PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 34, Fasc. 2 (2014), pp. 215–235

A UNIFIED WEIGHTED FAMILY OF DISTRIBUTIONS WHICH CONTAINS THE SKEW-ELLIPTICAL FAMILY

BY

M. T. ALODAT (DOHA), S. O. AL-SALMAN (IRBID),
AND M. Y. AL-RAWWASH (SHARJAH)

Abstract. In this paper, we introduce a new family of multivariate distributions, called the *unified weighted family*, as a generalization to the skew-elliptical family. We study some properties of the proposed family and show that it subsumes many important subfamilies such as the families arisen from the selection and hidden truncation ideas. Although the proposed family is very general, we focus on the multivariate weighted normal family which is regarded as a promising candidate in statistical inference.

2000 AMS Mathematics Subject Classification: Primary: 62E15; Secondary: 62H10.

Key words and phrases: Multivariate weighted normal distribution, unified weighted family of distributions, skew-elliptical family.

1. INTRODUCTION

Although the multivariate normal distribution has nice properties that make it popular in statistical modeling, yet it shows some deficiencies in skew data set-ups which should not be ignored. In fact, using Gaussian assumptions to model skewed data leads to unrealistic or nonsensical results (see Ganjali et al. [16]). The authors of [16] presented their concerns where such unrealistic results may occur while investigating the parameter estimates of a linear mixed model. Through the last two decades several articles and books introduced distributions to model skewed data. Azzalini [9] introduced the skewness concept to the normal distribution in terms of modifying the normal probability density function by a multiplicative skewing function. This idea has been applied to elliptical distributions to produce the so-called skew-elliptical distribution. Numerous articles discussed new classes of multivariate distributions having properties which coincide with those of normal class or are close to them. Examples of such classes include the skew-normal family proposed by Azzalini and Dalla Valle [13], the fundamental skew-normal distribution of Arellano-Valle and Genton [5], the unified skew-normal family studied by Arellano-Valle and Azzalini [3], the closed skew-normal distribution proposed by Domínguez-Molina et al. [15], the selection skewed distribution proposed by Arellano-Valle et al. [4], and the unified skew-elliptical class proposed by Arellano-Valle and Genton [7], among others. We propose a new family of distributions and focus on a subfamily called the *multivariate weighted normal distribution*. This article is organized as follows. In Section 2, we revise the skew-normal distribution of Azzalini [9] and its variations. In Section 3, we present a unified weighted family of distributions and discuss several special cases. In Section 4, we introduce a weighted normal distribution and discuss its properties and more special cases. Finally, we discuss our findings in Section 5, and we outline some detailed work on the theory of this article in the Appendix.

2. THE SKEW-NORMAL DISTRIBUTION AND ITS EXTENSIONS

Azzalini [9] introduced the univariate skew-normal distribution that proves to be significant in skewed data modeling. Azzalini and Capitanio [11] and Azzalini and Dalla Valle [13] extended this distribution to the following multivariate version:

$$f_{\mathbf{Y}}(\mathbf{y}) = 2\varphi_n(\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Sigma})\Phi\big(\boldsymbol{\alpha}'\boldsymbol{\omega}^{-1}(\mathbf{y}-\boldsymbol{\mu})\big), \quad \mathbf{y}\in\mathbb{R}^n, \tag{2.1}$$

where $\varphi_n(\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Sigma})$ is the multivariate normal probability density function (pdf) with mean $\boldsymbol{\mu}\in\mathbb{R}^n$ and $(n\times n)$ covariance matrix $\boldsymbol{\Sigma}$, $\Phi(\cdot)$ is the cumulative distribution function of standard univariate normal distribution, $\boldsymbol{\alpha}\in\mathbb{R}^n$ is a vector that controls the skewness, and $\boldsymbol{\omega}$ is a diagonal matrix formed by using the standard deviations of $\boldsymbol{\Sigma}$.

As a generalization of the normal law, the skew-normal distribution is considered a natural choice in all practical skewed data sets. A particularly valuable property is the continuity of the passage from the normal case to skewed distributions. From the theoretical point of view, the skew-normal (SN) class has the advantage of being mathematically tractable and having a good number of properties in common with the standard normal distribution. The SN class introduced by Azzalini and Dalla Valle [13] lacks the closure under both conditional distribution and convolution, which is a major pitfall of the class. Domínguez-Molina et al. [15] extended the SN class to a larger class, called the *multivariate closed skew-normal class*, which admits the closure under several properties. Azzalini and Capitanio [12] proposed a new set of non-symmetric densities which subsumes several special cases including skew-elliptical densities. Several articles discussed extensively the skew-normal distribution and recent modifications related to multivariate families. For more details, see [2]–[4], [10], and [14].

2.1. Closed skew-normal distribution. The closed skew-normal (CSN) family of distributions was proposed by Domínguez-Molina et al. [15] and reported in González-Farias et al. [17]. It shares several properties with normal distribution, including the closure under marginalization, and conditionally based on statistical modeling. Domínguez-Molina et al. [15] discussed the closure of CSN family under linear transformation and moment generating function. Moreover, it is closed under the sum of independent CSN random vectors and closed for the joint distribution of independent CSN distributions. The expressions of the CSN marginal and conditional densities are similar to their normal distribution counterparts. The pdf of CSN distribution is given by

$$g_{p,q}(\mathbf{x}) = C\varphi_p(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})\Phi_q\big(\mathbf{D}(\mathbf{x} - \boldsymbol{\mu}); \boldsymbol{\nu}, \boldsymbol{\Delta}\big), \quad \mathbf{x} \in \mathbb{R}^p, \tag{2.2}$$

where $C^{-1} = \Phi(\mathbf{0}; \boldsymbol{\nu}, \boldsymbol{\Delta} + \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^T)$, $p > 1$, $q > 1$, $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathbb{R}^q$, $\mathbf{D}$ is an arbitrary $(q \times p)$ matrix, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Delta}$ are positive definite matrices of dimensions $p \times p$ and $q \times q$, respectively. The CSN pdf defined in (2.2) is more general than the SN distribution considered in (2.1) and it might be used to fit non-normal data. The absence of analytical representations of $\Phi_q(\mathbf{0}; \boldsymbol{\nu}, \boldsymbol{\Delta})$ and its derivatives is a major challenge for practitioners from the inferential point of view. Arellano-Valle and Azzalini [3] discussed the CSN family and pointed out that this distribution is over-parameterized. Some extensions and redefinitions of this family are presented in Arellano-Valle and Azzalini [3] and Arellano-Valle and Genton [5].

2.2. Selection distributions. Let $\mathbf{U} \in \mathbb{R}^q$ and $\mathbf{V} \in \mathbb{R}^p$ be two random vectors and $\mathcal{C}$ be a measurable subset of $\mathbb{R}^q$. The random vector $\mathbf{X} \stackrel{d}{=} (\mathbf{V}|\mathbf{U} \in \mathcal{C})$ is said to have a *selection distribution* (see [5]), denoted by $\mathbf{X} \sim SLC_{p,q}$, if its pdf is given by

$$f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{V}}(\mathbf{x})\frac{P\{\mathbf{U} \in \mathcal{C}|\mathbf{V} = \mathbf{x}\}}{P\{\mathbf{U} \in \mathcal{C}\}} = \frac{\int_{\mathcal{C}} f_{\mathbf{U},\mathbf{V}}(\mathbf{u}, \mathbf{x})d\mathbf{u}}{\int_{\mathcal{C}} f_{\mathbf{U}}(\mathbf{u})d\mathbf{u}}, \tag{2.3}$$

where $f_{\mathbf{U}}(\cdot)$ is the marginal pdf of $\mathbf{U}$ and $f_{\mathbf{V},\mathbf{U}}(\cdot, \cdot)$ is the joint pdf of $\mathbf{V}$ and $\mathbf{U}$. Arellano-Valle et al. [4] showed that several distributions arise from (2.3) by varying the set $\mathcal{C}$ and they studied several examples when $(\mathbf{U}', \mathbf{V}')'$ is normally distributed. The most common specification of $\mathcal{C}$ is $\mathcal{C} = \{\mathbf{u} \in \mathbb{R}^q : \mathbf{u} > \mathbf{0}\}$, where the inequality between vectors is defined componentwisely. In such a case, the distribution corresponding to (2.3) is called a *fundamental skew-normal* (FUSN) distribution. Conditions under which invariant properties hold for the selection distribution class (SLC) are studied in Arellano-Valle and Genton [6]. If the distribution of $\mathbf{V}$ and the conditional distribution of $\mathbf{U}$ given $\mathbf{V}$ are both elliptical or both normal, then the resulting FUSN distribution is said to be *unified skew-elliptical* (SUE) or *unified skew-normal* (SUN), respectively. The main computational drawback of SLC is how to obtain $P\{\mathbf{U} \in \mathcal{C}\}$ for $q > 2$, especially with high dimensional data sets.

2.3. A unified skew-elliptical distribution. A continuous p-dimensional random vector $\mathbf{Y}$ has a multivariate unified skew-elliptical distribution, denoted by

$\mathbf{Y} \sim SUE_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Omega}, h^{(q+p)}, \boldsymbol{\tau}, \boldsymbol{\Gamma})$, if its pdf is given as

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{F_q(\boldsymbol{\tau}; \boldsymbol{\Gamma} + \boldsymbol{\Lambda}\bar{\boldsymbol{\Omega}}\boldsymbol{\Lambda}', h^{(q)})} f_p(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Omega}, h^{(p)}) F_q(\boldsymbol{\Lambda}\mathbf{z} + \boldsymbol{\tau}; \boldsymbol{\Gamma}, h^{(q)}_{Q(\mathbf{z})}),$$

where $f_p(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Omega}, h^{(p)}) = |\boldsymbol{\Omega}|^{-1/2} h^{(p)}\big(Q(\mathbf{z})\big)$ is the pdf of an elliptical distribution with parameters $\boldsymbol{\mu}$, $\boldsymbol{\Omega}$ and density generator function $h^{(p)}$, $F_q(\cdot, h^{(q)}_{Q(\mathbf{z})})$ is the cdf of q-dimensional elliptical distribution, $h^{(q)}_{Q(\mathbf{z})}(u) = h^{(p+q)}\big(u + Q(\mathbf{z})\big)/h^{(p)}\big(Q(\mathbf{z})\big)$, $\mathbf{z} = \omega^{-1}(\mathbf{y} - \boldsymbol{\mu})$, $Q(\mathbf{z}) = \mathbf{z}'\bar{\boldsymbol{\Omega}}^{-1}\mathbf{z}$, $\bar{\boldsymbol{\Omega}} = \omega^{-1}\boldsymbol{\Omega}\omega^{-1}$, $\omega = \mathrm{diag}(\boldsymbol{\Omega})^{1/2}$, $\boldsymbol{\Lambda}$ is a $(q \times p)$ real matrix controlling shape, $\boldsymbol{\tau} \in \mathbb{R}^q$ is an extension parameter, and $\boldsymbol{\Gamma}$ is a $(q \times q)$ positive definite correlation matrix. For more details, see [3] and [7].

3. A UNIFIED WEIGHTED FAMILY OF DISTRIBUTIONS

In this section, we define a unified weighted family of distributions and explore its probabilistic properties. Also, we consider some important subfamilies of distributions.

DEFINITION 3.1. Assume that an $(n \times 1)$ random vector $\mathbf{U}$ and an $(m \times 1)$ random vector $\mathbf{V}$ have the joint pdf $f_{\mathbf{U},\mathbf{V}}(\mathbf{u}, \mathbf{v})$ and let $g : \mathbb{R}^n \to \mathbb{R}^+$ be an integrable and measurable function. Also, assume that $\mathcal{C}$ is a measurable subset of $\mathbb{R}^n$ such that $\{\mathbf{u} \in \mathcal{C} : g(\mathbf{u}) > 0\}$ has a positive probability. We define the *unified weighted family of distributions* as those with the pdf

$$h(\mathbf{v}) = \frac{\int_{\mathcal{C}} g(\mathbf{u}) f_{\mathbf{U},\mathbf{V}}(\mathbf{u}, \mathbf{v}) d\mathbf{u}}{\int_{\mathcal{C}} g(\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}}, \quad \mathbf{v} \in \mathbb{R}^m, \tag{3.1}$$

where $f_{\mathbf{U}}(\mathbf{u})$ is the marginal density of $\mathbf{U}$.

