $\Psi_Y (T) = E \left[ \text{etr} \left( iT^T Y \right) \right]$, where $T$ is an $n \times p$-dimensional matrix. Kim and Genton [29] derived the characteristic function of multivariate skew-normal distribution as presented below.

**Lemma 2.2** (Kim and Genton [29]) For an $n$-dimensional random vector $Z \sim SN_n (\mu, \Omega, \alpha)$, the characteristic function of $Z$ is

\[
\Psi_Z (t) = 2\exp \left( i \mu^T t - \frac{1}{2} t^T \Omega t \right) \Phi \left( i \delta^T t \right) = \exp \left( i \mu^T t - \frac{1}{2} t^T \Omega t \right) \left( 1 + i \tau \left( \delta^T t \right) \right), \tag{2.2}\]

where

\[
\delta = \left( 1 + \alpha^T \overline{\Omega} \alpha \right)^{-1/2} \omega \overline{\Omega} \alpha.
\]

\[
\tau(u) = \sqrt{\frac{2}{\pi}} \int_0^u \exp \left( z^2 / 2 \right) dz, \quad u > 0. \tag{2.3}\]

and $\tau(-u) = -\tau(u)$.
Remark 2.3 The parameter $\delta$ is determined from the skewness parameter $\alpha$ and the scale parameter $\Omega$. The relationship between $\delta$ and $\alpha$ can be shown as in (2.3) and also

$$\alpha = \left(1 - \delta^T \Omega^{-1} \delta \right)^{-1/2} \omega \Omega^{-1} \delta.$$ 

Thus, the parameter $\delta$ can be used for a different description of skewness, and is sometimes used to denote $SN_n(\mu, \Omega, \alpha)$ by $SN_n(\mu, \Omega, \alpha, \delta)$. So, we write $SN_n(\mu, \Omega, \alpha, \ast)$ or $SN_n(\mu, \Omega, \ast, \delta)$ if $\delta$ or $\alpha$ is not important in the ensuing discussion. For a discussion on the parameter $\delta$, one may refer to Azzalini and Capitanio [8].

Azzalini and Capitanio [7] proved that multivariate skew-normal distributions are closed under linear transformations. In fact, Shushi [40] recently proved that all generalized skew-elliptical distributions are closed under affine transformations.

Lemma 2.4 (Azzalini and Capitanio [7]) Suppose $Y \sim SN_n(\mu, \Omega, \alpha, \delta)$. Let $X$ be a linear transformation of $Y$, i.e., $X = A^T Y$, where $A$ is an $n \times p$ full rank matrix. Then, $X \sim SN_p(A^T \mu, \Omega_X, \alpha_X, \delta_X)$, where

$$\Omega_X = A^T \Omega A,$$

$$\alpha_X = \frac{\omega_X \Omega_X^{-1} B^T \alpha}{\sqrt{1 + \alpha^T \left(\Omega - B \Omega_X^{-1} B^T\right) \alpha}}, \quad \quad (2.4)$$

$$\delta_X = A^T \delta \quad \quad (2.5)$$

and

$$B = \omega^{-1} \Omega A.$$

Remark 2.5 Note that (2.5) can be derived by substituting (2.4) into (2.3).

The first-order and second-order moments of multivariate skew-normal distribution have been given by Genton et al. [22] as follows.

Lemma 2.6 (Genton et al. [22]) The mean vector and the second-order moment matrix of $Z \sim SN_n(\mu, \Omega, \ast, \delta)$ are as follows:

$$E(Z) = \mu + \sqrt{\frac{\delta}{\pi}},$$

$$E(ZZ^T) = \Omega + \mu \mu^T + \sqrt{\frac{\delta}{\pi}} \left(\mu \delta^T + \delta \mu^T\right).$$
In particular, if $\mu = 0$, then we deduce

$$E(Z) = \sqrt{\frac{2}{\pi}} \delta,$$

$$E(ZZ^T) = \Omega.$$

The following lemma, called the Fourier Inversion Theorem, can be used to show the relationship between (2.1) and (2.2).

**Lemma 2.7** Suppose $Z$ is an $n$-dimensional random vector, with density function $f_Z(z)$ and characteristic function $\Psi_Z(t)$. Then,

$$f_Z(z) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \exp\left(-it^Tz\right) \Psi_Z(t) dt.$$

If $N$ is an $n \times p$ random matrix, with density function $f_N(N)$ and characteristic function $\Psi_N(T)$, then

$$f_N(N) = (2\pi)^{-np/2} \int_{\mathbb{R}^{n \times p}} \text{etr}\left(-iT^TN\right) \Psi_N(T) dT.$$

The definition of multivariate skew-normal distribution in Definition 2.1 can be extended to the matrix variate case. Harrar and Gupta [24] refer to this distribution as matrix variate skew-normal distribution of Azzalini and Dalla Valle type. The following definition is basically from Harrar and Gupta [24], but we have added a location parameter $M$ to it.

**Definition 2.8** (Harrar and Gupta [24]) An $n \times p$ random matrix $Y$ is said to have matrix variate skew-normal distribution with location matrix $M$, scale matrix $V \otimes \Sigma$ and skewness matrix $B$, denoted by $Y \sim SN_{n \times p}(M, V \otimes \Sigma, B)$, if its probability density function is given by

$$f_Y(Y) = 2\phi_{n \times p}(Y; M, V \otimes \Sigma) \Phi\left(\text{vec}(B)^T \omega^{-1} \text{vec}(Y - M)\right),$$

where $M \in \mathbb{R}^{n \times p}$, $B \in \mathbb{R}^{n \times p}$, $V \in \mathbb{R}^{p \times p}$, $\Sigma \in \mathbb{R}^{n \times n}$, $\omega = v \otimes \sigma$, $v$ and $\sigma$ are defined by $\overline{V} = v^{-1}Vv^{-1}$ and $\overline{\Sigma} = \sigma^{-1}\Sigma\sigma^{-1}$, and $\phi_{n \times p}(Y; M, V \otimes \Sigma)$ denotes the probability density function of an $n \times p$-dimensional matrix variate normal distribution with location matrix $M$ and scale matrix $V \otimes \Sigma$.

The following lemma provides a necessary and sufficient condition for the matrix variate skew-normal distribution. In some works (see Ye et al. [42]), matrix variate skew-normal distributions are defined as in the following lemma.

**Lemma 2.9** (Harrar and Gupta [24]) Random matrix $Y \sim SN_{n \times p}(M, V \otimes \Sigma, B)$ if and only if $y = \text{vec}(Y) \sim SN_{np}(\text{vec}(M), V \otimes \Sigma, \text{vec}(B), \ast)$.

\(\copyright\) Springer
The following lemma provides an identity for the correlation matrix, and shows specifically that the Kronecker product of two correlation matrices is a correlation matrix.

