DISTRIBUTION OF THE PRODUCT OF SINGULAR WISHART MATRIX AND NORMAL VECTOR

UDC 519.21

T. BODNAR, S. MAZUR, AND Y. OKHRIN

ABSTRACT. In this paper we derive a very useful formula for the stochastic representation of the product of a singular Wishart matrix with a normal vector. Using this result, the expressions of the density function as well as of the characteristic function are established. Moreover, the derived stochastic representation is used to generate random samples from product which leads to a considerable improvement in the computation efficiency. Finally, we present several important properties of the singular Wishart distribution, like its characteristic function and distributional properties of the partitioned singular Wishart matrix.

1. INTRODUCTION

The theory of Wishart distribution contains numerous important and useful results which are applied in theoretical and applied statistics. The distributional properties of random matrices which follow a Wishart distribution, an inverse Wishart distribution and related quantities were established in Olkin and Roy (1954), Khatri(1959), Olkin and Rubin (1964), Gupta and Nagar (2000), Bodnar and Okhrin (2008), Drton *et al.* (2008) and others. Massam and Weselowski (2006) discussed in detail the characterization of the Wishart distribution and extended the results of Geiger and Heckerman (1998, 2002).

A k-dimensional symmetric positive semi-definite random matrix $\mathbf{V} = (v_{ij})_{i,j=1,\ldots,k}$ is Wishart distributed with n degrees of freedom and covariance matrix $\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j=1,\ldots,k}$, i.e. $\mathbf{V} \sim W_k(n, \boldsymbol{\Sigma})$, if the joint density function of its functionally independent elements, that is of $\mathbf{vech}(\mathbf{V}) = (v_{11}, \ldots, v_{n1}, v_{12}, \ldots, v_{n2}, \ldots, v_{nn})$, is given by

$$f(\mathbf{V}) = \frac{|\mathbf{V}|^{(n-k-1)/2}}{2^{kn/2}\Gamma_k(n/2)|\boldsymbol{\Sigma}|^{n/2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\mathbf{V}\right),$$

which is defined on a cone of positive definite matrices (cf. Andersson and Wojnar (2004)). The symbol $\Gamma_m(\cdot)$ denotes the multivariate gamma function expressed as

$$\Gamma_m(n/2) = \pi^{k(k-1)/4} \prod_{i=1}^{m} \Gamma((n-i+1)/2).$$

Gupta and Nagar (2000) discussed the relationship between the Wishart distribution and the matrix variate normal distribution. They proved that if $\mathbf{X} \sim N_{k,n}(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I}_n)$ (matrix variate normal distribution with zero mean matrix and covariance matrix $\boldsymbol{\Sigma} \otimes \mathbf{I}_n$ where $\mathbf{I}_n$ denotes the n-dimensional identity matrix), then $\mathbf{V} = \mathbf{X}\mathbf{X}^T \sim W_k(n, \boldsymbol{\Sigma})$ as

soon as $n \geq k$ (see Theorem 3.2.2 of Gupta and Nagar (2000)). Recently, Srivastava (2003) extended this result by deriving the distribution of the quadratic form

$$\mathbf{A} = \mathbf{X}\mathbf{X}^T \quad \text{with } \mathbf{X} \sim N_{k,n}(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I}_n), \quad k > n. \tag{1}$$

This distribution was called the singular Wishart distribution and it is denoted by $\mathbf{A} \sim W_k(n, \boldsymbol{\Sigma})$, $k > n$. Although the matrix $\mathbf{A}$ is singular, its density function was derived in terms of its functionally independent elements, i.e. $\mathbf{A}_{11}$ and $\mathbf{A}_{21}$, where $\mathbf{A}_{11}$ and $\mathbf{A}_{21}$ are obtained from the following partitioned matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \tag{2}$$

with $\dim(\mathbf{A}_{11}) = \dim(\boldsymbol{\Sigma}_{11}) = n \times n$, $k > n$. The density function of the singular Wishart distributed random matrix $\mathbf{A}$ is then given by (cf. Srivastava (2003, p. 1550))

$$f(\mathbf{A}) \equiv f(\mathbf{A}_{11}, \mathbf{A}_{21}) = \frac{\pi^{n(n-k)/2} 2^{-kn/2}}{\Gamma_n(n/2)|\boldsymbol{\Sigma}|^{n/2}} |\mathbf{A}_{11}|^{(n-k-1)/2} \exp\left(-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\mathbf{A}\right)\right). \tag{3}$$

Bodnar and Okhrin (2008) derived several distributional results for the functions calculated for the elements of the singular Wishart distributed matrix.

Although the (singular) Wishart distribution has been applied in different fields of science and a lot of papers were devoted to studying its distributional properties, the Wishart matrix usually does not appear alone but in a combination with a normally distributed random vector. For instance, Mathai and Provost (1992) discussed in detail the distribution of quadratic forms in normally distributed random matrices, while Bodnar and Okhrin (2011) derived the distribution of the product of an inverse Wishart random matrix and a normally distributed random vector. This combination is of great interest both in statistical theory and in the applications, for example to portfolio analysis. In statistical theory the product of an inverse Wishart matrix and a normal vector appears in the discriminant function (see, e.g. Rencher (2002)), whereas some types of optimal portfolio weights are determined by this product in portfolio theory (cf. Bodnar and Okhrin (2011)). Bodnar *et al.* (2013) considered expressions which depend on $\mathbf{Az}$, where $\mathbf{A}$ is a (non-singular) Wishart matrix and $\mathbf{z}$ is a Gaussian vector, which are independently distributed and derived a stochastic representation as well as the exact density function of $\mathbf{LAz}$ for an arbitrary deterministic matrix $\mathbf{L}$.

In this paper we extend the results of Bodnar *et al.* (2013) by investigating the distributional properties of the product of a singular Wishart matrix and a normal vector. The singularity of the Wishart distribution leads to substantial technical complications, which have to be solved. Particularly, this refers to a special type of partitioning which leads to singular matrix variate normal distributions. Furthermore, the singular Wishart distribution becomes more relevant nowadays, when the covariance matrix has to be estimated for high-dimensional data using a few observations. For example, this problem arises in portfolio theory, when we consider many assets and historical returns over a few past periods. The results established in this paper can tackle such problems in contrary to the results provided by Bodnar *et al.* (2013).

The rest of the paper is structured as follows. In Section 2 we consider several distributional properties of the singular Wishart random matrix. In particular, we derive its characteristic function in Theorem 1 and prove that the singular Wishart distribution is closed with respect to linear symmetric transformations given by $\mathbf{LAL}^T$ in Theorem 2. In Theorem 3 the distribution of the quadratic form based on the singular Wishart distribution is presented. In Section 3 main results are given. In Theorem 4 we obtain a very useful stochastic representation for the product $\mathbf{LAz}$, where $\mathbf{L}$ is a $p \times k$ matrix of constants and $\mathbf{z}$ is a k-dimensional normal vector which is independent of $\mathbf{A}$. Its density function is derived in Corollary 1. Important special cases are considered in Corollaries 2

and 3. The characteristic function for the product $\mathbf{Az}$ is given in Theorem 5. The results of numerical studies are discussed in Section 4. Section 5 summarizes the paper.