The pdf (3.1) defines a rich class of distributions that can sometimes be obtained by varying the function $g(\mathbf{x})$ and keeping the joint pdf $f_{\mathbf{U},\mathbf{V}}(\mathbf{u}, \mathbf{v})$ fixed. On the other hand, certain members of this class can be obtained by keeping $g(\mathbf{x})$ fixed and varying the joint pdf $f_{\mathbf{U},\mathbf{V}}(\mathbf{u}, \mathbf{v})$. In the proposed model, the variable of integration $\mathbf{u}$ is a dummy variable and we may condition on the variable $\mathbf{V} = \mathbf{v}$ instead of $\mathbf{U}$. If we choose the dimension of $\mathbf{u}$ to be small, it allows us to reduce the difficulties in estimating the parameters of $h(\mathbf{v})$. The class of unified weighted distributions defined in this paper subsumes the weighted multivariate elliptically distributed families introduced by Kim [19] and [20]. Although most distributions represented by (3.1) are asymmetric, we could be interested in situations where $f_{\mathbf{V}}(\mathbf{v})$ is symmetric. In such situations the pdf (3.1) remains asymmetric unless the following condition is satisfied for all $\mathbf{v}$:

$$\int_{\mathcal{C}} g(\mathbf{u}) f_{\mathbf{U},\mathbf{V}}(\mathbf{u}, \mathbf{v}) d\mathbf{u} = \int_{\mathcal{C}} g(\mathbf{u}) f_{\mathbf{U},\mathbf{V}}(\mathbf{u}, -\mathbf{v}) d\mathbf{u}.$$

Simulating a random sample from the pdf (3.1) is very important in practice. To achieve that, we first rewrite $h(\mathbf{v})$ as

$$h(\mathbf{v}) = \int_{\mathbb{R}^n} \chi(\mathbf{u}) f_{\mathbf{V}|\mathbf{U}=\mathbf{u}}(\mathbf{v}|\mathbf{u}) d\mathbf{u},$$

where

$$\chi(\mathbf{u}) = \frac{g^*(\mathbf{u}) f_{\mathbf{U}}(\mathbf{u})}{\int_{\mathbb{R}^n} g^*(\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}} \quad \text{with } g^*(\mathbf{u}) = g(\mathbf{u}) \mathbf{1}_{\mathcal{C}}(\mathbf{u}).$$

Then, the following algorithm can be implemented to draw a random observation from $h(\mathbf{v})$:

1. Simulate a random observation, say, $\mathbf{U}_0$, from $\chi(\mathbf{u})$.
2. Simulate a random observation, say, $\mathbf{V}_0$, from $f_{\mathbf{V}|\mathbf{U}=\mathbf{u}}(\mathbf{v}|\mathbf{u})$.
3. Deliver $\mathbf{V}_0$ as an observation from $h(\mathbf{v})$.

The simulations in Steps 1 and 2 depend on the functional forms $\chi(\mathbf{u})$ and $f_{\mathbf{V}|\mathbf{U}=\mathbf{u}}(\mathbf{v}|\mathbf{u})$. The accept-reject method or the Metropolis–Hastings algorithms are very general algorithms that can be used for such purposes.

THEOREM 3.1. *Suppose that the pdf $f_{\mathbf{U},\mathbf{V}}(\mathbf{u}, \mathbf{v})$ in equation* (3.1) *obeys the closure under both marginal and conditional distributions. Then $h(\mathbf{v})$ satisfies the same properties.*

The proof is outlined in the Appendix.

Some interesting cases of (3.1) can be obtained by choosing $g(\mathbf{x})$ and $\mathcal{C}$ so that the integration in (3.1) has a closed form. Given an $(n \times 1)$ random vector $\mathbf{U}$ and an $(n \times n)$ matrix $\mathbf{A}$ such that $\boldsymbol{\delta} = E(\mathbf{U})$ and $\boldsymbol{\Upsilon} = \operatorname{cov}(\mathbf{U})$, we may get the following theorem.

THEOREM 3.2. *Let $\mathbf{U}$ be an $(n\times 1)$ random vector with finite second moments, and $\mathbf{A}$ be an $(n \times n)$ symmetric and semipositive definite matrix. If $g(\mathbf{u}) = \mathbf{u}'\mathbf{A}\mathbf{u}$ and $\mathcal{C} = \mathbb{R}^n$, then* (3.1) *reduces to*

$$h(\mathbf{v}) = f_{\mathbf{V}}(\mathbf{v}) \frac{E(\mathbf{U}'\mathbf{A}\mathbf{U}|\mathbf{V} = \mathbf{v})}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Upsilon}) + \boldsymbol{\delta}'\mathbf{A}\boldsymbol{\delta}}, \quad \mathbf{v} \in \mathbb{R}^m.$$

In fact, the symmetry of $f_{\mathbf{V}}(\mathbf{v})$ does not guarantee the symmetry of the last family while if $E(\mathbf{U}'\mathbf{A}\mathbf{U}|\mathbf{V} = \mathbf{v}) = E(\mathbf{U}'\mathbf{A}\mathbf{U}|\mathbf{V} = -\mathbf{v})$ for all $\mathbf{v}$ and $f_{\mathbf{V}}(\mathbf{v})$, then the symmetry is guaranteed.

3.1. Special cases. To show the significant role of the pdf (3.1), we present two special cases of the pdf (3.1). The first example involves the crossing theory of random processes and fields discussed by Podgórski and Rychlik [22]. Let $V(\mathbf{t})$ and $X(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^k$, be two stationary stochastic processes taking values in $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. Also, assume that the process $X(\mathbf{t})$ is differentiable with matrix of partial derivatives denoted by $\ddot{\mathbf{X}}(\mathbf{0})$. If the random variables $V(\mathbf{0})$, $X(\mathbf{0})$, and

$\ddot{\mathbf{X}}(\mathbf{0})$ admit a joint pdf denoted by $f_{V(\mathbf{0}),\ddot{\mathbf{X}}(\mathbf{0}),X(\mathbf{0})}(\mathbf{v},\ddot{\mathbf{x}},\mathbf{x})$, then the distribution of $V(\mathbf{t})$ on the contour $\mathcal{C}_{\mathbf{x}_0}=\{\mathbf{t}\in[0,1]^k : X(\mathbf{t})=\mathbf{x}_0\}$ will be written as

$$P\big(X(\mathbf{t})\in A|\mathbf{t}\in\mathcal{C}_{\mathbf{x}_0}\big)=\frac{\iint\limits_A \det g\big(\ddot{\mathbf{x}}(\mathbf{0})\big)f_{V(\mathbf{0}),\ddot{\mathbf{X}}(\mathbf{0}),X(\mathbf{0})}\big(\mathbf{v},\ddot{\mathbf{x}}(\mathbf{0}),\mathbf{x}_0\big)d\ddot{\mathbf{x}}(\mathbf{0})d\mathbf{v}}{\int \det g\big(\ddot{\mathbf{x}}(\mathbf{0})\big)f_{V(\mathbf{0}),\ddot{\mathbf{X}}(\mathbf{0}),X(\mathbf{0})}\big(\mathbf{v},\ddot{\mathbf{x}}(\mathbf{0}),\mathbf{x}_0\big)d\ddot{\mathbf{x}}(\mathbf{0})},$$

where $\det g(\cdot)$ means the generalized determinant function, i.e. $\det g(\mathbf{B})=\sqrt{|\mathbf{BB}'|}$.

Setting $A=(-\infty,x_1]\times\ldots\times(-\infty,x_n]$ in the last formula and differentiating with respect to $x_1,\ldots,x_n$, we get the pdf (3.1) provided that $\mathcal{C}=\mathbb{R}^m$, $g(\mathbf{x})=\det\big(\ddot{\mathbf{x}}(\mathbf{0})\big)$, and $f_{\mathbf{U},\mathbf{V}}(\mathbf{u},\mathbf{v})=f_{V(\mathbf{0}),\ddot{\mathbf{X}}(\mathbf{0})}\big(\mathbf{v},\ddot{\mathbf{x}}|X(\mathbf{0})=\mathbf{x}_0\big)$.

The second example is from Adler [1], where $X(\mathbf{t})$, $\mathbf{t}\in\mathbb{R}^n$, is assumed to be a stationary and ergodic random field which admits partial derivative up to the second order. Then, for any set of points $\mathbf{t}_1,\ldots,\mathbf{t}_k$, the conditional pdf of $\mathbf{V}=\big(X(\mathbf{t}_1),\ldots,X(\mathbf{t}_k)\big)'$ under the condition that $X(\mathbf{t})$ has a local maximum of height u at $\mathbf{t}=\mathbf{0}$ is given by

$$f_{\mathbf{V}}(\mathbf{v})=\frac{\int\limits_D \big|\det\big(\ddot{\mathbf{x}}(\mathbf{0})\big)\big|f_{V(\mathbf{0}),\ddot{\mathbf{X}}(\mathbf{0})}\big(\mathbf{v},\ddot{\mathbf{x}}|X(\mathbf{0})=u,\dot{X}(\mathbf{0})=\mathbf{0}\big)d\ddot{\mathbf{x}}(\mathbf{0})d\mathbf{v}}{\int\limits_D \big|\det\big(\ddot{\mathbf{x}}(\mathbf{0})\big)\big|f_{\ddot{\mathbf{X}}(\mathbf{0})}\big(\ddot{\mathbf{x}}|X(\mathbf{0})=u,\dot{X}(\mathbf{0})=\mathbf{0}\big)d\ddot{\mathbf{x}}(\mathbf{0})},$$

where $\dot{\mathbf{x}}$ is the gradient of $X(\mathbf{t})$ and D is the set of all $\mathbf{t}$ such that $\ddot{\mathbf{X}}(\mathbf{t})$ is a negative definite matrix. It can be noticed that the last pdf has the same form as (3.1).

To elaborate more on Theorem 3.2, we consider more special cases of $g(\mathbf{u})$ and assume that $\mathbf{U}$ and $\mathbf{V}$ have jointly a multivariate normal distribution, which leads us to the following cases (i)–(vii):

(i) If $g(\mathbf{u})=\mathbf{1}_{\mathcal{C}}(\mathbf{u})$, then (3.1) reduces to (2.3).

(ii) Let $\mathbf{A}_1$ and $\mathbf{A}_2$ be $(n\times n)$ semipositive definite matrices. Define $g(\mathbf{u})=\mathbf{u}'\mathbf{A}_1\mathbf{u}\times\mathbf{u}'\mathbf{A}_2\mathbf{u}$ such that $\mathbf{U}$ and $\mathbf{V}$ in (3.1) have jointly the following multivariate normal distribution:

$$\begin{pmatrix}\mathbf{U}\\ \mathbf{V}\end{pmatrix}\sim N_{n+m}\left(\begin{pmatrix}\boldsymbol{S}\\ \boldsymbol{\ell}\end{pmatrix},\begin{pmatrix}\mathbf{M} & \boldsymbol{\Gamma}^0\\ (\boldsymbol{\Gamma}^0)' & \mathbf{K}\end{pmatrix}\right).$$

Using Theorem 3.2 of Mathai and Provost [21], we may rewrite (3.1) as

$$h(\mathbf{v})=\frac{\varphi_m(\mathbf{v};\boldsymbol{\ell},\mathbf{K})\int\limits_{\mathbb{R}^n}\mathbf{u}'\mathbf{A}_1\mathbf{u}\times\mathbf{u}'\mathbf{A}_2\mathbf{u}\,\varphi_n\big(\mathbf{u};\boldsymbol{\tau}^0(\mathbf{v}),\mathbf{T}\big)d\mathbf{u}}{\int\limits_{\mathbb{R}^n}\mathbf{u}'\mathbf{A}_1\mathbf{u}\times\mathbf{u}'\mathbf{A}_2\mathbf{u}\,\varphi_n(\mathbf{u};\boldsymbol{S},\mathbf{M})d\mathbf{u}},\qquad \mathbf{v}\in\mathbb{R}^m,$$

where $\boldsymbol{\tau}^0(\mathbf{v})=\boldsymbol{S}+\boldsymbol{\Gamma}^0\mathbf{K}^{-1}(\mathbf{v}-\boldsymbol{\ell})$ and $\mathbf{T}=\mathbf{M}-\boldsymbol{\Gamma}^0\mathbf{K}^{-1}(\boldsymbol{\Gamma}^0)'$. Since the function $\varphi_n\big(\mathbf{u};\boldsymbol{\tau}^0(\mathbf{v}),\mathbf{T}\big)$ is the pdf of $N_n\big(\boldsymbol{\tau}^0(\mathbf{v}),\mathbf{T}\big)$, the last equation reduces to

$$h(\mathbf{v})=\varphi_m(\mathbf{v};\boldsymbol{\ell},\mathbf{T})\frac{E\left(\mathbf{U}'\mathbf{A}_1\mathbf{U}\times\mathbf{U}'\mathbf{A}_2\mathbf{U}|\mathbf{V}=\mathbf{v}_0\right)}{E\left(\mathbf{U}'\mathbf{A}_1\mathbf{U}\times\mathbf{U}'\mathbf{A}_2\mathbf{U}\right)}.$$

Applying Theorem 3.2 d3 in Mathai and Provost [21], p. 75, we get

$$h(\mathbf{v}) = \varphi_m(\mathbf{v}; \boldsymbol{\ell}, \mathbf{T}) \frac{k_1 + k_2}{k_3 + k_4},$$

where

$$\begin{aligned}
k_1 &= 2\mathrm{tr}(\mathbf{A}_1\mathbf{T}\mathbf{A}_2\mathbf{T}) + 4(\boldsymbol{\tau}^0)'(\mathbf{v})\mathbf{A}_1\mathbf{T}\mathbf{A}_2\boldsymbol{\tau}^0(\mathbf{v}),\\
k_2 &= \big((\boldsymbol{\tau}^0)'(\mathbf{v})\mathbf{A}_1\boldsymbol{\tau}^0(\mathbf{v}) + \mathrm{tr}(\mathbf{A}_1\mathbf{T})\big)\big((\boldsymbol{\tau}^0)'(\mathbf{v})\mathbf{A}_2\boldsymbol{\tau}^0(\mathbf{v}) + \mathrm{tr}(\mathbf{A}_2\mathbf{T})\big),\\
k_3 &= 2\mathrm{tr}(\mathbf{A}_1\mathbf{M}\mathbf{A}_2\mathbf{M}) + 4\boldsymbol{S}'\mathbf{A}_1\mathbf{M}\mathbf{A}_2\boldsymbol{S},\\
k_4 &= \big(\boldsymbol{S}'\mathbf{A}_1\boldsymbol{S} + \mathrm{tr}(\mathbf{A}_1\mathbf{M})\big)\big(\boldsymbol{S}'\mathbf{A}_2\boldsymbol{S} + \mathrm{tr}(\mathbf{A}_2\mathbf{M})\big).
\end{aligned}$$

Another special case assumes $g(\mathbf{u}) = (\mathbf{u}'\mathbf{A}\mathbf{u})^r$, where r is an integer, and the closed form is obtained by making use of Theorems 3.2 b2 and 3.2 b4 of Mathai and Provost [21].