**Lemma 2.10** We assume matrices \( V_{p \times p} = (v_{ij}) \) and \( \Sigma_{n \times n} = (\sigma_{ij}) \) are both positive definite and their diagonal elements are positive, and let matrix \( \Omega_{np \times np} = V \otimes \Sigma = (\omega_{ij}) \). Further, let \( v = \text{diag}(\sqrt{\omega_{11}} \ldots \sqrt{\omega_{pp}}), \sigma = \text{diag}(\sqrt{\sigma_{11}} \ldots \sqrt{\sigma_{nn}}) \) and \( \omega = \text{diag}(\sqrt{\omega_{n1}} \ldots \sqrt{\omega_{nnp}}) \). Then, \( \Omega\omega^{-1} = (Vv^{-1}) \otimes (\Sigma\sigma^{-1}) \) and \( \omega^{-1}\Omega\omega^{-1} = (v^{-1}Vv^{-1}) \otimes (\sigma^{-1}\Sigma\sigma^{-1}) \).

**Proof** We have
\[
\Omega\omega^{-1} = (V \otimes \Sigma) (v^{-1} \otimes \sigma^{-1}) = (Vv^{-1}) \otimes (\Sigma\sigma^{-1}),
\]
\[
\omega^{-1}\Omega\omega^{-1} = (v^{-1} \otimes \sigma^{-1}) (V \otimes \Sigma) (v^{-1} \otimes \sigma^{-1}) = (v^{-1}Vv^{-1}) \otimes (\sigma^{-1}\Sigma\sigma^{-1}).
\]
\(\square\)

**Remark 2.11** Using Lemma 2.10, it can be shown that \( \Omega = V \otimes \Sigma \).

### 2.2 Integral Stochastic Orders

Integral stochastic orders seek orderings between \( X \) and \( Y \) by comparing \( Ef (Y) \) and \( Ef (X) \).

**Definition 2.12** (Denuit et al. [19]) Let \( F \) be a class of measurable functions \( f : \mathbb{R}^{n \times p} \to \mathbb{R} \), and \( X \) and \( Y \) be \( n \)-dimensional random matrices. Then, we say that \( X \leq_F Y \) if \( Ef (X) \leq Ef (Y) \) holds for all \( f \in F \), whenever the expectations are well defined.

Jamali et al. [26] and Yin [43] have discussed \( F \)-class integral stochastic orders through multivariate functions. The definition of \( F \)-class integral stochastic order is extended here to matrix variate functions as we are interested in comparing random matrices. It is easy to verify this extension by means of matrix vectorization.

**Definition 2.13** For a function \( f : \mathbb{R}^{n \times p} \to \mathbb{R} \), the difference operator is defined as
\[
\Delta^\epsilon_{i,j} f (X) = f (X + \epsilon E_{i,j}) - f (X),
\]
where \( E_{i,j} \) is the \((i,j)\)-th unit basis matrix of \( \mathbb{R}^{n \times p} \), for \( i = 1, 2, \ldots, n \), \( j = 1, 2, \ldots, p \), and \( \epsilon > 0 \). Then,

1. \( f \) is supermodular if \( \Delta^\epsilon_{k,l} \Delta^\delta_{i,j} f (X) \geq 0 \) holds for all \( X \in \mathbb{R}^{n \times p}, 1 \leq k < l \leq n, 1 \leq i < j \leq p, \epsilon, \delta > 0 \);
2. \( f \) is directionally convex if \( \Delta^\epsilon_{k,l} \Delta^\delta_{i,j} f (X) \geq 0 \) holds for all \( X \in \mathbb{R}^{n \times p}, k, l = 1, 2, \ldots, n, i, j = 1, 2, \ldots, p, \epsilon, \delta > 0 \);
3. \( f \) is \( \Delta \)-monotone if \( \Delta^\epsilon_{l_1,j_1} \Delta^\epsilon_{l_2,j_2} \cdots \Delta^\epsilon_{l_d,j_d} f (X) \geq 0 \) holds for all \( X \in \mathbb{R}^{n \times p} \), for any subset \( \{(l_1, j_1), (l_2, j_2), \ldots, (l_d, j_d)\} \subseteq \{(k, i) \mid k = 1, 2, \ldots, n, i = 1, 2, \ldots, p\} \) and \( \epsilon_m > 0, m = 1, \ldots, d \).
It needs to be pointed out that the italic letters like $X, Y$ denote random matrices, while roman letters like $X, Y$ denote constants and nonrandom matrices.

**Definition 2.14** 1. Usual stochastic order: $X \leq_{st} Y$ if $F$ is the class of increasing functions;
2. Convex order: $X \leq_{cx} Y$ if $F$ is the class of convex functions;
3. Increasing convex order: $X \leq_{icx} Y$ if $F$ is the class of increasing convex functions;
4. Upper orthant order: $X \leq_{uo} Y$ if $F$ is the class of $\Delta_1$-monotone functions;
5. Supermodular order: $X \leq_{sm} Y$ if $F$ is the class of supermodular functions;
6. Directionally convex order: $X \leq_{dcx} Y$ if $F$ is the class of directionally convex functions.

We consider the functions in the class $F$ to be twice differentiable. The gradient vector and the Hessian matrix of a twice differentiable function $f$ are defined by (see Magnus and Neudecker [31])

$$
\nabla f (X) = \frac{\partial f (X)}{\partial (\text{vec} X)^T} = \left( \frac{\partial}{\partial x_{11}^1} f (X), \frac{\partial}{\partial x_{11}^2} f (X), \ldots, \frac{\partial}{\partial x_{1p}} f (X), \frac{\partial}{\partial x_{21}^1} f (X), \frac{\partial}{\partial x_{21}^2} f (X), \ldots, \frac{\partial}{\partial x_{2p}} f (X), \ldots \right)
$$

and

$$
H_f (X) = \frac{\partial^2 f (X)}{\partial \text{vec} X \partial (\text{vec} X)^T}.
$$

If we assume $\text{vec} X = (x_1, x_2, \ldots, x_{np})$, then

$$
H_f (X) = \begin{pmatrix}
\frac{\partial^2}{\partial x_{11}^1 \partial x_{11}^1} f (X) & \frac{\partial^2}{\partial x_{11}^1 \partial x_{12}^1} f (X) & \ldots & \frac{\partial^2}{\partial x_{11}^1 \partial x_{np}^1} f (X) \\
\frac{\partial^2}{\partial x_{12}^1 \partial x_{11}^1} f (X) & \frac{\partial^2}{\partial x_{12}^1 \partial x_{12}^1} f (X) & \ldots & \frac{\partial^2}{\partial x_{12}^1 \partial x_{np}^1} f (X) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial^2}{\partial x_{np}^1 \partial x_{11}^1} f (X) & \frac{\partial^2}{\partial x_{np}^1 \partial x_{12}^1} f (X) & \ldots & \frac{\partial^2}{\partial x_{np}^1 \partial x_{np}^1} f (X)
\end{pmatrix}_{np \times np}.
$$

**Theorem 2.15** Let the function $f : \mathbb{R}^{n \times p} \rightarrow \mathbb{R}$ be a twice differentiable function. Then,