2. Preliminary results

In this section we present several distributional properties of the elements of the singular Wishart random matrix which are used in the proofs of our main results of Section 3. We start with presenting Lemma 1 which extends the results of Srivastava (2003, Corollary 3.4) and Bodnar and Okhrin (2008, Lemma 1) to an arbitrary partitioning in (2).

Lemma 1. *Let* $\mathbf{A} \sim W_k(n, \boldsymbol{\Sigma})$, $k > n$ *where* $\mathbf{A}$ *and* $\boldsymbol{\Sigma}$ *are partitioned as follows*

$$\mathbf{A} = \begin{bmatrix} \widetilde{\mathbf{A}}_{11} & \widetilde{\mathbf{A}}_{12} \\ \widetilde{\mathbf{A}}_{21} & \widetilde{\mathbf{A}}_{22} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \widetilde{\boldsymbol{\Sigma}}_{11} & \widetilde{\boldsymbol{\Sigma}}_{12} \\ \widetilde{\boldsymbol{\Sigma}}_{21} & \widetilde{\boldsymbol{\Sigma}}_{22} \end{bmatrix} \tag{4}$$

with $\dim(\widetilde{\mathbf{A}}_{11}) = \dim(\widetilde{\boldsymbol{\Sigma}}_{11}) = p \times p$. *Then, it holds that*

$$\widetilde{\mathbf{A}}_{21} | \widetilde{\mathbf{A}}_{11} \sim N_{k-p,p}(\widetilde{\boldsymbol{\Sigma}}_{21} \widetilde{\boldsymbol{\Sigma}}_{11}^{-1} \widetilde{\mathbf{A}}_{11}, \widetilde{\boldsymbol{\Sigma}}_{22\cdot1} \otimes \widetilde{\mathbf{A}}_{11}),$$

where $\widetilde{\boldsymbol{\Sigma}}_{22\cdot1} = \widetilde{\boldsymbol{\Sigma}}_{22} - \widetilde{\boldsymbol{\Sigma}}_{21} \widetilde{\boldsymbol{\Sigma}}_{11}^{-1} \widetilde{\boldsymbol{\Sigma}}_{12}$.

Proof. The statement of the lemma was proved for $p = n$ by Srivastava (2003, Corollary 3.4) and for $p < n$ by Bodnar and Okhrin (2008, Lemma 1). Thus, we deal with the case $p > n$ only.

Because $\mathbf{A}_{22} = \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}$, we get

$$\mathbf{A} = \begin{bmatrix} \widetilde{\mathbf{A}}_{11} & \widetilde{\mathbf{A}}_{12} \\ \widetilde{\mathbf{A}}_{21} & \widetilde{\mathbf{A}}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12;1} & \mathbf{A}_{12;2} \\ \mathbf{A}_{21;1} & \mathbf{A}_{21;1} \mathbf{A}_{11}^{-1} \mathbf{A}_{12;1} & \mathbf{A}_{21;1} \mathbf{A}_{11}^{-1} \mathbf{A}_{12;2} \\ \mathbf{A}_{21;2} & \mathbf{A}_{21;2} \mathbf{A}_{11}^{-1} \mathbf{A}_{12;1} & \mathbf{A}_{21;2} \mathbf{A}_{11}^{-1} \mathbf{A}_{12;2} \end{bmatrix},$$

where $\mathbf{A}_{11}$ and $\mathbf{A}_{21} = \begin{bmatrix} \mathbf{A}_{21;1} \\ \mathbf{A}_{21;2} \end{bmatrix}$ are given in (2). Moreover, let $\boldsymbol{\Sigma}_{21} = \begin{bmatrix} \boldsymbol{\Sigma}_{21;1} \\ \boldsymbol{\Sigma}_{21;2} \end{bmatrix}$. We are interested in the conditional distribution of

$$\begin{bmatrix} \mathbf{A}_{21;2} & \mathbf{A}_{21;2} \mathbf{A}_{11}^{-1} \mathbf{A}_{12;1} \end{bmatrix} = \mathbf{A}_{21;2} \begin{bmatrix} \mathbf{I} & \mathbf{A}_{11}^{-1} \mathbf{A}_{12;1} \end{bmatrix} \tag{5}$$

given $\mathbf{A}_{11}$ and $\mathbf{A}_{12;1}$. Using (5) we first derive the distribution of $\mathbf{A}_{21;2} | \mathbf{A}_{11}, \mathbf{A}_{12;1}$. The application of Corollary 3.4 by Srivastava (2003) leads to

$$\mathbf{A}_{21} = \begin{bmatrix} \mathbf{A}_{21;1} \\ \mathbf{A}_{21;2} \end{bmatrix} \mid \mathbf{A}_{11} \sim N_{k-n,n} \left(\boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \mathbf{A}_{11}, \boldsymbol{\Sigma}_{22\cdot1} \otimes \mathbf{A}_{11} \right).$$

Let

$$\mathbf{B} = \boldsymbol{\Sigma}_{22\cdot1} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \quad \text{and} \quad \mathbf{B}_{22\cdot1} = \mathbf{B}_{22} - \mathbf{B}_{21} \mathbf{B}_{11}^{-1} \mathbf{B}_{12}.$$

Using Theorem 2.21 by Gupta *et al.* (2013) for the case of matrix variate normal distribution we get

$$\mathbf{A}_{21;2} \mid \mathbf{A}_{11}, \mathbf{A}_{21;1}$$
$$\sim N_{k-p,n} \left(\boldsymbol{\Sigma}_{21;2} \boldsymbol{\Sigma}_{11}^{-1} \mathbf{A}_{11} + \mathbf{B}_{21} \mathbf{B}_{11}^{-1} \left(\mathbf{A}_{21;1} - \boldsymbol{\Sigma}_{21;1} \boldsymbol{\Sigma}_{11}^{-1} \mathbf{A}_{11} \right), \mathbf{B}_{22\cdot1} \otimes \mathbf{A}_{11} \right).$$

Now, the application of Theorem 2.2 in Gupta *et al.* (2013) in the same case of matrix variate normal distribution leads to

$$\widetilde{\mathbf{A}}_{21} \mid \widetilde{\mathbf{A}}_{11} = \mathbf{A}_{21;2} \begin{bmatrix} \mathbf{I} & \mathbf{A}_{11}^{-1} \mathbf{A}_{12;1} \end{bmatrix} \mid \mathbf{A}_{11}, \mathbf{A}_{12;1}$$
$$\sim N_{k-p,p} \left(\left(\boldsymbol{\Sigma}_{21;2} \boldsymbol{\Sigma}_{11}^{-1} \mathbf{A}_{11} + \mathbf{B}_{21} \mathbf{B}_{11}^{-1} \left(\mathbf{A}_{21;1} - \boldsymbol{\Sigma}_{21;1} \boldsymbol{\Sigma}_{11}^{-1} \mathbf{A}_{11} \right) \right) \begin{bmatrix} \mathbf{I} & \mathbf{A}_{11}^{-1} \mathbf{A}_{12;1} \end{bmatrix},$$
$$\mathbf{B}_{22\cdot1} \otimes \left(\begin{bmatrix} \mathbf{I} \\ \mathbf{A}_{21;1} \mathbf{A}_{11}^{-1} \end{bmatrix} \mathbf{A}_{11} \begin{bmatrix} \mathbf{I} & \mathbf{A}_{11}^{-1} \mathbf{A}_{12;1} \end{bmatrix} \right) \right).$$