(iii) Let $g(\mathbf{u}) = \mathbf{a}'\mathbf{u}$, where $\mathbf{a}$ is a fixed vector in $\mathbb{R}^n$, and let $\mathcal{C} = \{\mathbf{u} \in \mathbb{R}^n : \mathbf{a}'\mathbf{u} > 0\}$. If $\binom{\mathbf{U}}{\mathbf{V}}$ has the same distribution as in (ii), then $(\mathbf{a}'\mathbf{U}|\mathbf{V} = \mathbf{v}) \sim N(\ddot{\mu}, \ddot{\sigma}^2)$ and $\mathbf{a}'\mathbf{U} \sim N(\mu, \sigma^2)$, where $\ddot{\mu} = E(\mathbf{a}'\mathbf{U}|\mathbf{V} = \mathbf{v})$, $\ddot{\sigma}^2 = \mathrm{var}(\mathbf{a}'\mathbf{U}|\mathbf{V} = \mathbf{v})$, $\mu = \mathbf{a}'\boldsymbol{S}$, and $\sigma^2 = \mathbf{a}'\mathbf{M}\mathbf{a}$. Finally, by putting $x^+ = \max(0, x)$, the pdf (3.1) will reduce to

$$\begin{aligned}
h(\mathbf{v}) &= \varphi_m(\mathbf{v}; \boldsymbol{\ell}, \mathbf{K}) \frac{E\big((\mathbf{a}'\mathbf{U})^+|\mathbf{V} = \mathbf{v}\big)}{E\big((\mathbf{a}'\mathbf{U})^+\big)}\\
&= \varphi_m(\mathbf{v}; \boldsymbol{\ell}, \mathbf{K}) \frac{\ddot{\mu}\Phi(\ddot{\mu}/\ddot{\sigma}) + (\ddot{\sigma}/\sqrt{2\pi})\exp(-\mu^2/\sigma^2)}{\mu\Phi(\mu/\sigma) + (\sigma/\sqrt{2\pi})\exp(-\mu^2/\sigma^2)}.
\end{aligned}$$

(iv) Let $g(\mathbf{u}) = \Phi(a + \mathbf{b}'\mathbf{u})$ and $\mathcal{C} = \mathbb{R}^n$, where $a \in \mathbb{R}$ and $\mathbf{b}$ is a fixed vector in $\mathbb{R}^n$. Then

$$\begin{aligned}
h(\mathbf{v}) &= \frac{\int_{\mathcal{C}} \Phi(a + \mathbf{b}'\mathbf{u}) f_{\mathbf{U},\mathbf{V}}(\mathbf{u}, \mathbf{v}) d\mathbf{u}}{\int_{\mathcal{C}} \Phi(a + \mathbf{b}'\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}}\\
&= \varphi_m(\mathbf{v}; \boldsymbol{\ell}, \mathbf{K}) \frac{\int_{\mathbb{R}^n} \Phi(a + \mathbf{b}'\mathbf{u})\varphi_n\big(\mathbf{u}; \boldsymbol{\tau}^0(\mathbf{v}), \mathbf{T}\big) d\mathbf{u}}{\int_{\mathbb{R}^n} \Phi(a + \mathbf{b}'\mathbf{u})\varphi_n(\mathbf{u}; \boldsymbol{S}, \mathbf{M}) d\mathbf{u}}\\
&= \varphi_m(\mathbf{v}; \boldsymbol{\ell}, \mathbf{K}) \frac{E\big(\Phi(a + \mathbf{b}'\mathbf{U})|\mathbf{V} = \mathbf{v}\big)}{E\big(\Phi(a + \mathbf{b}'\mathbf{U})\big)},
\end{aligned}$$

where $\boldsymbol{\tau}^0(\mathbf{v})$ and $\mathbf{T}$ are defined in (ii). Referring to Raluca [23] and using Lemma 2.1 in Arellano-Valle and Genton [5] to simplify the conditional expectation, we get the following closed form:

$$h(\mathbf{v}) = \varphi_m(\mathbf{v}; \boldsymbol{\ell}, \mathbf{K}) \frac{\Phi\big((a + \mathbf{b}'\boldsymbol{\tau}^0(\mathbf{v}))/\sqrt{1 + \mathbf{b}'\mathbf{T}\mathbf{b}}\big)}{\Phi\big((a + \mathbf{b}'\boldsymbol{S})/\sqrt{1 + \mathbf{b}'\mathbf{M}\mathbf{b}}\big)}.$$

Selecting the function $g(\mathbf{u})$ in this manner leads to the so-called extended skew-normal (ESN) distribution. In fact, according to Arnold and Beaver [8] and Rodrigues [24], a random vector $\mathbf{X}$ is said to have an *n-dimensional ESN distribution* with parameters $\boldsymbol{\mu}$, $\boldsymbol{\Omega}$, $\boldsymbol{\lambda}$ and τ, denoted by $\mathbf{X} \sim ESN_n(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\lambda}, \tau)$, if it has the pdf

$$f(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\lambda}, \tau) = \varphi_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Omega})\Phi\big(\alpha_0 + \boldsymbol{\lambda}'(\mathbf{x} - \boldsymbol{\mu})\big)/\Phi(\tau),$$

where $\boldsymbol{\lambda} = \boldsymbol{\Omega}^{-1}\boldsymbol{\delta}(1 - \boldsymbol{\delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\delta})^{-1/2}$, $\tau = a$, and $\alpha_0 = \tau(1 - \boldsymbol{\delta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\delta})^{-1/2}$. Setting $\boldsymbol{S} = \mathbf{0}$, $\boldsymbol{\ell} = \boldsymbol{\mu}$, $\boldsymbol{\Gamma}^0 = \boldsymbol{\delta}$, $\mathbf{M} = \boldsymbol{\Omega}$, and $\mathbf{K} = 1$, and noting that $\mathbf{b}'\mathbf{T}\mathbf{b} = \mathbf{b}'\mathbf{b} - \mathbf{b}'\boldsymbol{\delta}\boldsymbol{\Omega}^{-1}\boldsymbol{\delta}'\mathbf{b}$, we see that the pdf $h(\mathbf{v})$ reduces to the ESN distribution with

$$\alpha_0 = \frac{\tau}{\sqrt{1 + \mathbf{b}'\mathbf{b}}}\left(1 - \frac{\mathbf{b}'\boldsymbol{\delta}\boldsymbol{\Omega}^{-1}\boldsymbol{\delta}'\mathbf{b}}{1 + \mathbf{b}'\mathbf{b}}\right)^{-1/2}$$

and

$$\boldsymbol{\lambda} = \frac{\boldsymbol{\Omega}^{-1}\mathbf{b}'\boldsymbol{\delta}}{\sqrt{1 + \mathbf{b}'\mathbf{b}}}\left(1 - \frac{\mathbf{b}'\boldsymbol{\delta}\boldsymbol{\Omega}^{-1}\boldsymbol{\delta}'\mathbf{b}}{1 + \mathbf{b}'\mathbf{b}}\right)^{-1/2},$$

where $\boldsymbol{\delta}^* = \mathbf{b}'\boldsymbol{\delta}/\sqrt{1 + \mathbf{b}'\mathbf{b}}$.

(v) Let $g(\mathbf{u}) = \exp(\mathbf{c}'\mathbf{u})\Phi(a + \mathbf{b}'\mathbf{u})$, $\mathbf{u} \in \mathbb{R}^n$, $a \in \mathbb{R}$. Then the pdf (3.1) has the following closed form:

$$\begin{aligned} h(\mathbf{v}) = \varphi_m(\mathbf{v}; \boldsymbol{\ell}, \mathbf{K})&\left[\exp\left(\mathbf{c}'\boldsymbol{\tau}^0(\mathbf{v}) + \mathbf{c}'\mathbf{T}\mathbf{c}\right)\Phi\left(\frac{a + \mathbf{b}'\boldsymbol{\tau}^0(\mathbf{v}) + \mathbf{b}'\mathbf{T}\mathbf{c}}{\sqrt{1 + \mathbf{b}'\mathbf{T}\mathbf{b}}}\right)\right] \\ &\times \left[\exp(\mathbf{c}'\boldsymbol{S} + \mathbf{c}'\mathbf{M}\mathbf{c})\Phi\left(\frac{a + \mathbf{b}'\boldsymbol{S} + \mathbf{b}'\mathbf{M}\mathbf{c}}{\sqrt{1 + \mathbf{b}'\mathbf{M}\mathbf{b}}}\right)\right]^{-1}. \end{aligned}$$

(vi) Let $p^{(j)} > 0$, $\lambda_0^{(j)}$ be real numbers and $\boldsymbol{\lambda}^{(j)} \in \mathbb{R}^n$ for $j = 1, \ldots, k$. Define $g(\mathbf{u})$ as

$$g(\mathbf{u}) = \sum_{j=1}^{k} p^{(j)}\Phi(\lambda_0^{(j)} + \boldsymbol{\lambda}^{(j)\prime}\mathbf{u})$$

and let $\mathcal{C} = \mathbb{R}^n$. Then (3.1) reduces to

$$h(\mathbf{v}) = \frac{\int_{\mathcal{C}} \sum_{j=1}^{k} p^{(j)}\Phi(\lambda_0^{(j)} + \boldsymbol{\lambda}^{(j)\prime}\mathbf{u}) f_{\mathbf{U},\mathbf{V}}(\mathbf{u}, \mathbf{v}) d\mathbf{u}}{\int_{\mathcal{C}} \sum_{j=1}^{k} p^{(j)}\Phi(\lambda_0^{(j)} + \boldsymbol{\lambda}^{(j)\prime}\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}}.$$

Using Fubini's theorem, we may write this pdf as follows:

$$
\begin{aligned}
h(\mathbf{v}) &= \varphi_m(\mathbf{v};\boldsymbol{\ell},\mathbf{K})\frac{\sum_{j=1}^{k} p^{(j)} \int_{\mathcal{C}} \Phi(\lambda_0^{(j)} + \boldsymbol{\lambda}^{(j)\prime}\mathbf{u})\varphi_n\left(\mathbf{u};\boldsymbol{\tau}^0(\mathbf{v}),\mathbf{T}\right) d\mathbf{u}}{\sum_{j=1}^{k} p^{(j)} \int_{\mathcal{C}} \Phi(\lambda_0^{(j)} + \boldsymbol{\lambda}^{(j)\prime}\mathbf{u})\varphi_n\left(\mathbf{u};\boldsymbol{S},\mathbf{M}\right) d\mathbf{u}} \\
&= \varphi_m(\mathbf{v};\boldsymbol{\ell},\mathbf{K})\frac{\sum_{j=1}^{k} p^{(j)} E\Phi(\lambda_0^{(j)} + \boldsymbol{\lambda}^{(j)\prime}\mathbf{U}|\mathbf{V}=\mathbf{v})}{\sum_{j=1}^{k} p^{(j)} E\Phi(\lambda_0^{(j)} + \boldsymbol{\lambda}^{(j)\prime}\mathbf{U})} \\
&= \varphi_m(\mathbf{v};\boldsymbol{\ell},\mathbf{K})\frac{\sum_{j=1}^{k} p^{(j)} \Phi\big(\big(\lambda_0^{(j)} + \boldsymbol{\lambda}^{(j)\prime}\boldsymbol{\tau}^0(\mathbf{v})\big)/\sqrt{1+\boldsymbol{\lambda}^{(j)\prime}\mathbf{T}\boldsymbol{\lambda}^{(j)}}\big)}{\sum_{j=1}^{k} p^{(j)} \Phi\big((\lambda_0^{(j)} + \boldsymbol{\lambda}^{(j)\prime}\boldsymbol{S})/\sqrt{1+\boldsymbol{\lambda}^{(j)\prime}\mathbf{M}\boldsymbol{\lambda}^{(j)}}\big)}.
\end{aligned}
$$