1. $f$ is increasing if and only if $\frac{\partial}{\partial x_{ij}} f (X) \geq 0$ holds for all $X \in \mathbb{R}^{n \times p}$ and $i = 1, 2, \ldots, n, j = 1, 2, \ldots, p$;
2. $f$ is convex if and only if $H_f (X)$ is positive semi-definite, for all $X \in \mathbb{R}^{n \times p}$;
3. $f$ is increasing convex if and only if $\frac{\partial}{\partial x_{ij}} f (X) \geq 0$ holds and $H_f (X)$ is positive semi-definite, for all $X \in \mathbb{R}^{n \times p}$ and $i = 1, 2, \ldots, n, j = 1, 2, \ldots, p$;

Springer
4. $f$ is supermodular if and only if $\frac{\partial^2}{\partial x_{kl} \partial x_{ij}} f(X) \geq 0$ holds for all $X \in \mathbb{R}^{n \times p}$, $1 \leq k, l \leq n$, $1 \leq i, j \leq p$, and $(k, j) \neq (l, i)$;
5. $f$ is directionally convex if and only if $\frac{\partial^2}{\partial x_{kl} \partial x_{ij}} f(X) \geq 0$ holds for all $X \in \mathbb{R}^{n \times p}$, $1 \leq k, l \leq n$, $1 \leq i, j \leq p$.

**Proof** Matrix function $f(X) : \mathbb{R}^{n \times p} \rightarrow \mathbb{R}$ can be treated as a multivariate function $f(\text{vec}X) : \mathbb{R}^{np} \rightarrow \mathbb{R}$. So, the results in the theorem can be derived from the results for multivariate functions presented by Arlotto and Scarsini [5] and Denuit et al. [18].

3 Characteristic Function of the Matrix Variate Skew-normal Distribution

The following theorem presents the characteristic function of the matrix variate skew-normal distribution.

**Theorem 3.1** Let $Z \sim SN_{n \times p}(M, V \otimes \Sigma, B)$. Then, the characteristic function of $Z$ is

$$
\Psi_Z(T) = 2\exp\left(iM^T T - \frac{1}{2} T^T \Sigma T V \right) \Phi \left( i \frac{\text{tr} \left( V^{-1}B^T \sigma^{-1} \Sigma T \right)}{\sqrt{1 + B^T \Sigma BV}} \right),
$$

(3.1)

where $v = \text{diag}\{\sqrt{v_{11}}, \ldots, \sqrt{v_{pp}}\}$, $\sigma = \text{diag}\{\sqrt{\sigma_{11}}, \ldots, \sqrt{\sigma_{nn}}\}$, and $i = \sqrt{-1}$.

**Proof** Let $z = \text{vec}(Z)$. We then have $z \sim SN_{np}(\text{vec}(M), V \otimes \Sigma, \text{vec}(B))$ from Lemma 2.9. We choose matrix $T$ such that $t = \text{vec}(T)$. The characteristic function of $z$ is then

$$
\Psi_z(t) = 2\exp\left(i\text{vec}(M)^T t - \frac{1}{2} t^T (V \otimes \Sigma) t \right) \Phi \left( i \frac{\text{vec}(\Sigma \sigma^{-1} B^{-1} V)^T}{\sqrt{1 + \text{tr} (B^T \Sigma BV)}} \right).
$$

We then have

$$
\Psi_z(\text{vec}(T)) = 2\exp\left(iM^T T - \frac{1}{2} T^T \Sigma T V \right) \Phi \left( i \frac{\text{tr} \left( V^{-1}B^T \sigma^{-1} \Sigma T \right)}{\sqrt{1 + \text{tr} (B^T \Sigma BV)}} \right). \tag{3.2}
$$

Due to the fact that

$$
E\left(\text{etr} \left(iZ^T T \right) \right) = E\left(\exp \left(i\text{vec}(Z)^T \text{vec}(T) \right) \right).
$$
we have

\[ \Psi_Z(T) = 2\exp\left( iM^T T - \frac{1}{2} T^T \Sigma T V \right) \Phi \left( i \frac{\tr (Vv^{-1}B^T \sigma^{-1} \Sigma T)}{\sqrt{1 + \tr (B^T \Sigma BV)}} \right) \]

as required. \( \square \)

**Remark 3.2** If we set \( p = 1 \) in Theorem 3.1, the characteristic function of \( X \sim SN_{n \times 1} (M, 1 \otimes \Sigma, B) \) simplifies to

\[ \Psi_X(T) = 2\exp\left( iM^T T - \frac{1}{2} T^T \Sigma T V \right) \Phi \left( i \frac{B^T \sigma^{-1} \Sigma T}{\sqrt{1 + B^T \Sigma B}} \right). \]

If we regard \( SN_{n \times 1} (M, 1 \otimes \Sigma, B) \) as a multivariate skew-normal distribution, then its characteristic function is identical to the result in Lemma 2.2.

**Remark 3.3** If we set the skew matrix \( B = 0_{n \times p} \), then the matrix variate skew-normal distribution \( SN_{n \times p} (M, V \otimes \Sigma, B, \delta) \) will degenerate to a matrix variate normal distribution \( N_{n \times p} (M, V \otimes \Sigma) \), with characteristic function \( \exp \left( iM^T T - \frac{1}{2} T^T \Sigma T V \right) \), which is indeed the characteristic function of a matrix variate normal distribution.

**Remark 3.4** If we set

\[ \delta = \frac{(V \otimes \Sigma)(v^{-1} \otimes \sigma^{-1}) \vec{B}}{\sqrt{1 + \tr (B^T \Sigma BV)}} = \frac{\vec{\Sigma \sigma^{-1} Bv^{-1} V}}{\sqrt{1 + \tr (B^T \Sigma BV)}} \]

(3.1) can be expressed as

\[ \Psi_Z(T) = 2\exp\left( i \vec{M^T} \vec{T} - \frac{1}{2} \vec{T}^T (V \otimes \Sigma) \vec{T} \right) \Phi \left( i \delta^T \vec{T} \right). \]

As in the multivariate case, in the ensuing discussion, we write \( SN_{n \times p} (M, V \otimes \Sigma, B, \delta) \) for the matrix variate skew-normal distribution. We write \( SN_{n \times p} (M, V \otimes \Sigma, B, \ast) \) or \( SN_{n \times p} (M, V \otimes \Sigma, \ast, \delta) \) if parameters \( \delta \) or \( B \) are not important in the discussion.

**Remark 3.5** Let \( Z \sim SN_{n \times p} (M, V \otimes \Sigma, B, \delta) \). Then, we can deduce from Lemma 2.9 that \( \vec{Z} \sim SN_{np} (\vec{M}, V \otimes \Sigma, \vec{B}, \delta) \) and that they share the same parameter \( \delta \).