Let

$$\mathbf{Q} = \mathbf{\Sigma}_{22} = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix} \quad \text{with } \mathbf{Q}_{22} = \tilde{\mathbf{\Sigma}}_{22}.$$

Then

$$\tilde{\mathbf{\Sigma}}_{11} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{21;1} \\ \mathbf{\Sigma}_{21;1} & \mathbf{Q}_{11} \end{pmatrix} \tag{6}$$

and

$$\mathbf{B}_{ij} = \mathbf{Q}_{ij} - \mathbf{\Sigma}_{21;i}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12;j} \quad \text{for } i,j = 1,2. \tag{7}$$

Moreover, we get

$$\begin{bmatrix} \mathbf{I} \\ \mathbf{A}_{21;1}\mathbf{A}_{11}^{-1} \end{bmatrix} \mathbf{A}_{11} \begin{bmatrix} \mathbf{I} & \mathbf{A}_{11}^{-1}\mathbf{A}_{12;1} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12;1} \\ \mathbf{A}_{21;1} & \mathbf{A}_{21;1}\mathbf{A}_{11}^{-1}\mathbf{A}_{12;1} \end{bmatrix} = \widetilde{\mathbf{A}}_{11}, \tag{8}$$

and

$$\begin{aligned}
&\left(\mathbf{\Sigma}_{21;2}\mathbf{\Sigma}_{11}^{-1}\mathbf{A}_{11} + \mathbf{B}_{21}\mathbf{B}_{11}^{-1}\left(\mathbf{A}_{21;1} - \mathbf{\Sigma}_{21;1}\mathbf{\Sigma}_{11}^{-1}\mathbf{A}_{11}\right)\right) \begin{bmatrix} \mathbf{I} & \mathbf{A}_{11}^{-1}\mathbf{A}_{12;1} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{\Sigma}_{21;2}\mathbf{\Sigma}_{11}^{-1} - \mathbf{B}_{21}\mathbf{B}_{11}^{-1}\mathbf{\Sigma}_{21;1}\mathbf{\Sigma}_{11}^{-1} & \mathbf{B}_{21}\mathbf{B}_{11}^{-1} \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{A}_{11} \\ \mathbf{A}_{21;1} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A}_{11}^{-1}\mathbf{A}_{12;1} \end{bmatrix}}_{=\widetilde{\mathbf{A}}_{11}} \\
&= \begin{bmatrix} \mathbf{\Sigma}_{21;2}(\mathbf{\Sigma}_{11}^{-1} + \mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12;1}\mathbf{B}_{11}^{-1}\mathbf{\Sigma}_{21;1}\mathbf{\Sigma}_{11}^{-1}) - \mathbf{Q}_{21}\mathbf{B}_{11}^{-1}\mathbf{\Sigma}_{21;1}\mathbf{\Sigma}_{11}^{-1} \\ -\mathbf{\Sigma}_{21;2}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12;1}\mathbf{B}_{11}^{-1} + \mathbf{Q}_{21}\mathbf{B}_{11}^{-1} \end{bmatrix}^T \widetilde{\mathbf{A}}_{11} \\
&= \begin{bmatrix} \mathbf{\Sigma}_{21;1} & \mathbf{Q}_{21} \end{bmatrix} \begin{pmatrix} \mathbf{\Sigma}_{11}^{-1} + \mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12;1}\mathbf{B}_{11}^{-1}\mathbf{\Sigma}_{21;1}\mathbf{\Sigma}_{11}^{-1} & -\mathbf{B}_{11}^{-1}\mathbf{\Sigma}_{21;1}\mathbf{\Sigma}_{11}^{-1} \\ -\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12;1}\mathbf{B}_{11}^{-1} & \mathbf{B}_{11}^{-1} \end{pmatrix} \widetilde{\mathbf{A}}_{11} \\
&= \widetilde{\mathbf{\Sigma}}_{21}\widetilde{\mathbf{\Sigma}}_{11}^{-1}\widetilde{\mathbf{A}}_{11},
\end{aligned}$$

where in the third equality we use (8) and the fact that $\widetilde{\mathbf{\Sigma}}_{21} = \begin{bmatrix} \mathbf{\Sigma}_{21;1} & \mathbf{Q}_{21} \end{bmatrix}$. The last equality follows from the formula for the inverse of the partitioned matrix $\widetilde{\mathbf{\Sigma}}_{11}$ as in (6) (see, e.g. Harville (1997, Corollary 8.5.12)).

Finally, let

$$\mathbf{\Xi} = \mathbf{\Sigma}^{-1} = \begin{bmatrix} \mathbf{\Xi}_{11} & \mathbf{\Xi}_{12} \\ \mathbf{\Xi}_{21} & \mathbf{\Xi}_{22} \end{bmatrix} = \begin{bmatrix} \widetilde{\mathbf{\Xi}}_{11} & \widetilde{\mathbf{\Xi}}_{12} \\ \widetilde{\mathbf{\Xi}}_{21} & \widetilde{\mathbf{\Xi}}_{22} \end{bmatrix}$$

with $\dim(\mathbf{\Xi}_{11}) = n \times n$ and $\dim(\widetilde{\mathbf{\Xi}}_{11}) = p \times p$. Then $\mathbf{\Sigma}_{22\cdot1} = \mathbf{\Xi}_{22}^{-1}$. Let

$$\mathbf{D} = (\mathbf{\Xi}_{22}^{-1})^{-1} = \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{bmatrix} \quad \text{with } \mathbf{D}_{22} = \widetilde{\mathbf{\Xi}}_{22}.$$

Then $\mathbf{B}_{22\cdot1} = \mathbf{D}_{22}^{-1} = \widetilde{\mathbf{\Xi}}_{22}^{-1}$. On the other hand, we get

$$\widetilde{\mathbf{\Xi}}_{22}^{-1} = \widetilde{\mathbf{\Sigma}}_{22} - \widetilde{\mathbf{\Sigma}}_{21}\widetilde{\mathbf{\Sigma}}_{11}^{-1}\widetilde{\mathbf{\Sigma}}_{12} = \widetilde{\mathbf{\Sigma}}_{22\cdot1},$$

and, hence, $\mathbf{B}_{22\cdot1} = \widetilde{\mathbf{\Sigma}}_{22\cdot1}$ which completes the proof of the lemma. □

Because $\widetilde{\mathbf{A}}_{11}$ is singular, we obtain from Lemma 1 that the conditional distribution of $\widetilde{\mathbf{A}}_{21}$ given $\widetilde{\mathbf{A}}_{11}$ belongs to the family of singular matrix variate normal distributions. This is the only difference to the case of $p \leq n$ where the non-singular matrix variate normal distribution appears.