Setting $\boldsymbol{\lambda}^{(j)} = e_j$ and $\lambda_0^{(j)} = 0$ for all j gives the following pdf:

$$
h(\mathbf{v}) = \varphi_m(\mathbf{v};\boldsymbol{\ell},\mathbf{K})\frac{\sum_{j=1}^{k} p^{(j)} \Phi\big(\big(\mathbf{e}_j'\boldsymbol{\tau}^0(\mathbf{v})\big)/\sqrt{1+\boldsymbol{\lambda}^{(j)\prime}\mathbf{T}\boldsymbol{\lambda}^{(j)}}\big)}{\sum_{j=1}^{k} p^{(j)} \Phi\big((\mathbf{e}_j'\boldsymbol{S})/\sqrt{1+\boldsymbol{\lambda}^{(j)\prime}\mathbf{M}\boldsymbol{\lambda}^{(j)}}\big)}.
$$

(vii) Let $g(\mathbf{u}) = \exp(s\mathbf{u}'\mathbf{A}\mathbf{u})$, where $s \in \mathbb{R}$ and $\mathbf{A}$ is a symmetric and semi-positive definite matrix. Then

$$
\begin{aligned}
h(\mathbf{v}) &= \frac{\int_{\mathcal{C}} \exp(s\mathbf{u}'\mathbf{A}\mathbf{u}) f_{\mathbf{U},\mathbf{V}}(\mathbf{u},\mathbf{v}) d\mathbf{u}}{\int_{\mathcal{C}} \exp(s\mathbf{u}'\mathbf{A}\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}} \\
&= \varphi_m(\mathbf{v};\boldsymbol{\ell},\mathbf{K})\frac{\int_{\mathbb{R}^n} \exp(s\mathbf{u}'\mathbf{A}\mathbf{u})\varphi_n\left(\mathbf{u};\boldsymbol{\tau}^0(\mathbf{v}),\mathbf{T}\right) d\mathbf{u}}{\int_{\mathbb{R}^n} \exp(s\mathbf{u}'\mathbf{A}\mathbf{u})\varphi_n\left(\mathbf{u};\boldsymbol{S},\mathbf{M}\right) d\mathbf{u}}.
\end{aligned}
$$

Since $\varphi_n\big(\mathbf{u};\boldsymbol{\tau}^0(\mathbf{v}),\mathbf{T}\big)$ is the pdf of $N_n\big(\boldsymbol{\tau}^0(\mathbf{v}),\mathbf{T}\big)$, the pdf $h(\mathbf{v})$ turns out to be

$$
h(\mathbf{v}) = \varphi_m(\mathbf{v};\boldsymbol{\ell},\mathbf{K})\frac{E\big(\exp(s\mathbf{u}'\mathbf{A}\mathbf{u})|\mathbf{V}=\mathbf{v}\big)}{\big(\exp(s\mathbf{u}'\mathbf{A}\mathbf{u})\big)d\mathbf{u}} = \varphi_m(\mathbf{v};\boldsymbol{\ell},\mathbf{K})\frac{M_{Q^0}(s)}{M_Q(s)},
$$

where $M_{Q^0}(s)$ is the moment generating function (mgf) of $\mathbf{U}'\mathbf{A}\mathbf{U}$ given $\mathbf{V} = \mathbf{v}$, and $M_Q(s)$ is the mgf of $\mathbf{U}'\mathbf{A}\mathbf{U}$.

4. WEIGHTED NORMAL DISTRIBUTION

Throughout the rest of the paper, we assume that $f_{\mathbf{U},\mathbf{V}}(\mathbf{u},\mathbf{v})$ in equation (3.1) is a member of the multivariate normal class. To introduce the weighted normal (WN) distribution, let g be the function described in Definition 3.1 and $\varphi_{n+m}(\mathbf{x},\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Sigma})$ be the pdf of $N_{n+m}(\boldsymbol{\mu},\boldsymbol{\Sigma})$, $m>1, n>1$, $\boldsymbol{\mu}\in\mathbb{R}^{n+m}$, $\boldsymbol{\Sigma}$ is an $(n+m)\times(n+m)$ positive definite matrix such that

$$\boldsymbol{\mu}=\begin{pmatrix}\boldsymbol{\mu}_{\mathbf{x}}\\ \boldsymbol{\mu}_{\mathbf{y}}\end{pmatrix}=\begin{pmatrix}\boldsymbol{\theta}\\ \gamma\end{pmatrix}\quad\text{and}\quad\boldsymbol{\Sigma}=\begin{pmatrix}\boldsymbol{\Sigma}_{\mathbf{xx}} & \boldsymbol{\Sigma}_{\mathbf{xy}}\\ \boldsymbol{\Sigma}_{\mathbf{yx}} & \boldsymbol{\Sigma}_{\mathbf{yy}}\end{pmatrix}=\begin{pmatrix}\boldsymbol{\Psi} & \boldsymbol{\Gamma}\\ \boldsymbol{\Gamma}' & \boldsymbol{\Delta}\end{pmatrix}.$$

DEFINITION 4.1. A random variable $\mathbf{Y}$ is said to have a *WN distribution*, denoted by $WN_{m,n}(\boldsymbol{\mu},\boldsymbol{\Sigma},g)$, if the pdf of $\mathbf{Y}$ is given by

$$f_{\mathbf{Y}}(\mathbf{y})=\frac{\int_{\mathcal{C}} g(\mathbf{x})\varphi_{n+m}\left(\mathbf{x},\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Sigma}\right)d\mathbf{x}}{\int_{\mathcal{C}} g(\mathbf{x})\varphi_{n}\left(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi}\right)d\mathbf{x}},\quad \mathbf{y}\in\mathbb{R}^m, \tag{4.1}$$

where $\varphi_n(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})$ is the marginal pdf of $\mathbf{X}$.

Also, we can write the pdf of a WN random vector as

$$f_{\mathbf{Y}}(\mathbf{y})=\frac{\int_{\mathcal{C}} g(\mathbf{x})\varphi_{n}(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})\varphi_m\big(\mathbf{y};\boldsymbol{\mu}^*(\mathbf{x}),\boldsymbol{\Sigma}^*\big)d\mathbf{x}}{\int_{\mathcal{C}} g(\mathbf{x})\varphi_{n}(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})d\mathbf{x}}, \tag{4.2}$$

where $\boldsymbol{\mu}^*(\mathbf{x})=\boldsymbol{\gamma}+\boldsymbol{\Gamma}'\boldsymbol{\Psi}^{-1}(\mathbf{x}-\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}^*=\boldsymbol{\Delta}-\boldsymbol{\Gamma}'\boldsymbol{\Psi}^{-1}\boldsymbol{\Gamma}$.

4.1. Main results.

PROPOSITION 4.1. *The cdf associated with the pdf given in* (4.2) *is*

$$F_{\mathbf{Y}}(\mathbf{y})=\frac{\int_{\mathcal{C}} g(\mathbf{x})\varphi_{n}(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})\Phi_m\big(\mathbf{y};\boldsymbol{\mu}^*(\mathbf{x}),\boldsymbol{\Sigma}^*\big)d\mathbf{x}}{\int_{\mathcal{C}} g(\mathbf{x})\varphi_{n}(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})d\mathbf{x}},$$

where $\Phi_m\big(\mathbf{y};\boldsymbol{\mu}^*(\mathbf{x}),\boldsymbol{\Sigma}^*\big)$ *is the cdf of an* m*-dimensional normal distribution with mean vector* $\boldsymbol{\mu}^*(\mathbf{x})$ *and covariance matrix* $\boldsymbol{\Sigma}^*$*.*

The proofs of some theorems and propositions will be outlined in Appendix.

PROPOSITION 4.2. *If* $\mathbf{Y}\sim WN_{m,n}(\boldsymbol{\mu},\boldsymbol{\Sigma},g)$, *then the mgf of* $\mathbf{Y}$ *is*

$$M_{\mathbf{Y}}(\mathbf{u})=\frac{\int_{\mathcal{C}} g(\mathbf{x})\varphi_{n}(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})M_{\mathbf{Y}|\mathbf{x}}(\mathbf{u})d\mathbf{x}}{\int_{\mathcal{C}} g(\mathbf{x})\varphi_{n}(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})d\mathbf{x}},\quad \mathbf{u}\in\mathbb{R}^m, \tag{4.3}$$

where $M_{\mathbf{Y}|\mathbf{x}}(\mathbf{u})=\exp\left(\mathbf{u}'\boldsymbol{\mu}^*(\mathbf{x})+\frac{1}{2}\mathbf{u}'\boldsymbol{\Sigma}^*\mathbf{u}\right)$.

PROPOSITION 4.3. *Suppose that* $\mathbf{Y} \sim WN_{m,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A}$ *is an* $(m \times m)$ *matrix of rank* p, *and*

$$\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{I}_n & \mathbf{O}_{n\times m} \\ \mathbf{O}'_{n\times m} & \mathbf{A} \end{pmatrix} \quad \textit{and} \quad \tilde{\mathbf{b}} = \begin{pmatrix} \mathbf{O}_{n\times 1} \\ \mathbf{b} \end{pmatrix}.$$

Then $\mathbf{AY} + \mathbf{b} \sim WN_{p,m}(\tilde{\mathbf{A}}\boldsymbol{\mu} + \mathbf{b}, \tilde{\mathbf{A}}\boldsymbol{\Sigma}\tilde{\mathbf{A}}', g)$.

This proposition can be easily obtained since

$$\tilde{\mathbf{A}} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} + \tilde{\mathbf{b}} = \begin{pmatrix} \mathbf{X} \\ \mathbf{AY} + \mathbf{b} \end{pmatrix} \sim N_{n+p}\left(\begin{pmatrix} \boldsymbol{\theta} \\ \mathbf{A}\boldsymbol{\gamma} + \mathbf{b} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Psi} & \boldsymbol{\Gamma}\mathbf{A}' \\ \mathbf{A}\boldsymbol{\Gamma}' & \mathbf{A}\boldsymbol{\Delta}\mathbf{A}' \end{pmatrix} \right).$$

COROLLARY 4.1. *Under the assumptions of Proposition* 4.3 *and for* $\mathbf{Y} \sim WN_{m,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$, *the random variable* $\bar{\mathbf{Y}} = \mathbf{a}'\mathbf{Y}$, $\mathbf{a} \neq \mathbf{0}$, *is distributed as* $WN_{1,n}(\tilde{\mathbf{a}}'\boldsymbol{\mu}, \tilde{\mathbf{a}}'\boldsymbol{\Sigma}\tilde{\mathbf{a}}, g)$.

PROPOSITION 4.4. *Let* $\mathbf{Y} \sim WN_{m,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$ *and let us divide* $\mathbf{Y}$ *into two subvectors*: *a* k_1*-dimensional vector* $\mathbf{Y}_1$ *and a* k_2*-dimensional vector* $\mathbf{Y}_2$. *Moreover, let us consider the following corresponding partitions of* $\boldsymbol{\gamma}$, $\boldsymbol{\Gamma}$, *and* $\boldsymbol{\Delta}$:

$$\gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}, \quad \boldsymbol{\Gamma} = \begin{pmatrix} \boldsymbol{\Gamma}_1 & \boldsymbol{\Gamma}_2 \end{pmatrix}, \quad \textit{and} \quad \boldsymbol{\Delta} = \begin{pmatrix} \boldsymbol{\Delta}_{11} & \boldsymbol{\Delta}_{12} \\ \boldsymbol{\Delta}'_{12} & \boldsymbol{\Delta}_{22} \end{pmatrix}.$$

Then:

(i) $\mathbf{Y}_1 \sim WN_{k_1,n}(\tilde{\boldsymbol{\mu}}_1, \tilde{\boldsymbol{\Sigma}}_{11}, g)$, *where*

$$\tilde{\boldsymbol{\mu}}_1 = \begin{pmatrix} \boldsymbol{\theta}_1 \\ \gamma_1 \end{pmatrix} \quad \textit{and} \quad \tilde{\boldsymbol{\Sigma}}_{11} = \begin{pmatrix} \boldsymbol{\Psi} & \boldsymbol{\Gamma}_1 \\ \boldsymbol{\Gamma}'_1 & \boldsymbol{\Delta}_{11} \end{pmatrix}.$$

(ii) $(\mathbf{Y}_2 | \mathbf{Y}_1 = \mathbf{y}_1) \sim WN_{k_2,n}(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}, g)$, *where*

$$\tilde{\boldsymbol{\mu}} = \begin{pmatrix} \boldsymbol{\theta} \\ \gamma_2 \end{pmatrix} + \begin{pmatrix} \boldsymbol{\Gamma}_1 \\ \boldsymbol{\Delta}'_{12} \end{pmatrix} \boldsymbol{\Delta}_{11}^{-1} (\mathbf{y}_1 - \gamma_1)$$

and

$$\tilde{\boldsymbol{\Sigma}} = \begin{pmatrix} \boldsymbol{\Psi} - \boldsymbol{\Gamma}_1\boldsymbol{\Delta}_{11}^{-1}\boldsymbol{\Gamma}'_1 & \boldsymbol{\Gamma}_2 - \boldsymbol{\Gamma}_1\boldsymbol{\Delta}_{11}\boldsymbol{\Delta}_{12} \\ \boldsymbol{\Gamma}'_2 - \boldsymbol{\Delta}'_{12}\boldsymbol{\Delta}_{11}^{-1}\boldsymbol{\Gamma}'_1 & \boldsymbol{\Delta}_{22} - \boldsymbol{\Delta}'_{12}\boldsymbol{\Delta}_{11}^{-1}\boldsymbol{\Delta}_{12} \end{pmatrix}.$$

THEOREM 4.1. *Let* $\mathbf{Y} = (Y_1, \ldots, Y_m) \sim WN_{m,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$. *Then*:

(i) *We have*

$$E(Y_i) = \frac{\int_{\mathcal{C}} g(\mathbf{x})\varphi_n(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Psi})\, \boldsymbol{\mu}_i^*(x_i) d\mathbf{x}}{\int_{\mathcal{C}} g(\mathbf{x})\varphi_n(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Psi})\, d\mathbf{x}},$$

where $\mu_i^*(\mathbf{x})$ *is the* i*-th component of* $\mu^*(\mathbf{x})$.