**4 Main results**

The next theorem presents an identity for \( Ef(Y) - Ef(X) \), and it extends an identity of Müller [33] for the case of multivariate normal distribution.
Theorem 4.1 Suppose the \( n \times p \) skew-normal random matrices \( X \) and \( Y \) are distributed as
\[
X \sim SN_{n \times p} (M, \Omega, B, \delta), \quad Y \sim SN_{n \times p} (M', \Omega', B', \delta'),
\]
where \( M, M', B, B' \in \mathbb{R}^{n \times p} \), \( \Omega = V_{p \times p} \otimes \Sigma_{n \times n} \), \( \Omega' = V'_{p \times p} \otimes \Sigma'_{n \times n} \). \( V, V', \Sigma \) and \( \Sigma' \) are positive definite, and \( \delta, \delta' \in \mathbb{R}^{np} \). Suppose an \( n \times p \)-dimensional matrix variate skew-normal random variable \( Z_\lambda \) is distributed as
\[
Z_\lambda \sim SN_{n \times p} (M_\lambda, \Omega_\lambda, \ast, \delta_\lambda),
\]
where
\[
M_\lambda = \lambda M' + (1 - \lambda) M, \\
\Omega_\lambda = \lambda \Omega' + (1 - \lambda) \Omega, \\
\delta_\lambda = \lambda \delta' + (1 - \lambda) \delta.
\]

Let the density functions of random matrices \( X, Y \) and \( Z_\lambda \) be \( \phi_0, \phi_1 \) and \( \phi_\lambda \), respectively, and the corresponding characteristic functions be \( \Psi_0, \Psi_1 \) and \( \Psi_\lambda \). Let \( \overline{\phi}_\lambda \) and \( \overline{\Psi}_\lambda \) be the density function and characteristic function of matrix variate normal distribution with mean matrix \( M_\lambda \) and covariance \( \Omega_\lambda - \delta_\lambda \delta_\lambda' \). Suppose the function \( f : \mathbb{R}^{n \times p} \rightarrow \mathbb{R} \) is twice differentiable and is such that
\[
\begin{align*}
1. \lim_{x_{i1} \rightarrow \pm \infty} f (X) \phi_\lambda (X) = 0, & \quad \lim_{x_{i1} \rightarrow \pm \infty} f (X) \overline{\phi}_\lambda (X) = 0, \quad \forall X \in \mathbb{R}^{n \times p}, 0 \leq \lambda \leq 1, l = 1, 2, \ldots, n \text{ and } i = 1, 2, \ldots, p, \\
2. \lim_{x_{i1} \rightarrow \pm \infty} f (X) \frac{\partial}{\partial x_{kj}} \phi_\lambda (X) = 0, & \quad \lim_{x_{i1} \rightarrow \pm \infty} f (X) \frac{\partial}{\partial x_{kj}} \overline{\phi}_\lambda (X) = 0, \quad \forall X \in \mathbb{R}^{n \times p}, 0 \leq \lambda \leq 1, k, l = 1, 2, \ldots, n \text{ and } i = 1, 2, \ldots, p, \\
3. \lim_{x_{i1} \rightarrow \pm \infty} \phi_\lambda (X) \frac{\partial}{\partial x_{kj}} f (X) = 0, & \quad \lim_{x_{i1} \rightarrow \pm \infty} \overline{\phi}_\lambda (X) \frac{\partial}{\partial x_{kj}} f (X) = 0, \quad \forall X \in \mathbb{R}^{n \times p}, 0 \leq \lambda \leq 1, k, l = 1, 2, \ldots, n \text{ and } i, j = 1, 2, \ldots, p.
\end{align*}
\]

Then,
\[
\begin{align*}
E [ f (Y) - f (X) ] &= \int_0^1 \int_{\mathbb{R}^{n \times p}} \left( \nabla f (X) \text{vec} (M' - M) + \frac{1}{2} \text{tr} \left( (\Omega' - \Omega) H_f (X) \right) \right) \phi_\lambda (X) \\
& \quad + \frac{2}{\sqrt{2\pi}} \left( \nabla f (X) (\delta' - \delta) \right) \overline{\phi}_\lambda (X) dX d\lambda.
\end{align*}
\]

Proof Let \( g (\lambda) = \int_{\mathbb{R}^{n \times p}} f (X) \phi_\lambda (X) dX \). Then, we have \( g (0) = Ef (X), g (1) = Ef (Y) \) and
\[
Ef (Y) - Ef (X) = \int_0^1 \frac{\partial}{\partial \lambda} g (\lambda) d\lambda = \int_0^1 \int_{\mathbb{R}^{n \times p}} f (X) \frac{\partial}{\partial \lambda} \phi_\lambda (X) dX d\lambda.
\]

According to the Fourier inversion formula in Lemma 2.7, we have
\[
\frac{\partial}{\partial \lambda} \phi_\lambda (X) = \frac{1}{(2\pi)^{np}} \int_{\mathbb{R}^{n \times p}} \text{etr} \left( -iT'^T X \right) \frac{\partial}{\partial \lambda} \Psi_\lambda (T) dT.
\]

 Springer
Using Theorem 3.1, we have

\[
\frac{\partial}{\partial \lambda} \Psi_\lambda (T) = \frac{\partial}{\partial \lambda} 2 \exp \left( i \text{vec} (M_\lambda)^T \text{vec} (T) - \frac{1}{2} \text{vec} (T)^T \Omega_\lambda \text{vec} (T) \right) \Phi \left( i \delta^T_\lambda \text{vec} (T) \right) \\
= \{ i \text{vec} (T)^T \text{vec} (M' - M) - \frac{1}{2} \text{vec} (T)^T (\Omega' - \Omega) \text{vec} (T) \} \Psi_\lambda (T) \\
+ \frac{2}{\sqrt{2\pi}} \left( i \left( \delta' - \delta \right)^T \text{vec}(T) \right) \overline{\Psi}_\lambda (T).
\]

Upon substituting (4.5) into (4.4), we get

\[
\frac{\partial}{\partial \lambda} \phi_\lambda (X) = \frac{1}{2} \sum_{k,l=1}^{n} \sum_{i,j=1}^{p} \left( v'_{ij} \sigma'_{kl} - v_{ij} \sigma_{kl} \right) \frac{\partial^2}{\partial x_{ki} \partial x_{lj}} \phi_\lambda (X) \\
- \sum_{k=1}^{n} \sum_{j=1}^{p} \left( m'_{kj} - m_{kj} \right) \frac{\partial}{\partial x_{kj}} \phi_\lambda (X) \\
- \frac{2}{\sqrt{2\pi}} \sum_{k=1}^{n} \sum_{j=1}^{p} \left( \delta'_{k+n(j-1)} - \delta_{k+n(j-1)} \right) \frac{\partial}{\partial x_{kj}} \overline{\phi}_\lambda (X).
\]