Next, we derive the characteristic function of $\mathbf{A}$.

Theorem 1. *Let $\mathbf{A}$ be k-dimensional singular Wishart distributed, i.e. $\mathbf{A} \sim W_k(n, \mathbf{\Sigma})$, $k > n$. Then the characteristic function of $\mathbf{A}$ is given by*

$$\Psi(\mathbf{T}) = \left|\mathbf{I}_n + 4\mathbf{T}_{12}\mathbf{\Sigma}_{22\cdot1}\mathbf{T}_{21}\mathbf{\Sigma}_{11} - 2i(\mathbf{T}_{11}\mathbf{\Sigma}_{11} + \mathbf{T}_{12}\mathbf{\Sigma}_{21} + \mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{T}_{21}\mathbf{\Sigma}_{11})\right|^{-n/2}, \tag{9}$$

where $\mathbf{T} = (t_{ij})_{i,j=1,\ldots,k}$ with $t_{ij} = \frac{1}{2}(1 + \delta_{ij})\tau_{ij}$, $\tau_{ij} = \tau_{ji}$, for $i, j = 1, \ldots, k$ and $\tau_{ij} = 0$ for $i, j = n + 1, \ldots, k$. The symbol δ_{ij} denotes the Kronecker delta given by

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

Proof. Let $\mathbf{A}$ and $\boldsymbol{\Sigma}$ be partitioned as in (2) with $\dim(\mathbf{A}_{11}) = n \times n$ and let

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{bmatrix} \tag{10}$$

with $\dim(\mathbf{T}_{11}) = n \times n$ and $\mathbf{T}_{22} = \mathbf{0}$.

Then the characteristic function of $\mathbf{A}$ is given by

$$\begin{aligned} \Psi(\mathbf{T}) &= \mathsf{E}\left[\text{etr}\left(i\mathbf{T}^T\mathbf{A}\right)\right] = \mathsf{E}\left[\text{etr}\left(i\mathbf{T}_{11}^T\mathbf{A}_{11} + 2i\mathbf{T}_{21}^T\mathbf{A}_{21}\right)\right] \\ &= \mathsf{E}\left[\mathsf{E}\left(\text{etr}\left(i\mathbf{T}_{11}^T\mathbf{A}_{11} + 2i\mathbf{T}_{21}^T\mathbf{A}_{21}\right) \mid \mathbf{A}_{11}\right)\right] \\ &= \mathsf{E}\left[\text{etr}\left(i\mathbf{T}_{11}^T\mathbf{A}_{11}\right)\mathsf{E}\left(\text{etr}\left(2i\mathbf{T}_{21}^T\mathbf{A}_{21}\right) \mid \mathbf{A}_{11}\right)\right]. \end{aligned}$$

Using Lemma 1 and Theorem 2.3.2 of Gupta and Nagar (2000) we get

$$\begin{aligned} \Psi(\mathbf{T}) &= \mathsf{E}\left[\text{etr}\left(i\mathbf{T}_{11}^T\mathbf{A}_{11}\right)\text{etr}\left(2i\mathbf{T}_{21}^T\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A}_{11} - 2\mathbf{T}_{21}^T\boldsymbol{\Sigma}_{22\cdot1}\mathbf{T}_{21}\mathbf{A}_{11}\right)\right] \\ &= \mathsf{E}\left[\text{etr}\left(i\left(\mathbf{T}_{11}^T + 2\mathbf{T}_{21}^T\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\right)\mathbf{A}_{11}\right)\text{etr}\left(-2\mathbf{T}_{21}^T\boldsymbol{\Sigma}_{22\cdot1}\mathbf{T}_{21}\mathbf{A}_{11}\right)\right] \\ &= \frac{1}{|\boldsymbol{\Sigma}_{11}|^{n/2}|\boldsymbol{\Sigma}_{11}^{-1} + 4\mathbf{T}_{12}\boldsymbol{\Sigma}_{22\cdot1}\mathbf{T}_{21}|^{n/2}}\mathsf{E}\left[\text{etr}\left(i\left(\mathbf{T}_{11} + 2\mathbf{T}_{12}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\right)\widetilde{\mathbf{A}}_{11}\right)\right], \end{aligned}$$

where $\widetilde{\mathbf{A}} \sim W_n\left(n, (\boldsymbol{\Sigma}_{11}^{-1} + 4\mathbf{T}_{12}\boldsymbol{\Sigma}_{22\cdot1}\mathbf{T}_{21})^{-1}\right)$ and we use that $\mathbf{T}_{11}^T = \mathbf{T}_{11}$ and $\mathbf{T}_{21}^T = \mathbf{T}_{12}$.

Because

$$\begin{aligned} \text{tr}\left[\left(\mathbf{T}_{11} + 2\mathbf{T}_{12}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\right)\widetilde{\mathbf{A}}_{11}\right] &= \text{tr}\left[\widetilde{\mathbf{A}}_{11}^T\left(\mathbf{T}_{11} + 2\mathbf{T}_{12}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\right)^T\right] \\ &= \text{tr}\left[\left(\mathbf{T}_{11} + 2\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\mathbf{T}_{21}\right)\widetilde{\mathbf{A}}_{11}\right] \end{aligned}$$

we get

$$\begin{aligned} \text{tr}\left[\left(\mathbf{T}_{11} + 2\mathbf{T}_{12}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\right)\widetilde{\mathbf{A}}_{11}\right] &= \frac{1}{2}\text{tr}\left[\left(\mathbf{T}_{11} + 2\mathbf{T}_{12}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\right)\right]\widetilde{\mathbf{A}}_{11} \\ &\quad + \frac{1}{2}\text{tr}\left[\left(\mathbf{T}_{11} + 2\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\mathbf{T}_{21}\right)\widetilde{\mathbf{A}}_{11}\right] \\ &= \text{tr}\left[\left(\mathbf{T}_{11} + \mathbf{T}_{12}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} + \boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\mathbf{T}_{21}\right)\widetilde{\mathbf{A}}_{11}\right], \end{aligned}$$

where $\mathbf{T}_{11} + \mathbf{T}_{12}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} + \boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\mathbf{T}_{21}$ is a symmetric matrix. Hence,

$$\begin{aligned} \Psi(\mathbf{T}) &= \frac{1}{|\boldsymbol{\Sigma}_{11}|^{n/2}|\boldsymbol{\Sigma}_{11}^{-1} + 4\mathbf{T}_{12}\boldsymbol{\Sigma}_{22\cdot1}\mathbf{T}_{21}|^{n/2}} \\ &\quad \times \mathsf{E}\left[\text{etr}\left[i\left(\mathbf{T}_{11} + \mathbf{T}_{12}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} + \boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\mathbf{T}_{21}\right)^T\widetilde{\mathbf{A}}_{11}\right]\right], \end{aligned} \tag{11}$$