(ii) $\operatorname{cov}(Y_i, Y_j) = \sigma_{ij}^* - E(Y_i)E(Y_j)$, *where* σ_{ij}^* *is the covariance between* $(Y_i|X_i = x_i)$ *and* $(Y_j|X_j = x_j)$, $i, j = 1, \ldots, m$.

REMARK. If $g(\mathbf{x})$ is symmetric about $\boldsymbol{\theta}$ for all $\mathbf{x}$ and $\mathcal{C} = \mathbb{R}^n$, then $E(Y_i) = \gamma_i$.

Note that if $\mathbf{A}$ is an $(n \times n)$ semipositive definite matrix and $g(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$, then the pdf of $\mathbf{Y} \sim W_{m,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$ can be written as

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= \frac{\int_{\mathbb{R}^n} \mathbf{x}'\mathbf{A}\mathbf{x}\varphi_{n+m}(\mathbf{x}, \mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma})d\mathbf{x}}{\int_{\mathbb{R}^n} \mathbf{x}'\mathbf{A}\mathbf{x}\varphi_n(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Psi})d\mathbf{x}} \\ &= \varphi_m(\mathbf{y}; \gamma, \boldsymbol{\Delta}) \frac{\int_{\mathbb{R}^n} \mathbf{x}'\mathbf{A}\mathbf{x}\varphi_n\big(\mathbf{x}; \boldsymbol{\tau}(\mathbf{y}), \boldsymbol{\eta}\big)d\mathbf{x}}{\int_{\mathbb{R}^n} \mathbf{x}'\mathbf{A}\mathbf{x}\varphi_n(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Psi})d\mathbf{x}}, \end{aligned}$$

where $\mathbf{y} \in \mathbb{R}^m$, $\boldsymbol{\tau}(\mathbf{y}) = \boldsymbol{\theta} + \boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1}(\mathbf{y} - \gamma)$, and $\boldsymbol{\eta} = \boldsymbol{\Psi} - \boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1}\boldsymbol{\Gamma}'$. Since it follows that $\varphi_n\big(\mathbf{x}; \boldsymbol{\tau}(\mathbf{y}), \boldsymbol{\eta}\big)$ is the pdf of $N_n\big(\boldsymbol{\tau}(\mathbf{y}), \boldsymbol{\eta}\big)$, we have

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= \varphi_m(\mathbf{y}; \gamma, \boldsymbol{\Delta}) \frac{E\left(\mathbf{X}'\mathbf{A}\mathbf{X}|\mathbf{Y} = \mathbf{y}\right)}{E\left(\mathbf{X}'\mathbf{A}\mathbf{X}\right)} \\ &= \varphi_m(\mathbf{y}; \gamma, \boldsymbol{\Delta}) \frac{\operatorname{tr}(\mathbf{A}\boldsymbol{\eta}) + \boldsymbol{\tau}(\mathbf{y})'\mathbf{A}\boldsymbol{\tau}(\mathbf{y})}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}}. \end{aligned} \tag{4.4}$$

PROPOSITION 4.5. *Let* $\mathbf{A}$ *be an* $(n \times n)$ *semipositive definite matrix and assume that* $\mathbf{Y} \sim WN_{m,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$ *with* $g(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$. *Then*

$$M_{\mathbf{Y}}(\mathbf{u}) = \frac{Q}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}} \exp\left(\mathbf{u}'\gamma + \frac{1}{2}\mathbf{u}'\boldsymbol{\Delta}\mathbf{u}\right),$$

where

$$\begin{aligned} Q = {} & \operatorname{tr}(\boldsymbol{\mathcal{C}}\boldsymbol{\Delta}) + (\boldsymbol{\Delta}\mathbf{u})'\boldsymbol{\mathcal{C}}(\boldsymbol{\Delta}\mathbf{u}) + 2\gamma'\boldsymbol{\mathcal{C}}\boldsymbol{\Delta}\mathbf{u} + \gamma'\boldsymbol{\mathcal{C}}\gamma \\ & + 2\boldsymbol{\mathcal{S}}'\mathbf{A}\boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1}(\boldsymbol{\Delta}\mathbf{u} + \gamma) + \boldsymbol{\mathcal{S}}'\mathbf{A}\boldsymbol{\mathcal{S}} + \operatorname{tr}(\mathbf{A}\boldsymbol{\eta}), \end{aligned}$$

$$\boldsymbol{\mathcal{S}} = \boldsymbol{\theta} - \boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1}\gamma \quad \textit{and} \quad \boldsymbol{\mathcal{C}} = \boldsymbol{\Delta}^{-1}\boldsymbol{\Gamma}'\mathbf{A}\boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1}.$$

THEOREM 4.2. *Let* $\mathbf{A}$ *be an* $(n \times n)$ *semipositive definite matrix. If* $\mathbf{Y} \sim WN_{m,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$, *with* $g(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$. *Then*

$$E(\mathbf{Y}) = \frac{\operatorname{tr}(\boldsymbol{\mathcal{C}}\boldsymbol{\Delta} + \mathbf{A}\boldsymbol{\eta}) + \gamma'\boldsymbol{\mathcal{C}}\gamma + 2(\boldsymbol{\mathcal{S}}'\mathbf{A}\boldsymbol{\Gamma}\gamma + \boldsymbol{\mathcal{S}}'\mathbf{A}\boldsymbol{\mathcal{S}})}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}} + \frac{2(\boldsymbol{\Delta}\boldsymbol{\mathcal{C}}\gamma + \boldsymbol{\Gamma}\mathbf{A}\boldsymbol{\mathcal{S}})}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}}$$

and

$$\begin{aligned} E(\mathbf{Y}\mathbf{Y}') = {} & \frac{2}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}} \big((\boldsymbol{\Delta}\boldsymbol{\mathcal{C}}\gamma + \boldsymbol{\Gamma}\mathbf{A}\boldsymbol{\mathcal{S}})\boldsymbol{\Delta} + \operatorname{tr}(\boldsymbol{\mathcal{C}}\boldsymbol{\Delta}) + \gamma'\boldsymbol{\mathcal{C}}\gamma \\ & + 2\mathbf{A}\boldsymbol{\mathcal{S}}\boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1}\gamma + 2\boldsymbol{\Delta}\boldsymbol{\mathcal{C}}\gamma + 2\boldsymbol{\Gamma}\mathbf{A}\boldsymbol{\mathcal{S}} + (\boldsymbol{\Delta}\boldsymbol{\mathcal{C}}\gamma + \boldsymbol{\Gamma}\mathbf{A}\boldsymbol{\mathcal{S}} + \boldsymbol{\Delta}\boldsymbol{\mathcal{C}}\boldsymbol{\Delta})\gamma'\big). \end{aligned}$$

THEOREM 4.3. *Let* $\mathbf{Y} \sim WN_{m,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$. *Then* $\mathbf{Y} \sim N_m(\boldsymbol{\gamma}, \boldsymbol{\Delta})$ *if and only if* $\boldsymbol{\Gamma} = \mathbf{0}$.

THEOREM 4.4. *Suppose that* $\mathbf{Y}_i,\ i = 1, 2, \ldots, r$, *are independent random vectors such that* $\mathbf{Y}_i \sim WN_{m_i,n_i}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, g_i)$. *Then*

$$\mathbf{Y} = (\mathbf{Y}_1', \ldots, \mathbf{Y}_r')' \sim WN_{m^*,n^*}(\boldsymbol{\mu}^\diamond, \boldsymbol{\Sigma}^\diamond, g^\diamond),$$

where $\mathcal{C}^\diamond = \mathcal{C}_1 \times \ldots \times \mathcal{C}_r$, $g^\diamond(\mathbf{x}_1, \ldots, \mathbf{x}_r) = \prod_{i=1}^r g_i(\mathbf{x}_i)$, $\boldsymbol{\mu}^\diamond = \mathrm{vec}(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_r)$, $\boldsymbol{\mu}_i = \binom{\boldsymbol{\theta}_i}{\boldsymbol{\gamma}_i}$, $\boldsymbol{\Sigma}^\diamond = \bigoplus_{i=1}^r \boldsymbol{\Sigma}_i$, $\boldsymbol{\theta}^\diamond = \mathrm{vec}(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_r)$, $\boldsymbol{\Gamma}^\diamond = \mathrm{vec}(\boldsymbol{\Gamma}_1, \ldots, \boldsymbol{\Gamma}_r)$, $\boldsymbol{\Psi}^\diamond = \bigoplus_{i=1}^r \boldsymbol{\Psi}_i$, $n^* = \sum_{i=1}^r n_i$, $m^* = \sum_{i=1}^r m_i$, *and the operators* $\oplus$ *and* $\mathrm{vec}(\cdot)$ *are respectively defined as*

$$\mathbf{A} \oplus \mathbf{B} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \quad \text{and} \quad \mathrm{vec}(\mathbf{a}, \mathbf{b}) = (\mathbf{a}', \mathbf{b}')'.$$

THEOREM 4.5. *If* $\mathbf{Y}_i, i = 1, \ldots, r$, *are independent random vectors such that* $\mathbf{Y}_i \sim WN_{m,n_i}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, g_i)$, *then* $\sum_{i=1}^r \mathbf{Y}_i \sim WN_{m,n^*}(\boldsymbol{\mu}^+, \boldsymbol{\Sigma}^+, g^+)$, *where* $g^+ : \mathbb{R}^{n^*} \to \mathbb{R}^+$ *with* $g^+(\mathbf{x}_1, \ldots, \mathbf{x}_r) = \prod_{i=1}^r g(\mathbf{x}_i)$, $\boldsymbol{\mu}^+ = \left(\boldsymbol{\theta}^{\diamond\prime}, \sum_{i=1}^r \boldsymbol{\gamma}_i'\right)'$, *and*

$$\boldsymbol{\Sigma}^+ = \begin{pmatrix} \boldsymbol{\Psi}^\diamond & \boldsymbol{\Gamma}^\diamond \\ \boldsymbol{\Gamma}^{\diamond\prime} & \sum\limits_{i=1}^r \boldsymbol{\Delta}_i \end{pmatrix}.$$

THEOREM 4.6. *A random vector* $\mathbf{Y}$ *has an* n*-dimensional weighted normal distribution if and only if* $\mathbf{a}'\mathbf{Y}$ *has a univariate weighted normal distribution for all non-zero vectors* $\mathbf{a} \in \mathbb{R}^n$.

COROLLARY 4.2. *Let* $\mathbf{Y}$ *and* $\boldsymbol{\epsilon}$ *be independent random vectors such that* $\mathbf{Y} \sim WN_{m,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$ *and* $\boldsymbol{\epsilon} \sim N_m(\mathbf{0}, \sigma^2 \mathbf{I}_m)$. *If* $\mathbf{Z} = \mathbf{Y} + \boldsymbol{\epsilon}$, *then* $\mathbf{Z}$ *is distributed as* $WN_{m,n}(\boldsymbol{\mu}, \boldsymbol{\Omega}^*, g)$, *where* $\boldsymbol{\Omega}^*$ *is the same as* $\boldsymbol{\Sigma}$ *except that* $\boldsymbol{\Delta}$ *is replaced by* $\boldsymbol{\Delta} + \sigma^2 \mathbf{I}_m$.

Gupta and Huang [18] reported a similar result to Theorem 4.6, discussing the relationship between univariate and multivariate skew-normal distributions. On the other hand, Corollary 4.2 shows that the weighted normal distribution is closed under convolution with the normal distribution.