Substituting (4.6) back into (4.3), we obtain

\[
\int_{\mathbb{R}^n \times p} f (X) \frac{\partial}{\partial \lambda} \phi_\lambda (X) dX = \frac{1}{2} \sum_{k,l=1}^{n} \sum_{i,j=1}^{p} \left( v'_{ij} \sigma'_{kl} - v_{ij} \sigma_{kl} \right) \int_{\mathbb{R}^n \times p} f (X) \frac{\partial^2}{\partial x_{ki} \partial x_{lj}} \phi_\lambda (X) dX \\
+ \sum_{k=1}^{n} \sum_{j=1}^{p} \left( m'_{kj} - m_{kj} \right) \int_{\mathbb{R}^n \times p} f (X) \frac{\partial}{\partial x_{kj}} \phi_\lambda (X) dX \\
+ \frac{2}{\sqrt{2\pi}} \sum_{k=1}^{n} \sum_{j=1}^{p} \left( \delta'_{k+n(j-1)} - \delta_{k+n(j-1)} \right) \int_{\mathbb{R}^n \times p} f (X) \frac{\partial}{\partial x_{kj}} \overline{\phi}_\lambda (X) dX.
\]

Now, integrating by parts and then using the properties of \( f \), we get

\[
\int_{\mathbb{R}^n \times p} f (X) \frac{\partial}{\partial \lambda} \phi_\lambda (X) dX = \frac{1}{2} \sum_{k,l=1}^{n} \sum_{i,j=1}^{p} \left( v'_{ij} \sigma'_{kl} - v_{ij} \sigma_{kl} \right) \int_{\mathbb{R}^n \times p} \phi_\lambda (X) \frac{\partial^2}{\partial x_{ki} \partial x_{lj}} f (X) dX \\
+ \sum_{k=1}^{n} \sum_{j=1}^{p} \left( m'_{kj} - m_{kj} \right) \int_{\mathbb{R}^n \times p} \phi_\lambda (X) \frac{\partial}{\partial x_{kj}} f (X) dX \\
+ \frac{2}{\sqrt{2\pi}} \sum_{k=1}^{n} \sum_{j=1}^{p} \left( \delta'_{k+n(j-1)} - \delta_{k+n(j-1)} \right) \int_{\mathbb{R}^n \times p} \overline{\phi}_\lambda (X) \frac{\partial}{\partial x_{kj}} f (X) dX \\
= \int_{\mathbb{R}^n \times p} \left( \nabla f (X) \text{vec} (M' - M) \right) + \frac{1}{2} \text{tr} \left( (\Omega' - \Omega) H_f (X) \right) \phi_\lambda (X) \\
+ \frac{2}{\sqrt{2\pi}} \left( \nabla f (X) (\delta' - \delta) \right) \overline{\phi}_\lambda (X) dX.
\]

Substitution of (4.7) into (4.3) yields the result in (4.2).
Theorem 4.2 Suppose the \( n \times p \) skew-normal random matrices \( X \) and \( Y \) and the function \( f \) satisfy the conditions in Theorem 4.1. If

1. \( \sum_{k,l=1}^{n} \sum_{i,j=1}^{p} \left( v'_{ij} \sigma'_{kl} - v_{ij} \sigma_{kl} \right) \frac{\partial^2}{\partial x_{ki} \partial x_{lj}} f(X) \geq 0, \)
2. \( \sum_{k=1}^{n} \sum_{j=1}^{p} \left( m'_{kj} - m_{kj} \right) \frac{\partial}{\partial x_{kj}} f(X) \geq 0, \)
3. \( \sum_{k=1}^{n} \sum_{j=1}^{p} \left( \delta_{i+n(j-1)}' - \delta_{i+n(j-1)} \right) \frac{\partial}{\partial x_{kj}} f(X) \geq 0, \)

then \( Ef(Y) \geq Ef(X). \)

5 Stochastic Orderings

In the rest of this paper, we introduce some conditions to compare skew-normal matrices \( X \) and \( Y \). Usual stochastic order is one of the most widely used groups of ordering relations and have been extensively used by insurance practitioners. Besides Definition 2.14, there is an equivalent condition for the usual stochastic order of random variables \( X \leq_{st} Y \) is that \( \overline{F}_X(t) \leq \overline{F}_Y(t) \) for all \( t \in \mathbb{R} \). Jamali et al. [26] gave a necessary and sufficient condition for comparing univariate skew-normal variables under usual stochastic order.

Lemma 5.1 (Jamali et al. [26]) Let \( X_1 \sim SN_1 \left( \mu_1, \sigma^2_1, *, \delta_1 \right) \) and \( X_2 \sim SN_1 \left( \mu_2, \sigma^2_2, *, \delta_2 \right) \). Then, \( X_1 \leq_{st} X_2 \) if and only if \( \mu_1 \leq \mu_2, \sigma_1 = \sigma_2 \) and \( \delta_1 \leq \delta_2. \)

In the following discussion, we say \( a > b \), where \( a \) and \( b \) are vectors, if and only if \( a_i > b_i \) for all \( i \). We say \( A > B \), where \( A \) and \( B \) are matrices, if and only if \( a_{ij} > b_{ij} \) for all \( i, j \). We suppose random matrices \( X \) and \( Y \) satisfy (4.1) and random matrices \( X_0 \) and \( Y_0 \) are such that \( X_0 \sim SN_{n \times p} \left( \mathbf{0}_{n \times p}, \Omega, *, \delta \right) \) and \( Y_0 \sim SN_{n \times p} \left( \mathbf{0}_{n \times p}, \Omega', *, \delta' \right) \).

The following theorem, which is an extension of Lemma 5.1, provides a necessary and sufficient condition for comparing skew-normal matrices under stochastic order.

Theorem 5.2 \( X \leq_{st} Y \) if and only if \( M \leq M', \delta \leq \delta' \) and \( \Omega = \Omega' \).

Proof Suppose \( M \leq M', \delta \leq \delta' \) and \( \Omega = \Omega' \). We know that for any increasing differentiable function \( f, \frac{\partial}{\partial x_{kj}} f(X) \geq 0, i = 1, 2, \ldots, n, j = 1, 2, \ldots, p \). Then, the conditions of Theorem 4.2 are satisfied.

To prove the converse, we assume \( X \leq_{st} Y \). It is then easy to see that \( X_{ij} \leq_{st} Y_{ij} \), \( i = 1, 2, \ldots, n \), \( j = 1, 2, \ldots, p \), where

\[
X_{ij} \sim SN_1 \left( m_{ij}, v_{ii} \sigma_{jj}, *, \delta_{i+n(j-1)} \right), Y_{ij} \sim SN_1 \left( m'_{ij}, v'_{ii} \sigma'_{jj}, *, \delta'_{i+n(j-1)} \right).
\]

Equation (5.1) is derived by Lemmas 2.4 and 2.9. According to Lemma 5.1, we know that \( m_{ij} \geq m'_{ij}, \delta_{i+n(j-1)} \geq \delta'_{i+n(j-1)} \), and \( v_{ii} \sigma_{jj} = v'_{ii} \sigma'_{jj} \).