where the expectation in (11) is the characteristic function of the n-dimensional Wishart distribution with n degrees of freedom and covariance matrix $(\boldsymbol{\Sigma}_{11}^{-1} + 4\mathbf{T}_{12}\boldsymbol{\Sigma}_{22\cdot1}\mathbf{T}_{21})^{-1}$ at $(\mathbf{T}_{11} + 2\mathbf{T}_{12}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})$. The application of Theorem 3.3.7 by Gupta and Nagar (2000) leads to

$$\begin{aligned} \Psi(\mathbf{T}) &= |\boldsymbol{\Sigma}_{11}|^{-n/2}|\boldsymbol{\Sigma}_{11}^{-1} + 4\mathbf{T}_{12}\boldsymbol{\Sigma}_{22\cdot1}\mathbf{T}_{21}|^{-n/2} \\ &\quad \times |\mathbf{I}_n - 2i(\mathbf{T}_{11} + \mathbf{T}_{12}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} + \boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\mathbf{T}_{21})(\boldsymbol{\Sigma}_{11}^{-1} + 4\mathbf{T}_{12}\boldsymbol{\Sigma}_{22\cdot1}\mathbf{T}_{21})^{-1}|^{-n/2} \\ &= |\mathbf{I}_n + 4\mathbf{T}_{12}\boldsymbol{\Sigma}_{22\cdot1}\mathbf{T}_{21}\boldsymbol{\Sigma}_{11} - 2i(\mathbf{T}_{11}\boldsymbol{\Sigma}_{11} + \mathbf{T}_{12}\boldsymbol{\Sigma}_{21} + \boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\mathbf{T}_{21}\boldsymbol{\Sigma}_{11})|^{-n/2}. \quad \square \end{aligned}$$

If $k \leq n$ then $\mathbf{A}$ has a Wishart distribution and, consequently, the characteristic function of $\mathbf{A}$ is given by

$$\Psi(\mathbf{T}) = |\mathbf{I}_k - 2i\mathbf{T}\mathbf{\Sigma}|^{-n/2}. \tag{12}$$

Next, using the results of Theorem 1 we prove that (12) is also the expression of the characteristic function for the singular Wishart distribution.

The application of the partitioned matrices from (2) and (10), the fact that $\mathbf{T}_{22} = 0$ and Theorem 13.3.8 of Harville (1997) lead to

$$
\begin{aligned}
|\mathbf{I}_k - 2i\mathbf{T}\mathbf{\Sigma}|^{-n/2} &= |\mathbf{I}_n - 2i(\mathbf{T}_{11}\mathbf{\Sigma}_{11} + \mathbf{T}_{12}\mathbf{\Sigma}_{21})|^{-n/2} \\
&\quad \times |\mathbf{I}_{k-n} - 2i\mathbf{T}_{21}\mathbf{\Sigma}_{12} + 4\mathbf{T}_{21}\mathbf{\Sigma}_{11}(\mathbf{I}_n - 2i(\mathbf{T}_{11}\mathbf{\Sigma}_{11} + \mathbf{T}_{12}\mathbf{\Sigma}_{21}))^{-1} \\
&\quad \times (\mathbf{T}_{11}\mathbf{\Sigma}_{12} + \mathbf{T}_{12}\mathbf{\Sigma}_{22})|^{-n/2} \\
&= |\mathbf{I}_n - 2i(\mathbf{T}_{11}\mathbf{\Sigma}_{11} + \mathbf{T}_{12}\mathbf{\Sigma}_{21})|^{-n/2} \\
&\quad \times |\mathbf{I}_n - 2i\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{T}_{21}\mathbf{\Sigma}_{11} + 4(\mathbf{I}_n - 2i(\mathbf{T}_{11}\mathbf{\Sigma}_{11} + \mathbf{T}_{12}\mathbf{\Sigma}_{21}))^{-1} \\
&\quad \times (\mathbf{T}_{11}\mathbf{\Sigma}_{12} + \mathbf{T}_{12}\mathbf{\Sigma}_{22})\mathbf{T}_{21}\mathbf{\Sigma}_{11}|^{-n/2},
\end{aligned}
$$

where the last identity is obtained by using Sylvester's determinant theorem (see, e.g. Harville (1997, p. 416)). Hence,

$$
\begin{aligned}
|\mathbf{I}_k &- 2i\mathbf{T}\mathbf{\Sigma}|^{-n/2} \\
&= |(\mathbf{I}_n - 2i(\mathbf{T}_{11}\mathbf{\Sigma}_{11} + \mathbf{T}_{12}\mathbf{\Sigma}_{21}))(\mathbf{I}_n - 2i\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{T}_{21}\mathbf{\Sigma}_{11}) \\
&\quad + 4(\mathbf{T}_{11}\mathbf{\Sigma}_{12} + \mathbf{T}_{12}\mathbf{\Sigma}_{22})\mathbf{T}_{21}\mathbf{\Sigma}_{11}|^{-n/2} \\
&= |\mathbf{I}_n - 2i(\mathbf{T}_{11}\mathbf{\Sigma}_{11} + \mathbf{T}_{12}\mathbf{\Sigma}_{21} + \mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{T}_{21}\mathbf{\Sigma}_{11}) - 4\mathbf{T}_{11}\mathbf{\Sigma}_{12}\mathbf{T}_{21}\mathbf{\Sigma}_{11} \\
&\quad - 4\mathbf{T}_{12}\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{T}_{21}\mathbf{\Sigma}_{11} + 4\mathbf{T}_{11}\mathbf{\Sigma}_{12}\mathbf{T}_{21}\mathbf{\Sigma}_{11} + 4\mathbf{T}_{12}\mathbf{\Sigma}_{22}\mathbf{T}_{21}\mathbf{\Sigma}_{11}|^{-n/2} \\
&= |\mathbf{I}_n - 2i(\mathbf{T}_{11}\mathbf{\Sigma}_{11} + \mathbf{T}_{12}\mathbf{\Sigma}_{21} + \mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{T}_{21}\mathbf{\Sigma}_{11}) + 4\mathbf{T}_{12}\mathbf{\Sigma}_{22\cdot1}\mathbf{T}_{21}\mathbf{\Sigma}_{11}|^{-n/2},
\end{aligned}
$$

which coincides with the expression for the characteristic function of the singular Wishart distribution as presented in Theorem 1. We summarize the above results in the following corollary.

Corollary 1. *Under the conditions of Theorem 1, the characteristic function of the singular Wishart distribution is given by*

$$\Psi(\mathbf{T}) = |\mathbf{I}_k - 2i\mathbf{T}\mathbf{\Sigma}|^{-n/2}. \tag{13}$$

The application of the characteristic function (13) allows us to prove that the class of singular Wishart distributions is closed with respect to linear symmetric transformations given by $\mathbf{L}\mathbf{A}\mathbf{L}^T$. This is done in Theorem 2.