5. CONCLUDING REMARK AND FUTURE STUDIES

In this paper, a new family of distributions, called the *unified weighted family*, is introduced as an extension to the skew-elliptical family and other families. A rich subfamily of the proposed class, called the *weighted normal* (WN) family, is studied extensively and it showed a high flexibility to obey several nice properties such as closure under conditional distributions, marginal distributions, affine transformation as well as closure under convolution with normal variates. Such tractable properties give us a strong motivation to redevelop several statistical techniques under a weighted skew-normal assumption. Although most distributions of the unified weighted family are asymmetric, we showed that, under some conditions, the family includes some symmetric cases. There is much room for future study including:

1. Estimating the parameters of the WN distribution.
2. Studying the properties of the quadratic form of WN random vector.
3. Exploring the role of $g(\mathbf{x})$ in controlling skewness, kurtosis, and modality.

Acknowledgments. The authors thank two anonymous referees and the editor for their helpful comments that helped them make substantial improvements.

6. APPENDIX

6.1. Proof of Theorem 3.1.

(i) Suppose that $f_{\mathbf{U},\mathbf{V}}(\mathbf{u},\mathbf{v})$ is closed under marginal distribution and consider the partition $\mathbf{v} = \binom{\mathbf{v}_1}{\mathbf{v}_2}$, where $\mathbf{v}_1$ is an r-dimensional vector and $\mathbf{v}_2$ is an $(m-r)$-dimensional vector. Then

$$h(\mathbf{v}_1,\mathbf{v}_2) = \frac{\int\limits_{\mathcal{C}} g(\mathbf{u}) f_{\mathbf{U},\mathbf{V}}(\mathbf{u},\mathbf{v}_1,\mathbf{v}_2)d\mathbf{u}}{\int\limits_{\mathcal{C}} g(\mathbf{u}) f_{\mathbf{U}}(\mathbf{u})d\mathbf{u}}, \quad \mathbf{v}\in\mathbb{R}^m,\ \mathbf{u}\in\mathbb{R}^n.$$

Integrating $\mathbf{v}_2$ out gives the following marginal pdf of $\mathbf{v}_1$:

$$h_1(\mathbf{v}_1) = \frac{\int\limits_{\mathcal{C}} g(\mathbf{u}) f_{\mathbf{U},\mathbf{V}_1}(\mathbf{u},\mathbf{v}_1)d\mathbf{u}}{\int\limits_{\mathcal{C}} g(\mathbf{u}) f_{\mathbf{U}}(\mathbf{u})d\mathbf{u}}, \quad \mathbf{v}_1\in\mathbb{R}^r.$$

Since $f_{\mathbf{U},\mathbf{V}_1}(\mathbf{u},\mathbf{v}_1)$ has the same form as $f_{\mathbf{U},\mathbf{V}}(\mathbf{u},\mathbf{v})$, it follows that $h_1(\mathbf{v}_1)$ has the same form as that of $h(\mathbf{v})$.

(ii) Suppose $f_{\mathbf{U},\mathbf{V}}(\mathbf{u},\mathbf{v})$ and $f_{\mathbf{U}}(\mathbf{u})$ are closed under conditional distribution. Then

$$f_{\mathbf{V}_2|\mathbf{V}_1=\mathbf{v}_1}(\mathbf{v}_2) = \frac{f_{\mathbf{V}}(\mathbf{v}_1,\mathbf{v}_2)}{f_{\mathbf{V}_1}(\mathbf{v}_1)} = \frac{\int\limits_{\mathcal{C}} g(\mathbf{u}) f_{\mathbf{U},\mathbf{V}}(\mathbf{u},\mathbf{v}_1,\mathbf{v}_2)d\mathbf{u}}{\int\limits_{\mathcal{C}} g(\mathbf{u}) f_{\mathbf{U},\mathbf{V}_1}(\mathbf{u},\mathbf{v}_1)d\mathbf{u}}, \quad \mathbf{v}_1\in\mathbb{R}^r.$$

Dividing the numerator and denominator of the last equation by $f_{\mathbf{V}_1}(\mathbf{v}_1)$ yields

$$f_{\mathbf{V}_2|\mathbf{V}_1=\mathbf{v}_1}(\mathbf{v}_2) = \frac{f_{\mathbf{V}}(\mathbf{v}_1,\mathbf{v}_2)}{f_{\mathbf{V}_1}(\mathbf{v}_1)} = \frac{\int\limits_{\mathcal{C}} g(\mathbf{u}) f_{\mathbf{U},\mathbf{V}_2|\mathbf{V}_1=\mathbf{v}_1}(\mathbf{u},\mathbf{v}_2|\mathbf{v}_1)d\mathbf{u}}{\int\limits_{\mathcal{C}} g(\mathbf{u}) f_{\mathbf{U}|\mathbf{V}_1=\mathbf{v}_1}(\mathbf{u}|\mathbf{v}_1)d\mathbf{u}}.$$

Since $f_{\mathbf{U},\mathbf{V}}(\mathbf{u},\mathbf{v})$ is closed by conditioning, it follows that $f_{\mathbf{U},\mathbf{V}_2|\mathbf{V}_1=\mathbf{v}_1}(\mathbf{u},\mathbf{v}_2|\mathbf{v}_1)$ and $f_{\mathbf{U}|\mathbf{V}_1=\mathbf{v}_1}(\mathbf{u}|\mathbf{v}_1)$ have the same forms as $f_{\mathbf{U},\mathbf{V}}(\mathbf{u},\mathbf{v})$ and $f_{\mathbf{U}}(\mathbf{u})$, respectively. Therefore, $f_{\mathbf{V}_2|\mathbf{V}_1=\mathbf{v}_1}(\mathbf{v}_2)$ has the same form as that of $h(\mathbf{v})$. ■

6.2. Proof of Proposition 4.1. Since the pdf of an n-dimensional vector $\mathbf{Y} \sim WN_{m,n}(\boldsymbol{\mu},\boldsymbol{\Sigma},g)$ is given by

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\int\limits_{\mathcal{C}} g(\mathbf{x})\varphi_{n+m}(\mathbf{x},\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Sigma})d\mathbf{x}}{\int\limits_{\mathcal{C}} g(\mathbf{x})\varphi_{n}(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})d\mathbf{x}}, \quad \mathbf{y} \in \mathbb{R}^m,$$

we have

$$F_{\mathbf{Y}}(\mathbf{y}) = \int\limits_{-\infty}^{y_1} \cdots \int\limits_{-\infty}^{y_m} \frac{\int\limits_{\mathcal{C}} g(\mathbf{x})\varphi_{n+m}(\mathbf{x},\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Sigma})d\mathbf{x}}{\int\limits_{\mathcal{C}} g(\mathbf{x})\varphi_{n}(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})d\mathbf{x}} dy_m \dots dy_1.$$

Writing $\varphi_{n+m}(\mathbf{x},\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \varphi_m\big(\mathbf{y};\boldsymbol{\mu}^*(\mathbf{x}),\boldsymbol{\Sigma}^*\big)\varphi_n(x;\boldsymbol{\theta},\boldsymbol{\Psi})$ and applying Fubini's theorem we get

$$F_{\mathbf{Y}}(\mathbf{y}) = \frac{\int\limits_{\mathcal{C}} g(\mathbf{x})\Phi_m\big(\mathbf{y};\boldsymbol{\mu}^*(\mathbf{x}),\boldsymbol{\Sigma}^*\big)\varphi_n(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})d\mathbf{x}}{\int\limits_{\mathcal{C}} g(\mathbf{x})\varphi_{n}(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})d\mathbf{x}}. \ ■$$

6.3. Proof of Proposition 4.2. If $\mathbf{u} \in \mathbf{R}^m$, then the mgf of $\mathbf{Y}$ is

$$M_{\mathbf{Y}}(\mathbf{u}) = \frac{\int\limits_{\mathbb{R}^m}\int\limits_{\mathcal{C}} g(\mathbf{x})\exp(\mathbf{u}'\mathbf{y})\varphi_{n+m}(\mathbf{x},\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Sigma})d\mathbf{x}d\mathbf{y}}{\int\limits_{\mathcal{C}} g(\mathbf{x})\varphi_{n}(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})d\mathbf{x}}.$$

Interchanging the order of integrations and using the equality $\varphi_{n+m}(\mathbf{x},\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \varphi_m\big(\mathbf{y};\boldsymbol{\mu}^*(\mathbf{x}),\boldsymbol{\Sigma}^*\big)\ \varphi_n(x;\boldsymbol{\theta},\boldsymbol{\Psi})$, we obtain

$$M_{\mathbf{Y}}(\mathbf{u}) = \frac{\int\limits_{\mathcal{C}} g(\mathbf{x})\varphi_n(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})M_{\mathbf{Y}|\mathbf{x}}(\mathbf{u})d\mathbf{x}}{\int\limits_{\mathcal{C}} \varphi_{n}(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})d\mathbf{x}}. \ ■$$

6.4. Proof of Proposition 4.4. Suppose that $\mathbf{Y} \sim WN_{m,n}(\boldsymbol{\mu},\boldsymbol{\Sigma},g)$ and consider the partition $\mathbf{Y} = \binom{\mathbf{Y}_1}{\mathbf{Y}_2}$. Then

(i) We have

$$f_{\mathbf{Y}}(\mathbf{y}_1,\mathbf{y}_2) = \frac{\int\limits_{\mathcal{C}} g(\mathbf{x})\varphi_{n+k_1+k_2}(\mathbf{x},\mathbf{y}_1,\mathbf{y}_2;\boldsymbol{\mu},\boldsymbol{\Sigma})d\mathbf{x}}{\int\limits_{\mathcal{C}} g(\mathbf{x})\varphi_{n}(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi})d\mathbf{x}}, \quad \mathbf{y}_1 \in \mathbb{R}^{k_1},\ \mathbf{y}_2 \in \mathbb{R}^{k_2}.$$

Integrating $\mathbf{Y}_2$ out, we get the pdf of $\mathbf{Y}_1$:

$$f_{\mathbf{Y}_1}(\mathbf{y}_1) = \frac{\int_{\mathcal{C}} g(\mathbf{x})\varphi_{n+k_1}(\mathbf{x}, \mathbf{y}_1; \tilde{\boldsymbol{\mu}}_1, \tilde{\boldsymbol{\Sigma}}_{11}) d\mathbf{x}}{\int_{\mathcal{C}} g(\mathbf{x})\varphi_n(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Psi}) d\mathbf{x}}.$$

Thus $\mathbf{Y}_1$ is distributed as WN, which concludes the closure under marginalization.

(ii) The conditional distribution of $(\mathbf{Y}_2|\mathbf{Y}_1 = \mathbf{y}_1)$ is obtained as follows:

$$f_{\mathbf{Y}_2|\mathbf{Y}_1=\mathbf{y}_1}(\mathbf{y}_2) = \frac{\int_{\mathcal{C}} g(\mathbf{x})\varphi_{n+k_1+k_2}(\mathbf{x}, \mathbf{y}_1, \mathbf{y}_2; \boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x}}{\int_{\mathcal{C}} g(\mathbf{x})\varphi_n(\mathbf{x}; \tilde{\boldsymbol{\mu}}_1, \tilde{\boldsymbol{\Sigma}}_{11}) d\mathbf{x}}.$$

Dividing by $\varphi_{k_1}(\mathbf{y}_1; \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$, we obtain

$$f_{\mathbf{Y}_2|\mathbf{Y}_1=\mathbf{y}_1}(\mathbf{y}_2) = \frac{\int_{\mathcal{C}} g(\mathbf{x})\varphi_{n+k_2}(\mathbf{x}, \mathbf{y}_2; \tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}) d\mathbf{x}}{\int_{\mathcal{C}} g(\mathbf{x})\varphi_{n+k_1}(\mathbf{x}, \mathbf{y}_1; \boldsymbol{\theta}^*, \boldsymbol{\Psi}^*) d\mathbf{x}},$$

where $\boldsymbol{\theta}^* = \boldsymbol{\theta} + \boldsymbol{\Gamma}_1\boldsymbol{\Delta}_{11}^{-1}(\mathbf{y}_1 - \boldsymbol{\gamma}_1)$ and $\boldsymbol{\Psi}^* = \boldsymbol{\Psi} - \boldsymbol{\Gamma}_1\boldsymbol{\Delta}_{11}^{-1}\boldsymbol{\Gamma}_1'$. ∎

6.5. Proof of Theorem 4.1.

(i) The first moment of Y_i is obtained by using Proposition 4.2. Indeed, we have

$$\frac{\partial M_{\mathbf{Y}}(\mathbf{u})}{\partial u_i}$$

$$= \frac{\partial}{\partial u_i} \frac{\int_{\mathcal{C}} g(\mathbf{x})\varphi(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Psi}) \exp\left(\sum_{i=1}^{m} \left(u_i\boldsymbol{\mu}_i^*(\mathbf{x}) + \frac{1}{2}u_i^2\sigma_{ii}^*\right) + \sum_{i=1}^{m}\sum_{i=1}^{m} u_i u_j \sigma_{ij}^*\right) d\mathbf{x}}{\int_{\mathcal{C}} g(\mathbf{x})\varphi(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Psi}) d\mathbf{x}}$$

for $u_i \in \mathbb{R}$, where σ_{ij}^* is the covariance between $(Y_i|X_i = x_i)$ and $(Y_j|X_j = x_j)$. Interchanging the integral sign and the derivative and setting $\mathbf{u} = \mathbf{0}$ yields

$$\left.\frac{\partial M_{\mathbf{Y}}(\mathbf{u})}{\partial u_i}\right|_{\mathbf{u}=\mathbf{0}} = \frac{\int_{\mathcal{C}} g(\mathbf{x})\boldsymbol{\mu}_i^*(\mathbf{x})\varphi_n(x; \boldsymbol{\theta}, \boldsymbol{\Psi}) d\mathbf{x}}{\int_{\mathcal{C}} g(\mathbf{x})\varphi_n(\mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\Psi}) d\mathbf{x}},$$

where $\boldsymbol{\mu}_i^*(\mathbf{x})$ is the i-th entry of the vector $\boldsymbol{\mu}^*(\mathbf{x})$.