\( \square \) Springer
Also we know that $X_{ij} + X_{kl} \leq_{st} Y_{ij} + Y_{kl}$ since $X \leq_{st} Y$, where

$$
X_{ij} + X_{kl} \sim SN_1 \left( m_{ij} + m_{kl}, v_{ii}\sigma_{jj} + v_{kk}\sigma_{ll} + 2v_{ki}\sigma_{lj}, *, \delta_{i+n(j-1)} + \delta_{k+n(l-1)} \right),
$$

(5.2)

$$
Y_{ij} + Y_{kl} \sim SN_1 \left( m'_{ij} + m'_{kl}, v'_i\sigma'_{jj} + v'_k\sigma'_{ll} + 2v'_k\sigma'_{lj}, *, \delta'_i + \delta'_j + \delta'_k + \delta'_l \right).
$$

(5.3)

(5.2) and (5.3) are derived by using Lemmas 2.4 and 2.9. According to Lemma 5.1, we know that $v_{ki}\sigma_{lj} = v'_k\sigma'_{lj}$, which means that $\Omega = \Omega'$.

Suppose $\Omega = V \otimes \Sigma$ and $\Omega' = V' \otimes \Sigma'$. Then, we have $\Omega = \Omega'$ if and only if there exists $a \in \mathbb{R}, a \neq 0$, such that $V = aV'$, $\Sigma = a^{-1}\Sigma'$. This is called the uniqueness of Kronecker product.

**Corollary 5.3** If convex function $f$ is twice differentiable, then Hessian matrix $H_f(X)$ is positive semi-definite. As $\Omega' - \Omega$ is positive semi-definite, there exists a matrix $A_{np \times np}$ such that $\Omega' - \Omega = AA^T$. Suppose $A = (a_1, a_2, \ldots, a_{np})$, where $a_i$ is an $np$-dimensional column vector, for $i = 1, 2, \ldots, np$. Then, according to the properties of trace of matrices, we have

$$
\text{tr} \left( (\Omega - \Omega') H_f(X) \right) = \text{tr} \left( A^T H_f(X) A \right) = \sum_{i=1}^{np} a_i^T H_f(X) a_i \geq 0.
$$

According to Theorem 4.2, we have $Ef(Y) - Ef(X) \geq 0$ for all twice differentiable convex functions $f$, which means $X \leq_{cx} Y$. 

\(\Box\) Springer
2. & 3. By letting $f(X) = x_{ij}$ and $f(Y) = -x_{ij}$ and using Definitions 2.12 and 2.14, we can easily see that the mean matrices of $X$ and $Y$ are identical. According to Lemma 2.6, $\delta = \delta'$ can be derived from $M = M'$ and vice versa.

If $\Omega' - \Omega$ is not positive semi-definite, then there exists an $np$-dimensional column vector $b$ such that $b^T \Omega b > b^T \Omega' b$. Let $f(X) = (b^T \text{vec}(X))^2 = b^T \text{vec}(X) \text{vec}(X)^T b$, which is obviously a convex function. Using Lemma 2.6, we have

$$E(f(Y)) - E(f(X)) = b^T \Omega' b - b^T \Omega b < 0,$$

which contradicts $X \leq_{cx} Y$. \hfill $\square$

**Corollary 5.6** $X_0 \leq_{cx} Y_0$ if and only if $\delta = \delta'$ and $\Omega - \Omega'$ is positive semi-definite.

Increasing convex order is also known as stop-loss order because of the fact that $X \leq_{icx} Y$ if and only if their stop-loss transforms $E((X - t)_+) \leq E((Y - t)_+)$ for all $t \in \mathbb{R}$. In the perspective of insurance, stop-loss order can be interpreted as a comparison of stop-loss premiums. The following theorem provides conditions for comparing skew-normal random matrices under increasing convex order. We recall that a matrix $A$ is said to be copositive if $x^T A x \geq 0$ for all $x \geq 0$; see Hadeler [23] for details.

**Theorem 5.7**

1. If $M \preceq M'$, $\delta \preceq \delta'$ and $\Omega' - \Omega$ is positive semi-definite, then $X \preceq_{icx} Y$;

2. If $X \preceq_{icx} Y$ and $M = M'$, then $\delta \preceq \delta'$ and $\Omega' - \Omega$ is copositive;

3. If $X \preceq_{icx} Y$ and $\delta = \delta'$, then $M \preceq M'$ and $\Omega' - \Omega$ is copositive.

**Proof**

1. If $f$ is an increasing convex function, then we have $\frac{\partial}{\partial x_{ij}} f(X) \geq 0$ and $H_f(X)$ is positive semi-definite. As we know that $M \preceq M'$, $\delta \preceq \delta'$ and $\Omega - \Omega'$ is positive semi-definite. Then, the conditions of Theorem 4.2 are satisfied, which means that $X \preceq_{icx} Y$.

2. & 3. By letting $f(X) = x_{ij}$, which is an increasing convex function, we see that $EX_0 \leq EY_0$. If $M = M'$, we get $\delta \preceq \delta'$ by using Lemma 2.6 and vice versa.

Let $f(X) = a^T \text{vec}(X)$, $a \geq 0$ and $a \neq 0$, which is an increasing convex function. Then, we know that $X \preceq_{icx} Y$, where

$$X = a^T \text{vec}(X) \sim SN_1(a^T \text{vec}(M), a^T \Omega a, *, a^T \delta), \quad (5.4)$$

$$Y = a^T \text{vec}(Y) \sim SN_1(a^T \text{vec}(M), a^T \Omega' a, *, a^T \delta'), \quad (5.5)$$

which are derived by using Lemma 2.4. Let $\sigma^2 = a^T \Omega a$ and $(\sigma')^2 = a^T \Omega' a$. We now claim that $\sigma \leq \sigma'$. If $\sigma > \sigma'$, then
\[
\lim_{t \to +\infty} \frac{E((\overline{Y} - t)_+)}{E((\overline{X} - t)_+)} = \lim_{t \to +\infty} \int_t^{+\infty} \frac{1 - F_Y(x)}{1 - F_X(x)} \, dx = \lim_{t \to +\infty} \frac{F_Y(t) - 1}{F_X(t) - 1} = \lim_{t \to +\infty} f_Y(t) f_X(t) = 0,
\]

where the last equality comes from Definition 2.1. This contradicts the fact that \(\overline{X} \leq_{icx} \overline{Y}\) since both \(E((\overline{Y} - t)_+\) and \(E((\overline{X} - t)_+)\) are nonnegative. This means that \(\Omega - \Omega'\) is copositive. \(\square\)

Directionally convex order, which is presented in Definition 2.14, can also be considered as an extension of convex order by means of difference operators. The directionally convex orders not only compare the dependence structures of two random vectors, but also the variability of the marginals. Interested readers may refer to [11] for more details. For directionally convex order, we have the following result.

**Theorem 5.8** \(X_0 \leq_{dcx} Y_0\ if\ and\ only\ if\ \delta = \delta' \) and \(\Omega' - \Omega \geq 0\).

**Proof** As the sufficiency is obvious, we just prove the necessity. \(\delta = \delta'\) can be proved by using the same argument as in the proof of Theorem 5.5, since both \(f(X) = x_{ij}\) and \(f(X) = -x_{ij}\) are directionally convex functions.