Theorem 2. *Let $\mathbf{A} \sim W_k(n, \mathbf{\Sigma})$, $k > n$, and $\mathbf{L}$ be a $p \times k$ arbitrary deterministic matrix with $\mathrm{rank}(\mathbf{L}) = p$. Then $\mathbf{L}\mathbf{A}\mathbf{L}^T \sim W_p(n, \mathbf{L}\mathbf{\Sigma}\mathbf{L}^T)$ if $p \leq n$ and $\mathbf{L}\mathbf{A}\mathbf{L}^T \sim W_p(n, \mathbf{L}\mathbf{\Sigma}\mathbf{L}^T)$, $p > n$, otherwise.*

Proof. The characteristic function of $\mathbf{L}\mathbf{A}\mathbf{L}^T$ is given by

$$
\begin{aligned}
\Psi(\mathbf{T}) &= \mathsf{E}\left[\mathrm{etr}\left(i\mathbf{T}^T\mathbf{L}\mathbf{A}\mathbf{L}^T\right)\right] = \mathsf{E}\left[\mathrm{etr}\left(i\mathbf{L}^T\mathbf{T}^T\mathbf{L}\mathbf{A}\right)\right] \\
&= \left|\mathbf{I}_k - 2i\mathbf{L}^T\mathbf{T}^T\mathbf{L}\mathbf{\Sigma}\right|^{-n/2} \\
&= \left|\mathbf{I}_p - 2i\mathbf{T}^T\mathbf{L}\mathbf{\Sigma}\mathbf{L}^T\right|^{-n/2},
\end{aligned}
\tag{14}
$$

where the second line follows from Corollary 1 and the third one is obtained by applying Sylvester's determinant theorem (see, e.g. Harville (1997, p. 416)). Now the result of Theorem 2 follows from the observation that the last line in (14) is the characteristic function of $W_p(n, \mathbf{L}\mathbf{\Sigma}\mathbf{L}^T)$ if $p \leq n$ and of $W_p(n, \mathbf{L}\mathbf{\Sigma}\mathbf{L}^T)$, $p > n$, otherwise. $\square$

A very important results is obtained in the special case when $p = 1$.

Theorem 3. *Let* $\mathbf{A} \sim W_k(n, \boldsymbol{\Sigma})$, $k > n$, *and* $\mathbf{y}$ *be any* $k \times 1$ *random vector which is independent of* $\mathbf{A}$. *Then* $\mathbf{y}^T \mathbf{A} \mathbf{y} / \mathbf{y}^T \boldsymbol{\Sigma} \mathbf{y} \sim \chi_n^2$ *and is independent of* $\mathbf{y}$.

Proof. Because $\mathbf{A}$ and $\mathbf{y}$ are independent, the conditional distribution of $\mathbf{y}^T \mathbf{A} \mathbf{y} / \mathbf{y}^T \boldsymbol{\Sigma} \mathbf{y}$ given $\mathbf{y} = \mathbf{y}^*$ coincides with the distribution of $\mathbf{y}_*^T \mathbf{A} \mathbf{y}_* / \mathbf{y}_*^T \boldsymbol{\Sigma} \mathbf{y}_*$. From Theorem 2 we get that

$$\mathbf{y}_*^T \mathbf{A} \mathbf{y}_* \sim W_1\left(n, \mathbf{y}_*^T \boldsymbol{\Sigma} \mathbf{y}_*\right),$$

which leads to

$$\frac{\mathbf{y}_*^T \mathbf{A} \mathbf{y}_*}{\mathbf{y}_*^T \boldsymbol{\Sigma} \mathbf{y}_*} \sim \chi_n^2. \tag{15}$$

Because the conditional distribution of $\mathbf{y}_*^T \mathbf{A} \mathbf{y}_* / \mathbf{y}_*^T \boldsymbol{\Sigma} \mathbf{y}_*$ is independent of $\mathbf{y}_*$, we get that (15) also determines the unconditional distribution of $\mathbf{y}^T \mathbf{A} \mathbf{y} / \mathbf{y}^T \boldsymbol{\Sigma} \mathbf{y}$. □

3. MAIN RESULTS

In this section the main findings are presented. Particularly, we derive the stochastic representation, the density and the characteristic function of a weighted product of a singular Wishart matrix and a normal random vector. Thus we extend the results of Bodnar and Okhrin (2011) and Bodnar *et al.* (2013) who considered a non-singular Wishart matrix only.

Let $\mathbf{A}$ be k-dimensional singular Wishart distributed with n degrees of freedom and covariance matrix $\boldsymbol{\Sigma}$, i.e. $\mathbf{A} \sim W_k(n, \boldsymbol{\Sigma})$, $k > n$. Let $\mathbf{z} \sim N_k(\boldsymbol{\mu}, \lambda \boldsymbol{\Sigma})$ (k-dimensional multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$) with $\lambda > 0$. Let $\mathbf{L}$ be a $p \times k$ matrix of constants with $\text{rank}(\mathbf{L}) = p$. We are interested in the distribution of $\mathbf{L} \mathbf{A} \mathbf{z}$ when $\mathbf{A}$ and $\mathbf{z}$ are independent.

Let $\widetilde{\mathbf{L}}_\mathbf{z} = (\mathbf{L}^T, \mathbf{z})^T$ and define $\mathbf{W}_\mathbf{z} = \widetilde{\mathbf{L}}_\mathbf{z} \mathbf{A} \widetilde{\mathbf{L}}_\mathbf{z}^T$. Because $\mathbf{A}$ and $\mathbf{z}$ are independent, we get that the distribution of $\mathbf{W}_\mathbf{z}$ given $\mathbf{z} = \mathbf{z}^*$ is the same as the distribution of $\mathbf{W}_{\mathbf{z}^*}$. Moreover, the application of Theorem 2 leads to

$$\mathbf{W}_{\mathbf{z}^*} \sim W_{p+1}\left(n, \widetilde{\mathbf{L}}_{\mathbf{z}^*} \boldsymbol{\Sigma} \widetilde{\mathbf{L}}_{\mathbf{z}^*}^T\right).$$

If $p + 1 \le n$ we get that $\mathbf{W}_{\mathbf{z}^*}$ follows a Wishart distribution (cf. Lemma 1). Hence, from Theorem 1 of Bodnar et al. (2013) the following stochastic representation of $\mathbf{L} \mathbf{A} \mathbf{z}$ we obtain

$$\begin{aligned}
\mathbf{L} \mathbf{A} \mathbf{z} \overset{d}{=} \; & \xi \left(\mathbf{L} \boldsymbol{\Sigma} \mathbf{L}^T\right)^{1/2} \mathbf{y}_1 \\
& + \sqrt{\xi} \left(\mathbf{L} \boldsymbol{\Sigma} \mathbf{L}^T\right)^{1/2} \left[\sqrt{\mathbf{y}_1^T \mathbf{y}_1 + \eta} \mathbf{I}_p - \frac{\sqrt{\mathbf{y}_1^T \mathbf{y}_1 + \eta} - \sqrt{\eta}}{\mathbf{y}_1^T \mathbf{y}_1} \mathbf{y}_1 \mathbf{y}_1^T\right] \mathbf{z}_0,
\end{aligned} \tag{16}$$

where $\xi \sim \chi_n^2$, $\mathbf{z}_0 \sim N_p(\mathbf{0}, \mathbf{I}_p)$, and

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} \sim N_k\left(\begin{pmatrix} \mathbf{S}_1 \boldsymbol{\Sigma}^{1/2} \boldsymbol{\mu} \\ \mathbf{S}_2 \boldsymbol{\Sigma}^{1/2} \boldsymbol{\mu} \end{pmatrix}, \lambda \begin{pmatrix} \mathbf{S}_1 \boldsymbol{\Sigma}^2 \mathbf{S}_1^T & \mathbf{S}_1 \boldsymbol{\Sigma}^2 \mathbf{S}_2^T \\ \mathbf{S}_2 \boldsymbol{\Sigma}^2 \mathbf{S}_1^T & \mathbf{S}_2 \boldsymbol{\Sigma}^2 \mathbf{S}_2^T \end{pmatrix}\right) \quad \text{with } \eta = \mathbf{y}_2^T \mathbf{y}_2;$$

ξ, $\mathbf{z}_0$ are independent of $\mathbf{y}$.