(ii) Since

$$E(Y_iY_j) = \left.\frac{\partial^2 M_{\mathbf{Y}}(\mathbf{u})}{\partial u_i \partial u_j}\right|_{\mathbf{u}=\mathbf{0}},$$

it clear that $E(Y_iY_j) = \sigma_{ij}^*$. ∎

6.6. Proof of Proposition 4.5. We know that if $\mathbf{Y} \sim WN_{m,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$ with $g(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$, then the mgf of $\mathbf{Y}$ is

$$M_{\mathbf{Y}}(\mathbf{u}) = \frac{\operatorname{tr}(\mathbf{A}\boldsymbol{\eta}) \int\limits_{\mathbb{R}^m} \exp(\mathbf{u}'\mathbf{y})\varphi_m(\mathbf{y}; \gamma, \boldsymbol{\Delta})d\mathbf{y}}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}} + \frac{\int\limits_{\mathbb{R}^m} \boldsymbol{\tau}(\mathbf{y})'\mathbf{A}\boldsymbol{\tau}(\mathbf{y}) \exp(\mathbf{u}'\mathbf{y})\varphi_m(\mathbf{y}; \gamma, \boldsymbol{\Delta})d\mathbf{y}}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}}. \tag{6.1}$$

The first term in (6.1) simplifies to

$$\frac{\operatorname{tr}(\mathbf{A}\boldsymbol{\eta}) \exp\left(\mathbf{u}'\gamma + \frac{1}{2}\mathbf{u}'\boldsymbol{\Delta}\mathbf{u}\right)}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}} \tag{6.2}$$

while the second term in (6.1) takes the form

$$\begin{aligned} &\frac{\boldsymbol{\mathcal{S}}'\mathbf{A}\boldsymbol{\mathcal{S}}}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}} \int\limits_{\mathbb{R}^m} \exp(\mathbf{u}'\mathbf{y})\varphi_m(\mathbf{y}; \gamma, \boldsymbol{\Delta})d\mathbf{y} \\ &\quad + \frac{2\boldsymbol{\mathcal{S}}'\mathbf{A}\boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1}}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}} \int\limits_{\mathbb{R}^m} \mathbf{y} \exp(\mathbf{u}'\mathbf{y})\varphi_m(\mathbf{y}; \gamma, \boldsymbol{\Delta})d\mathbf{y} \\ &\quad + \frac{2\boldsymbol{\mathcal{S}}'\mathbf{A}\boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1}}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}} \int\limits_{\mathbb{R}^m} \mathbf{y}'\boldsymbol{\mathcal{C}}\mathbf{y} \exp(\mathbf{u}'\mathbf{y})\varphi_m(\mathbf{y}; \gamma, \boldsymbol{\Delta})d\mathbf{y}, \end{aligned} \tag{6.3}$$

where $\boldsymbol{\mathcal{S}} = \boldsymbol{\theta} - \boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1}\gamma$ and $\boldsymbol{\mathcal{C}} = \boldsymbol{\Delta}^{-1}\boldsymbol{\Gamma}'\mathbf{A}\boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1}$. The first term in (6.3) becomes

$$\frac{\boldsymbol{\mathcal{S}}'\mathbf{A}\boldsymbol{\mathcal{S}} \exp\left(\mathbf{u}'\gamma + \frac{1}{2}\mathbf{u}'\boldsymbol{\Delta}\mathbf{u}\right)}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}}. \tag{6.4}$$

Now, define $\mathbf{Z} = \mathbf{Y} - \gamma$, where $\mathbf{Y} \sim N_m(\gamma, \boldsymbol{\Delta})$. Then

$$\begin{aligned} &\frac{2\boldsymbol{\mathcal{S}}'\mathbf{A}\boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1}}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}} \int\limits_{\mathbf{R}^m} \exp(\mathbf{u}'\mathbf{y})\varphi_m(\mathbf{y}; \gamma, \boldsymbol{\Delta})d\mathbf{y} \\ &\qquad = \frac{2\boldsymbol{\mathcal{S}}\mathbf{A}\boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1} \exp\left(\mathbf{u}'\gamma + \frac{1}{2}\mathbf{u}'\boldsymbol{\Delta}\mathbf{u}\right)}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}} E(\mathbf{W} + \gamma), \end{aligned}$$

where $W \sim N_m(\boldsymbol{\Delta}\mathbf{u}, \boldsymbol{\Delta})$. Using the multivariate normal properties, we get

$$\begin{aligned} &\frac{2\boldsymbol{\mathcal{S}}'\mathbf{A}\boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1}}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}} \int\limits_{\mathbb{R}^m} \mathbf{y} \exp(\mathbf{u}'\mathbf{y})\varphi_m(\mathbf{y}; \gamma, \boldsymbol{\Delta})d\mathbf{y} \\ &\qquad = \frac{2\boldsymbol{\mathcal{S}}'\mathbf{A}\boldsymbol{\Gamma}\boldsymbol{\Delta}^{-1}}{\operatorname{tr}(\mathbf{A}\boldsymbol{\Psi}) + \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}} (\boldsymbol{\Delta}\mathbf{u} + \gamma) \exp\left(\mathbf{u}'\gamma + \frac{1}{2}\mathbf{u}'\boldsymbol{\Delta}\mathbf{u}\right). \end{aligned} \tag{6.5}$$

The last term in (6.3) takes the form

$$\frac{\int_{\mathbb{R}^m} \mathbf{y}'\mathcal{C}\mathbf{y}\exp(\mathbf{u}'\mathbf{y})\varphi_m(\mathbf{y};\gamma,\mathbf{\Delta})d\mathbf{y}}{\mathrm{tr}(\mathbf{A}\mathbf{\Psi})+\boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}} \tag{6.6}$$

$$= \frac{\exp\big(\mathbf{u}'\gamma + \frac{1}{2}\mathbf{u}'\mathbf{\Delta}\mathbf{u}\big)}{\mathrm{tr}(\mathbf{A}\mathbf{\Psi})+\boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}}\big(\mathrm{tr}(\mathcal{C}\mathbf{\Delta}) + (\mathbf{\Delta}\mathbf{u})'\mathcal{C}(\mathbf{\Delta}\mathbf{u}) + 2\gamma'\mathcal{C}\mathbf{\Delta}\mathbf{u} + \gamma'\mathcal{C}\gamma\big).$$

Finally, by (6.2)–(6.6) we complete the proof of the theorem. ∎

6.7. Proof of Theorem 4.3. Given $\mathbf{\Gamma}=\mathbf{0}$, it follows that for $\mathbf{y}\in\mathbb{R}^m$ we have $\varphi_{n+m}(\mathbf{x},\mathbf{y};\boldsymbol{\mu},\mathbf{\Sigma}) = \varphi_n(\mathbf{x};\boldsymbol{\theta},\mathbf{\Psi})\varphi_m(\mathbf{y};\gamma,\mathbf{\Delta})$. Hence $f_{\mathbf{Y}}(\mathbf{y})=\varphi_m(\mathbf{y};\gamma,\mathbf{\Delta})$. Conversely, given $\mathbf{Y}\sim N_m(\gamma,\mathbf{\Delta})$, we have

$$f_{\mathbf{Y}}(\mathbf{y}) = \varphi_m(\mathbf{y};\gamma,\mathbf{\Delta})\frac{\int_{\mathcal{C}} g(\mathbf{x})\varphi\big(\mathbf{x};\boldsymbol{\tau}(\mathbf{y}),\boldsymbol{\eta}\big)d\mathbf{x}}{\int_{\mathcal{C}} g(\mathbf{x})\varphi(\mathbf{x};\boldsymbol{\theta},\mathbf{\Psi})d\mathbf{x}}.$$

Hence,

$$f_{\mathbf{Y}}(\mathbf{y}) = \varphi_m(\mathbf{y};\gamma,\mathbf{\Delta})\frac{\int_{\mathcal{C}} g(\mathbf{x})\exp\big(-\frac{1}{2}\big(\mathbf{x}-\boldsymbol{\tau}(\mathbf{y})\big)'\boldsymbol{\eta}^{-1}\big(\mathbf{x}-\boldsymbol{\tau}(\mathbf{y})\big)\big)d\mathbf{x}}{(2\pi)^{n/2}|\boldsymbol{\eta}|^{1/2}\int_{\mathcal{C}} g(\mathbf{x})\varphi_n(\mathbf{x};\boldsymbol{\theta},\mathbf{\Psi})d\mathbf{x}},$$

where $\boldsymbol{\tau}(\mathbf{y}) = \boldsymbol{\theta} + \mathbf{\Gamma}'\mathbf{\Delta}^{-1}(\mathbf{y}-\gamma)$ and $\boldsymbol{\eta} = \mathbf{\Psi} - \mathbf{\Gamma}\mathbf{\Delta}^{-1}\mathbf{\Gamma}'$. Integrating with respect to $\mathbf{y}$ over the space $\mathcal{R}^m$ yields

$$E\left(\frac{\int_{\mathcal{C}} g(\mathbf{x})\exp\big(-\frac{1}{2}\big(\mathbf{x}-\boldsymbol{\tau}(\mathbf{y})\big)'\boldsymbol{\eta}^{-1}\big(\mathbf{x}-\boldsymbol{\tau}(\mathbf{y})\big)\big)d\mathbf{x}}{(2\pi)^{n/2}|\boldsymbol{\eta}|^{1/2}\int_{\mathcal{C}} g(\mathbf{x})\varphi_n(\mathbf{x};\boldsymbol{\theta},\mathbf{\Psi})d\mathbf{x}}\right) = 1,$$

since the last equality is true for every $\mathbf{x}\in\mathbb{R}^n$. Then $\mathbf{\Gamma}=\mathbf{0}$. ∎

6.8. Proof of Theorem 4.4. The joint pdf of $\mathbf{Y}_1,\dots,\mathbf{Y}_r$ is given by

$$f(\mathbf{y}_1,\dots,\mathbf{y}_r) = \frac{\prod_{i=1}^r \int_{\mathcal{C}_i} g_i(\mathbf{x}_i)\varphi_{n_i+m_i}(\mathbf{x}_i,\mathbf{y}_i;\boldsymbol{\mu}_i,\mathbf{\Sigma}_i)d\mathbf{x}_i}{\prod_{i=1}^r \int_{\mathcal{C}_i} g_i(\mathbf{x}_i)\varphi_{n_i}(\mathbf{x}_i;\boldsymbol{\theta}_i,\mathbf{\Psi}_i)d\mathbf{x}_i},$$

where

$$\boldsymbol{\mu}_i = \begin{pmatrix}\boldsymbol{\theta}_i\\ \gamma_i\end{pmatrix} \quad\text{and}\quad \mathbf{\Sigma}_i = \begin{pmatrix}\mathbf{\Psi}_i & \mathbf{\Gamma}_i\\ \mathbf{\Gamma}_i' & \mathbf{\Delta}_i\end{pmatrix},$$

and $g^*(\mathbf{x}_1,\ldots,\mathbf{x}_r)=\prod_{i=1}^r g_i(\mathbf{x}_i)$. Now

$$\begin{aligned}\prod_{i=1}^r &\varphi_{n_i+m_i}(\mathbf{z}_i;\boldsymbol{\mu}_i,\boldsymbol{\Sigma}_i)\\ &=(2\pi)^{-\frac{1}{2}(n^*+m^*)}\prod_{i=1}^r|\boldsymbol{\Sigma}_i|^{-1/2}\exp\bigg(-\frac{1}{2}\sum_{i=1}^r(\mathbf{z}_i-\boldsymbol{\mu}_i)\boldsymbol{\Sigma}_i^{-1}(\mathbf{z}_i-\boldsymbol{\mu}_i)\bigg)\\ &=(2\pi)^{-\frac{1}{2}(n^*+m^*)}|\boldsymbol{\Sigma}^\diamond|^{-1/2}\exp\bigg(-\frac{1}{2}(\mathbf{z}-\boldsymbol{\mu}^\diamond)(\boldsymbol{\Sigma}^\diamond)^{-1}(\mathbf{z}-\boldsymbol{\mu}^\diamond)\bigg),\end{aligned}$$