\(E(X_{0ij}X_{0kl}) \leq E(Y_{0ij}Y_{0kl})\) can be shown by letting \(f(X) = x_{ij}x_{kl}\). Thus, it can be shown that \(E(\text{vec}(X_0)\text{vec}(X_0)^T) \leq E(\text{vec}(Y_0)\text{vec}(Y_0)^T)\), which yields \(\Omega' - \Omega \geq 0\) by using Lemma 2.6. \(\square\)

Let \(\overline{\Omega} = \overline{V} \otimes \overline{\Sigma}\) and \(\overline{\Omega}' = \overline{V}' \otimes \overline{\Sigma}'\). Under the setup of correlation matrix, the two decompositions are unique.

**Corollary 5.9** \(X_0 \leq_{dcx} Y_0\ if\ and\ only\ if\ \delta = \delta', \overline{V} - \overline{V} \geq 0\ and\ \overline{\Sigma}' - \overline{\Sigma} \geq 0\).

Azzalini and Capitanio [8] have presented a conditional stochastic representation of multivariate skew-normal distribution as follows.

**Lemma 5.10** (Azzalini and Capitanio [8]) If \(X \sim SN_n(\mu, \Omega, \alpha, \ast)\), then \(X \overset{d}{=} \mu + U \mid \{V < \alpha^T \omega^{-1} U\}\), where \(U \sim N_n(0, \Omega)\) and is independent of the random variable \(V \sim N(0, 1)\).

The upper orthant order, given in Definition 2.14, can also be defined through a comparison of upper orthants. Specifically, \(X \leq_{uo} Y\) holds if and only if \(P(X > T) \leq P(Y > T)\) holds for all matrices \(T\). The two definitions can be shown to be equivalent. The following theorem provides conditions for comparing matrix variate skew-normal distributions under upper orthant order.

**Theorem 5.11** 1. If \(M \leq M', \delta \leq \delta', \omega_{ii} = \omega_{ii}'\) and \(\omega_{ij} \leq \omega_{ij}'\), then \(X \leq_{uo} Y\); 2. If \(X \leq_{uo} Y\), then \(M \leq M', \delta \leq \delta'\) and \(\omega_{ii} = \omega_{ii}'\).

\(\square\) Springer
Proof 1. Let $X$ and $Y$ be distributed as in (4.1), and that $X$ and $Y$ have skewness parameters $B$ and $B'$, respectively. Then, we have $\text{vec} (X) \sim SN_{np} (\text{vec} (M), \Omega, \text{vec} (B), \delta)$ and $\text{vec} (Y) \sim SN_{np} (\text{vec} (M'), \Omega', \text{vec} (B'), \delta')$ by using Lemma 2.9 and Remark 3.5. According to Lemma 5.10, we have $\text{vec} (X) \overset{d}{=} \text{vec} (M) + U | \{ V < \text{vec} (B)^T \omega^{-1} U \}$ and $\text{vec} (Y) \overset{d}{=} \text{vec} (M') + U' | \{ V' < \text{vec} (B')^T \omega'^{-1} U' \}$. Then,

$$P (\text{vec} (X) > t) = 2P (Z > (t - \text{vec} (M), 0)),$$

$$P (\text{vec} (Y) > t) = 2P (Z' > (t - \text{vec} (M'), 0)),$$

where

$$Z \sim N_{np+1} \left( 0, \begin{pmatrix} \Omega & \delta \\ \delta^T & 1 \end{pmatrix} \right), \quad Z' \sim N_{np+1} \left( 0, \begin{pmatrix} \Omega' & \delta' \\ \delta'^T & 1 \end{pmatrix} \right).$$

Considering the conditions $M \leq M'$, $\delta \leq \delta'$, $\omega_{ii} = \omega'_{ii}$ and $\omega_{ij} \leq \omega'_{ij}$ and using Slepian’s inequality (Theorem 2.1.1 in Tong [41]), we can conclude that $P (X > T) \leq P (Y > T)$, which means $X \leq_{uo} Y$.

2. $X_{ij} \leq_{st} Y_{ij}$ can be obtained from $X \leq_{uo} Y$ due to the equivalence between stochastic order and upper orthant order in the univariate case. Then, $M \leq M'$, $\delta \leq \delta'$ and $\omega_{ii} = \omega'_{ii}$ can be proved by using the same idea as in the proof of Theorem 5.2. □

Supermodular orders are important for a wide range of scientific and industrial processes. Several practical applications for supermodular orders, like applications in genetic selection, are stated in Bäuerle [12]. The following result generalizes Theorem 11 in Müller [33] for the multivariate normal case to the setting considered here.

**Theorem 5.12** $X_0 \leq_{sm} Y_0$ if and only if $X_0$ and $Y_0$ have the same marginals and $\omega_{ij} \leq \omega'_{ij}$.

**Proof** If $X_0 \leq_{sm} Y_0$, then $X_0$ and $Y_0$ have the same marginals (see Müller [34]). This yields $\delta = \delta'$ since

$$X_{0ij} \sim SN_1 \left( 0, v_{ii} \sigma_{jj}, *, \delta_{l+n(j-1)} \right), \quad Y_{0ij} \sim SN_1 \left( 0, v'_{ii} \sigma'_{jj}, *, \delta'_{l+n(j-1)} \right). \quad (5.6)$$

We can prove $E (X_{0ij}X_{0kl}) \leq E (Y_{0ij}Y_{0kl})$ by letting $f (X) = x_{ij} x_{kl}$, which is a supermodular function. Then, $E (\text{vec} (X_0) \text{vec} (X_0)^T) \leq E (\text{vec} (Y_0) \text{vec} (Y_0)^T)$, which yields that $\omega_{ij} \leq \omega'_{ij}$ by using Lemma 2.6.

Conversely, if $X_0$ and $Y_0$ have the same marginals and $\omega_{ij} \leq \omega'_{ij}$, then $X_0 \leq_{sm} Y_0$ can be obtained from Theorem 4.2, since $\frac{\partial^2}{\partial x_{ij} \partial x_{kl}} f (X) \geq 0$ for supermodular function $f$ and $\delta = \delta'$.

### 6 Concluding Remarks

In this paper, the characteristic function of the matrix variate skew-normal distribution has been derived and has been utilized to establish integral stochastic orderings of
matrix variate skew-normal distributions. Considering six different important integral stochastic orderings, some necessary and sufficient conditions have been derived. It will naturally be of interest to further generalize these results to matrix variate skew-elliptical family of distributions. We are currently working on this problem and hope to report the findings in a future paper.

Acknowledgements The authors would like to thank the anonymous reviewers for their valuable comments and suggestions. This research was supported by the National Natural Science Foundation of China (No. 12071251, 11571198, 11701319).