If $p + 1 > n$ then $\mathbf{W}_{\mathbf{z}^*}$ has a singular Wishart distribution. However, also in this case the stochastic representation (16) remains valid as it is proven in Theorem 4.

Theorem 4. *Let* $\mathbf{A} \sim W_k(n, \boldsymbol{\Sigma})$, $k > n$, *and let* $\mathbf{z} \sim N_k(\boldsymbol{\mu}, \lambda \boldsymbol{\Sigma})$ *with* $\lambda > 0$. *We assume that* $\boldsymbol{\Sigma}$ *is positive definite and that* $\mathbf{A}$ *and* $\mathbf{z}$ *are independent. Let* $\mathbf{L}$ *be a* $p \times k$ *matrix of*

constants with $\text{rank}(\mathbf{L}) = p < k$, *and let* $\mathbf{S}_1 = (\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T)^{-1/2}\mathbf{L}\boldsymbol{\Sigma}^{1/2}$, $\mathbf{S}_2 = (\mathbf{I}_p - \mathbf{S}_1^T\mathbf{S}_1)^{1/2}$
with $\mathbf{S}_2 : (k - p) \times k$. *Then the stochastic representation of* $\mathbf{LAz}$ *is given by*

$$\mathbf{LAz} \stackrel{d}{=} \xi \left(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T\right)^{1/2} \mathbf{y}_1$$

$$+ \sqrt{\xi} \left(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T\right)^{1/2} \left[\sqrt{\mathbf{y}_1^T\mathbf{y}_1 + \eta}\mathbf{I}_p - \frac{\sqrt{\mathbf{y}_1^T\mathbf{y}_1 + \eta} - \sqrt{\eta}}{\mathbf{y}_1^T\mathbf{y}_1}\mathbf{y}_1\mathbf{y}_1^T\right] \mathbf{z}_0, \quad (17)$$

where $\xi \sim \chi_n^2$, $\mathbf{z}_0 \sim N_p(\mathbf{0}, \mathbf{I}_p)$, *and*

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} \sim N_k \left(\begin{pmatrix} \mathbf{S}_1\boldsymbol{\Sigma}^{1/2}\boldsymbol{\mu} \\ \mathbf{S}_2\boldsymbol{\Sigma}^{1/2}\boldsymbol{\mu} \end{pmatrix}, \lambda \begin{pmatrix} \mathbf{S}_1\boldsymbol{\Sigma}^2\mathbf{S}_1^T & \mathbf{S}_1\boldsymbol{\Sigma}^2\mathbf{S}_2^T \\ \mathbf{S}_2\boldsymbol{\Sigma}^2\mathbf{S}_1^T & \mathbf{S}_2\boldsymbol{\Sigma}^2\mathbf{S}_2^T \end{pmatrix} \right) \text{ with } \eta = \mathbf{y}_2^T\mathbf{y}_2;$$

ξ, $\mathbf{z}_0$ *are independent of* $\mathbf{y}$.

Proof. The proof for $p + 1 \leq n$ is given above. Thus, we deal only with the case $p + 1 > n$ here.

Let $\widetilde{\mathbf{L}}_{\mathbf{z}^*}\mathbf{A}\widetilde{\mathbf{L}}_{\mathbf{z}^*}^T$ and $\widetilde{\mathbf{L}}_{\mathbf{z}^*}\boldsymbol{\Sigma}\widetilde{\mathbf{L}}_{\mathbf{z}^*}^T$ be partitioned as

$$\widetilde{\mathbf{L}}_{\mathbf{z}^*}\mathbf{A}\widetilde{\mathbf{L}}_{\mathbf{z}^*}^T = \begin{bmatrix} \mathbf{LAL}^T & \mathbf{LAz}^* \\ \mathbf{z}^{*T}\mathbf{AL}^T & \mathbf{z}^{*T}\mathbf{Az}^* \end{bmatrix} \quad \text{and} \quad \widetilde{\mathbf{L}}_{\mathbf{z}^*}\boldsymbol{\Sigma}\widetilde{\mathbf{L}}_{\mathbf{z}^*}^T = \begin{bmatrix} \mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T & \mathbf{L}\boldsymbol{\Sigma}\mathbf{z}^* \\ \mathbf{z}^{*T}\boldsymbol{\Sigma}\mathbf{L}^T & \mathbf{z}^{*T}\boldsymbol{\Sigma}\mathbf{z}^* \end{bmatrix}.$$

Using Lemma 1 the conditional distribution of $\mathbf{LAz}$ is given by

$$\mathbf{LAz} \mid \mathbf{z}^T\mathbf{Az}, \mathbf{z} = \mathbf{z}^* \sim N_p \left(\mathbf{L}\boldsymbol{\Sigma}\mathbf{z}^*\frac{\mathbf{z}^{*T}\mathbf{Az}^*}{\mathbf{z}^{*T}\boldsymbol{\Sigma}\mathbf{z}^*}, \mathbf{z}^{*T}\mathbf{Az}^* \left(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T - \frac{\mathbf{L}\boldsymbol{\Sigma}\mathbf{z}^*\mathbf{z}^{*T}\boldsymbol{\Sigma}\mathbf{L}^T}{\mathbf{z}^{*T}\boldsymbol{\Sigma}\mathbf{z}^*} \right) \right).$$

Let $\xi = \mathbf{z}^{*T}\mathbf{Az}^*/\mathbf{z}^{*T}\boldsymbol{\Sigma}\mathbf{z}^*$. Then ξ and $\mathbf{z}$ are independent and $\xi \sim \chi_n^2$ (see Theorem 3).
As a result, we get

$$\mathbf{LAz} \stackrel{d}{=} \xi\mathbf{L}\boldsymbol{\Sigma}\mathbf{z} + \sqrt{\xi} \left(\mathbf{z}^T\boldsymbol{\Sigma}\mathbf{z}\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T - \mathbf{L}\boldsymbol{\Sigma}\mathbf{z}\mathbf{z}^T\boldsymbol{\Sigma}\mathbf{L}^T \right)^{1/2} \mathbf{z}_0,$$

where $\xi \sim \chi_n^2$, $\mathbf{z}_0 \sim N_p(\mathbf{0}, \mathbf{I}_p)$ and $\mathbf{z} \sim N_k(\boldsymbol{\mu}, \lambda\boldsymbol{\Sigma})$; ξ, $\mathbf{z}_0$, $\mathbf{z}$ are independent.