where

$$\mathbf{z}_i=\begin{pmatrix}\mathbf{x}_i\\ \mathbf{y}_i\end{pmatrix},\quad n^*=\sum_{i=1}^r n_i,\quad m^*=\sum_{i=1}^r m_i,$$

and $\mathbf{z}=\mathrm{vec}(\mathbf{z}_1,\ldots,\mathbf{z}_r)$, $\boldsymbol{\mu}^\diamond=\mathrm{vec}(\boldsymbol{\mu}_1,\ldots,\boldsymbol{\mu}_r)$, $\boldsymbol{\Sigma}^\diamond=\bigoplus_{i=1}^r\boldsymbol{\Sigma}_i$. Similarly,

$$\begin{aligned}\prod_{i=1}^r &\varphi_{n_i}(\mathbf{x}_i;\boldsymbol{\theta}_i,\boldsymbol{\Psi}_i)\\ &=(2\pi)^{-\frac{1}{2}(n^*)}\prod_{i=1}^r|\boldsymbol{\Psi}_i|^{-1/2}\exp\bigg(-\frac{1}{2}\sum_{i=1}^r(\mathbf{x}_i-\boldsymbol{\theta}_i)\boldsymbol{\Psi}_i^{-1}(\mathbf{x}_i-\boldsymbol{\theta}_i)\bigg)\\ &=(2\pi)^{-\frac{1}{2}(n^*)}|\boldsymbol{\Psi}^\diamond|^{-1/2}\exp\bigg(-\frac{1}{2}(\mathbf{x}^\diamond-\boldsymbol{\theta}^\diamond)(\boldsymbol{\Psi}^\diamond)^{-1}(\mathbf{x}^\diamond-\boldsymbol{\theta}^\diamond)\bigg),\end{aligned}$$

where $\mathbf{x}^\diamond=\mathrm{vec}(\mathbf{x}_1,\ldots,\mathbf{x}_r)$, $\boldsymbol{\theta}^\diamond=\mathrm{vec}(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_r)$, and $\boldsymbol{\Psi}^\diamond=\bigoplus_{i=1}^r\boldsymbol{\Psi}_i$. Finally, we get the pdf

$$f(\mathbf{y}_1,\ldots,\mathbf{y}_r)=\frac{\int\limits_{\mathcal{C}_1\times\ldots\times\mathcal{C}_r} g^*(\mathbf{x}^\diamond)\varphi_{n^*+m^*}(\mathbf{x}^\diamond,\mathbf{y}^\diamond;\boldsymbol{\mu}^\diamond,\boldsymbol{\Sigma}^\diamond)d\mathbf{x}^\diamond}{\int\limits_{\mathcal{C}_1\times\ldots\times\mathcal{C}_r} g^*(\mathbf{x}^\diamond)\varphi_{n^*}(\mathbf{x}^\diamond;\boldsymbol{\theta}^\diamond,\boldsymbol{\Psi}^\diamond)d\mathbf{x}^\diamond},\quad \mathbf{y}^\diamond\in\mathbb{R}^{n^*+m^*},$$

which completes the proof. ■

6.9. Proof of Theorem 4.5. Under the notation: $\mathbf{X}^\diamond=\mathrm{vec}(\mathbf{X}_1,\ldots,\mathbf{X}_r)$, $\mathbf{Y}^\diamond=\mathrm{vec}(\mathbf{Y}_1,\ldots,\mathbf{Y}_r)$, $\boldsymbol{\theta}^\diamond=\mathrm{vec}(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_r)$, $\boldsymbol{\gamma}^\diamond=\mathrm{vec}(\boldsymbol{\gamma}_1,\ldots,\boldsymbol{\gamma}_r)$, $\boldsymbol{\Psi}^\diamond=\bigoplus_{i=1}^r\boldsymbol{\Psi}_i$, $\boldsymbol{\Gamma}^\diamond=\mathrm{vec}(\boldsymbol{\Gamma}_1,\ldots,\boldsymbol{\Gamma}_r)$, $\boldsymbol{\Delta}^*=\sum_{i=1}^r\boldsymbol{\Delta}_i$, $\gamma^*=\sum_{i=1}^r\gamma_i$, $\boldsymbol{\Gamma}^*=\bigoplus_{i=1}^r\gamma_i$, $n^*=\sum_{i=1}^r n_i$, and $\boldsymbol{\Delta}^\diamond=\bigoplus_{i=1}^r\boldsymbol{\Delta}_i$, we have

$$\begin{pmatrix}\mathbf{X}^\diamond\\ \mathbf{Y}^\diamond\end{pmatrix}\sim N_{n^*,m^*}\left(\begin{pmatrix}\boldsymbol{\theta}^\diamond\\ \gamma^\diamond\end{pmatrix},\begin{pmatrix}\boldsymbol{\Psi}^\diamond & \boldsymbol{\Gamma}^*\\ (\boldsymbol{\Gamma}^*)' & \boldsymbol{\Delta}^\diamond\end{pmatrix}\right).$$

This implies that

$$\begin{pmatrix}\mathbf{X}^\diamond\\ \sum_{i=1}^r\mathbf{Y}_i\end{pmatrix}\sim N_{n^*+m^*}\left(\begin{pmatrix}\boldsymbol{\theta}^\diamond\\ \gamma^*\end{pmatrix},\begin{pmatrix}\boldsymbol{\Psi}^\diamond & \boldsymbol{\Gamma}^\diamond\\ (\boldsymbol{\Gamma}^\diamond)' & \boldsymbol{\Delta}^*\end{pmatrix}\right).\tag{6.7}$$

From (6.7) we conclude that $\sum_{i=1}^r\mathbf{Y}_i\sim WN_{m,n^*}(\boldsymbol{\mu}^+,\boldsymbol{\Sigma}^+,g^+)$, which means that the weighted normal distribution is closed under convolution. ■

6.10. Proof of Theorem 4.6. If $T = \mathbf{a}'\mathbf{Y}$ and $\tilde{\mathbf{a}}' = (\mathbf{0}'_{n\times 1}, \mathbf{a}')$, then

$$T \sim WN_{1,n}(\tilde{\mathbf{a}}'\boldsymbol{\mu}, \tilde{\mathbf{a}}'\boldsymbol{\Sigma}^*\tilde{\mathbf{a}}, g)$$

and the mgf of T is

$$M_T(u) = \frac{\int_{\mathcal{C}} g(\mathbf{x})\varphi_n(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi}) \exp\left(u\mathbf{a}'\boldsymbol{\mu}^*(\mathbf{x}) + \frac{1}{2}u^2\mathbf{a}'\boldsymbol{\Sigma}^*\mathbf{a}\right) d\mathbf{x}}{\int_{\mathcal{C}} g(\mathbf{x})\varphi_n(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi}) d\mathbf{x}}, \quad u \in \mathbb{R}.$$

Setting $u = 1$, we have, for every $\mathbf{a}$ and for $u \in \mathbb{R}$,

$$E\left(\exp(\mathbf{a}'\mathbf{Y})\right) = \frac{\int_{\mathcal{C}} g(\mathbf{x})\varphi_n(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi}) \exp\left(\mathbf{a}'\boldsymbol{\mu}^*(\mathbf{x}) + \frac{1}{2}\mathbf{a}'\boldsymbol{\Sigma}^*\mathbf{a}\right) d\mathbf{x}}{\int_{\mathcal{C}} g(\mathbf{x})\varphi_n(\mathbf{x};\boldsymbol{\theta},\boldsymbol{\Psi}) d\mathbf{x}} = M_{\mathbf{Y}}(\mathbf{a}),$$

which is the mgf of $WN_{m,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$. Conversely, if $\mathbf{Y} \sim WN_{m,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$, then using Corollary 4.1 we conclude that $\mathbf{a}'\mathbf{Y} \sim WN_{1,n}(\tilde{\mathbf{a}}'\boldsymbol{\mu}, \tilde{\mathbf{a}}'\boldsymbol{\Sigma}\tilde{\mathbf{a}}, g)$. ■

REFERENCES

[1] R. J. Adler, *The Geometry of Random Field*, Wiley, New York 1981.

[2] T. T. Alodat, *A generalized multivariate skew-normal distribution with application to spatial and regression predictions*, M.Sc. Thesis, Department of Statistics, Yarmouk University, Jordan, 2009.

[3] R. B. Arellano-Valle and A. Azzalini, *On the unification of families of skew-normal distributions*, Scand. J. Statist. 33 (2006), pp. 561–574.

[4] R. B. Arellano-Valle, M. D. Branco, and M. G. Genton, *A unified view on skewed distributions arising from selection*, Canad. J. Statist. 34 (4) (2006), pp. 581–601.

[5] R. B. Arellano-Valle and M. G. Genton, *On fundamental skew distributions*, J. Multivariate Anal. 96 (2005), pp. 93–116.

[6] R. B. Arellano-Valle and M. G. Genton, *An invariance property of quadratic forms in random vectors with a selection distribution, with application to sample variogram and covariogram estimators*, Ann. Inst. Statist. Math. 62 (2010), pp. 363–381.

[7] R. B. Arellano-Valle and M. G. Genton, *Multivariate unified skew-elliptical distributions*, Chil. J. Stat. 1 (1) (2010), pp. 17–33.

[8] B. C. Arnold and R. J. Beaver, *Hidden truncation models*, Sankhya A 62 (2000), pp. 22–35.

[9] A. Azzalini, *A class of distributions which includes the normal ones*, Scand. J. Statist. 12 (1985), pp. 171–178.

[10] A. Azzalini, *The skew-normal distribution and related multivariate families*, Scand. J. Statist. 32 (2005), pp. 159–188.

[11] A. Azzalini and A. Capitanio, *Statistical applications of the multivariate skew-normal distribution*, J. R. Stat. Soc. Ser. B 61 (3) (1999), pp. 579–602.

[12] A. Azzalini and A. Capitanio, *Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution*, J. R. Stat. Soc. Ser. B 65 (2) (2003), pp. 367–389.

[13] A. Azzalini and A. Dalla Valle, *The multivariate skew-normal distribution*, Biometrika 38 (1996), pp. 715–726.

[14] A. Capitanio, A. Azzalini, and E. Stanghellini, *Graphical models for skew-normal variates*, Scand. J. Statist. 30 (2003), pp. 129–144.

[15] J. A. Domínguez-Molina, G. González-Farias, and A. Gupta, *The multivariate closed skew normal distribution*, Technical Report 03-12, Department of Mathematics and Statistics, Bowling Green State University, USA, 2003, p. 21.

[16] M. Ganjali, T. Baghfalahki, and M. Khazaie, *A linear mixed model for analyzing longitudinal skew-normal responses with random dropout*, J. Korean Statist. Soc. 42 (2013), pp. 149–160.

[17] G. González-Farias, J. A. Domínguez-Molina, and A. K. M. Gupta, *The closed skew-normal distribution*, in: *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*, M. G. Genton (Ed.), Chapman and Hall/CRC, 2004.

[18] A. K. Gupta and W. J. Huang, *Quadratic forms in skew normal variates*, J. Math. Anal. Appl. 273 (2002), pp. 558–564.

[19] H.-J. Kim, *A class of weighted multivariate normal distributions and its properties*, J. Multivariate Anal. 99 (2008), pp. 1758–1771.

[20] H.-J. Kim, *A class of weighted multivariate elliptical models useful for robust analysis of nonnormal and bimodal data*, J. Korean Statist. Soc. 39 (2010), pp. 83–92.

[21] A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables: Theory and Applications*, Marcel Dekker, New York 1992.

[22] K. Podgórski and I. Rychlik, *Envelope crossing distributions for Gaussian fields*, Probab. Engineering Mechanics 23 (2008), pp. 364–377.

[23] V. Raluca, *On the multivariate skew-normal distribution and its scale mixtures*, An. Ştiinţ. Univ. “Ovidius” Constanţa Ser. Mat. 13 (2) (2005), pp. 83–96.

[24] J. Rodrigues, *A Bayesian inference for the extended skew-normal measurement error model*, Braz. J. Probab. Stat. 20 (2006), pp. 179–190.

M. T. Alodat
Department of Mathematics, Statistics and Physics
University of Qatar, Doha, Qatar
E-mail: alodatmts@yahoo.com
Permanent address:
Department of Statistics
Yarmouk University, Irbid, Jordan

S. O. Al-Salman
Department of Statistics
Yarmouk University, Irbid, Jordan
E-mail: alsalman85@yahoo.com

M. Y. Al-Rawwash
Department of Mathematics
Sharjah University, Sharjah, UAE
E-mail: malrawwash@sharjah.ac.ae

Received on 22.6.2012;
revised version on 7.2.2014