References

1. Abdi, M., Madadi, M., Balakrishnan, N., Jamalizadeh, A.: Family of mean-mixtures of multivariate normal distributions: properties, inference and assessment of multivariate skewness. J. Multiv. Anal. 181, 104679 (2021)
2. Amiri, M., Izadkhah, S., Jamalizadeh, A.: Linear orderings of the scale mixtures of the multivariate skew-normal distribution. J. Multiv. Anal. 179, 104647 (2020)
3. Anderlucci, L., Virol, C.: Covariance pattern mixture models for the analysis of multivariate heterogeneous longitudinal data. Annal. Appl. Statit. 9(2), 777–800 (2015)
4. Ansari, J., Rüschendorf, L.: Ordering results for elliptical distributions with applications to risk bounds. J. Multiv. Anal. 182, 104709 (2021)
5. Arlotto, A., Scarsini, M.: Hessian orders and multinormal distributions. J. Multiv. Anal. 100(10), 2324–2330 (2009)
6. Azzalini, A.: A class of distributions which includes the normal ones. Scandinavian J. Statit. 12(2), 171–178 (1985)
7. Azzalini, A., Capitanio, A.: Statistical applications of the multivariate skew normal distribution. J. R. Stat. Soc.: Ser. B 61(3), 579–602 (1999)
8. Azzalini, A., Capitanio, A.: The Skew-normal and related families. Cambridge University Press, England (2013)
9. Azzalini, A., Dalla, V.A.: The multivariate skew-normal distribution. Biometrika 83(4), 715–726 (1996)
10. Azzalini, A., Regoli, G.: Some properties of skew-symmetric distributions. Annal. Instit. Stat. Math. 64(4), 857–879 (2012)
11. Balakrishnan, N., Belzunce, F., Sordo, M.A., Suárez-Llorens, A.: Increasing directionally convex orderings of random vectors having the same copula, and their use in comparing ordered data. J. Multiv. Anal. 105(1), 45–54 (2012)
12. Bäuerle, N.: Inequalities for stochastic models via supermodular orderings. Stoch. Models 13(1), 181–201 (1997)
13. Belzunce, F.: An introduction to the theory of stochastic orders. Boletín de Estadística e Investigacion Operativa (SEIO) 26(1), 4–18 (2010)
14. Caro-Lopera, F.J., Leiva, V., Balakrishnan, N.: Connection between the Hadamard and matrix products with an application to matrix-variate Birnbaum-Saunders distributions. J. Multiv. Anal. 104(1), 126–139 (2012)
15. Chen, E.Y., Tsay, R.S., Chen, R.: Constrained factor models for high-dimensional matrix-variate time series. J. Am. Stat. Assoc. 115, 775–793 (2020)
16. Chen, J.T., Gupta, A.K.: Matrix variate skew normal distributions. Statistics 39(3), 247–253 (2005)
17. Davidov, O., Peddada, S.: The linear stochastic order and directed inference for multivariate ordered distributions. Annal. Stat. 41, 1–40 (2013)
18. Denuit, M., Müller, A.: Smooth generators of integral stochastic orders. Ann. Appl. Prob. 12(4), 1174–1184 (2002)
19. Denuit, M., Dhene, J., Goovaerts, M., Kaas, R.: Actuarial theory for dependent risks: measures orders and models. Wiley, Chichester, England (2006)
20. Ding, Y., Zhang, X.: Some stochastic orders of Kotz-type distributions. Stat. Probab. Lett. 69, 389–396 (2004)
21. Domínguez-Molina, J.A., González-Farías, G., Ramos-Quiroga, R., Gupta, A.K.: A matrix variate closed skew-normal distribution with applications to stochastic frontier analysis. Commun. Stat.-Theory Methods 36(9), 1691–1703 (2007)

22. Genton, M.G., He, L., Liu, X.: Moments of skew-normal random vectors and their quadratic forms. Stat. Probab. Lett. 51(4), 319–325 (2001)

23. Hadeler, K.-P.: On copositive matrices. Linear Algebra Appl. 49, 79–89 (1983)

24. Harrar, S.W., Gupta, A.K.: On matrix variate skew-normal distributions. Statistics 42(2), 179–194 (2008)

25. Hürlimann, W.: On likelihood ratio and stochastic order for skew-symmetric distributions with a common kernel. Int. J. Contemp. Math. Sci. 8(20), 957–967 (2013)

26. Jamali, D., Amiri, M., Jamalizadeh, A.: Comparison of the multivariate skew-normal random vectors based on the integral stochastic ordering. Commun. Stat.-Theory Methods 57, 1–13 (2020)

27. Jamali, D., Amiri, M., Jamalizadeh, A., Balakrishnan, N.: Integral stochastic ordering of the multivariate normal mean-variance and the skew-normal scale-shape mixture models. Stat., Optim. Inf. Comput. 8(1), 1–16 (2020)

28. Joe, H.: Multivariate concordance. J. Multiv. Anal. 35, 12–30 (1990)

29. Kim, H.-M., Genton, M.G.: Characteristic functions of scale mixtures of multivariate skew-normal distributions. J. Multiv. Anal. 102(7), 1105–1117 (2011)

30. Landsman, Z., Tsanakas, A.: Stochastic ordering of bivariate elliptical distributions. Stat. Probab. Lett. 76, 488–494 (2006)

31. Magnus, J.R., Neudecker, H.: Matrix differential calculus with applications in statistics and econometrics. Wiley, Hoboken, New Jersey (2019)

32. Müller, A.: Stochastic orders generated by integrals: a unified study. Adv. Appl. Probab. 35, 414–428 (1997)

33. Müller, A.: Stochastic ordering of multivariate normal distributions. Annal. Institut. Stat. Math. 53(3), 567–575 (2001)

34. Müller, A., Scarsini, M.: Some remarks on the supermodular order. J. Multiv. Anal. 73(1), 107–119 (2000)

35. Müller, A., Stoyan, D.: Comparison methods for stochastic models and risks. Wiley, New York (2002)

36. Ning, W., Gupta, A.K.: Matrix variate extended skew normal distributions. Random Op. Stoch. Eq. 20(4), 299–310 (2012)

37. Pan, X., Qiu, G., Hu, T.: Stochastic orderings for elliptical random vectors. J. Multiv. Anal. 148, 83–88 (2016)

38. Rezaei, A., Yousefzadeh, F., Arellano-Valle, R.B.: Scale and shape mixtures of matrix variate extended skew normal distributions. J. Multiv. Anal. 179, 104649 (2020)

39. Shaked, M., Shanthikumar, J.G.: Stochastic orders. Springer, New York (2007)

40. Shushi, T.: Generalized skew-elliptical distributions are closed under affine transformations. Stat. Probab. Lett. 134, 1–4 (2018)

41. Tong, Y.L.: Probability inequalities in multivariate distributions. Academic Press, Boston (2014)

42. Ye, R., Wang, T., Gupta, A.K.: Distribution of matrix quadratic forms under skew-normal settings. J. Multiv. Anal. 131, 229–239 (2014)

43. Yin, C.: Stochastic orderings of multivariate elliptical distributions. J. Appl. Probab. 58, 551–568 (2021)