The application of the equality

$$\left(\mathbf{F} - \mathbf{bb}^T \right)^{1/2} = \mathbf{F}^{1/2} \left(\mathbf{I}_p - c\mathbf{F}^{-1/2}\mathbf{bb}^T\mathbf{A}^{-1/2} \right),$$

with $c = \frac{1 - \sqrt{1 - \mathbf{b}^T\mathbf{A}^{-1}\mathbf{b}}}{\mathbf{b}^T\mathbf{A}^{-1}\mathbf{b}}$, $\mathbf{F} = \mathbf{z}^T\boldsymbol{\Sigma}\mathbf{z}\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T$, and $\mathbf{b} = \mathbf{L}\boldsymbol{\Sigma}\mathbf{z}$, leads to

$$\mathbf{LAz} \stackrel{d}{=} \xi\mathbf{L}\boldsymbol{\Sigma}\mathbf{z} + \sqrt{\xi} \left(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T\right)^{1/2}$$

$$\times \left(\sqrt{\mathbf{z}^T\boldsymbol{\Sigma}\mathbf{z}}\mathbf{I}_p - \frac{\sqrt{\mathbf{z}^T\boldsymbol{\Sigma}\mathbf{z}} - \sqrt{\mathbf{z}^T(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}^{1/2}\mathbf{Q}_1\boldsymbol{\Sigma}^{1/2})\mathbf{z}}}{\mathbf{z}^T\boldsymbol{\Sigma}^{1/2}\mathbf{Q}_1\boldsymbol{\Sigma}^{1/2}\mathbf{z}}\mathbf{S}_1\boldsymbol{\Sigma}^{1/2}\mathbf{z}\mathbf{z}^T\boldsymbol{\Sigma}^{1/2}\mathbf{S}_1^T \right) \mathbf{z}_0,$$

where $\mathbf{Q}_1 = \mathbf{S}_1^T\mathbf{S}_1$.

Using the facts that $\mathbf{Q}_1$ is a projection matrix with $\text{rank}(\mathbf{Q}_1) = p$ and $\mathbf{S}_1$ is a $p \times k$ matrix with $\text{rank}(\mathbf{S}_1) = p$ we get that $\text{rank}(\mathbf{I}_p - \mathbf{Q}_1) = k - p$ (see Harville (1997, Theorem 12.3.4)). As a result there exists the $(k - p) \times k$ matrix $\mathbf{S}_2 = (\mathbf{I}_p - \mathbf{Q}_1)^{1/2}$ such that $\mathbf{S}_2^T\mathbf{S}_2 = (\mathbf{I}_p - \mathbf{Q}_1)$ with $\text{rank}(\mathbf{S}_2) = k - p$. Making the transformation $\mathbf{y}_1 = \mathbf{S}_1\boldsymbol{\Sigma}^{1/2}\mathbf{z} \in \mathbb{R}^p$, $\mathbf{y}_2 = \mathbf{S}_2\boldsymbol{\Sigma}^{1/2}\mathbf{z} \in \mathbb{R}^{k-p}$ we get

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} \sim N_k \left(\begin{pmatrix} \mathbf{S}_1\boldsymbol{\Sigma}^{1/2}\boldsymbol{\mu} \\ \mathbf{S}_2\boldsymbol{\Sigma}^{1/2}\boldsymbol{\mu} \end{pmatrix}, \lambda \begin{pmatrix} \mathbf{S}_1\boldsymbol{\Sigma}^2\mathbf{S}_1^T & \mathbf{S}_1\boldsymbol{\Sigma}^2\mathbf{S}_2^T \\ \mathbf{S}_2\boldsymbol{\Sigma}^2\mathbf{S}_1^T & \mathbf{S}_2\boldsymbol{\Sigma}^2\mathbf{S}_2^T \end{pmatrix} \right).$$

Since $\mathbf{t}^T\mathbf{t} = \mathbf{t}^T\mathbf{Q}_1\mathbf{t} + \mathbf{t}^T\mathbf{S}_2^T\mathbf{S}_2\mathbf{t}$ for all $\mathbf{t} \in \mathbb{R}^k$, the stochastic representation of $\mathbf{LAz}$ is expressed as

$$\mathbf{LAz} \stackrel{d}{=} \xi \left(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T\right)^{1/2} \mathbf{y}_1 + \sqrt{\xi} \left(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T\right)^{1/2} \left[\sqrt{\mathbf{y}_1^T\mathbf{y}_1 + \eta}\mathbf{I}_p - \frac{\sqrt{\mathbf{y}_1^T\mathbf{y}_1 + \eta} - \sqrt{\eta}}{\mathbf{y}_1^T\mathbf{y}_1}\mathbf{y}_1\mathbf{y}_1^T\right] \mathbf{z}_0,$$

where $\eta = \mathbf{y}_2^T \mathbf{y}_2$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Next, we derive the density function of $\mathbf{LAz}$. Because $\mathbf{y}_2|\mathbf{y}_1 \sim N_{k-p}(\boldsymbol{\nu}, \lambda\boldsymbol{\Omega})$ with

$$\boldsymbol{\nu} = \mathbf{S}_2\boldsymbol{\Sigma}^{1/2}\boldsymbol{\mu} + \left(\mathbf{S}_2\boldsymbol{\Sigma}^2\mathbf{S}_1^T\right)\left(\mathbf{S}_1\boldsymbol{\Sigma}^2\mathbf{S}_1^T\right)^{-1}\left(\mathbf{y}_1 - \mathbf{S}_1\boldsymbol{\Sigma}^{1/2}\boldsymbol{\mu}\right)$$

and

$$\boldsymbol{\Omega} = \mathbf{S}_2\boldsymbol{\Sigma}^2\mathbf{S}_2^T - \mathbf{S}_2\boldsymbol{\Sigma}^2\mathbf{S}_1^T\left(\mathbf{S}_1\boldsymbol{\Sigma}^2\mathbf{S}_1^T\right)^{-1}\mathbf{S}_1\boldsymbol{\Sigma}^2\mathbf{S}_2^T,$$

the application of Theorem 4.2c.1 by Mathai and Provost (1992) leads to the density function of η expressed as

$$f_{\eta|\mathbf{y}_1}(y) = \sum_{i=0}^{\infty} c_i \frac{i!}{2\beta\Gamma\left(\frac{k-p}{2}+i\right)}\left(\frac{y}{2\beta}\right)^{\frac{k-p}{2}-1} e^{-y/2\beta} L_i^{\left(\frac{k-p}{2}-1\right)}\left(\frac{y}{2\beta}\right), \qquad y > 0,$$

$\beta > 0$ is an arbitrary constant,

$$L_i^{\left(\frac{k-p}{2}-1\right)}\left(\frac{y}{2\beta}\right) = \frac{(-\frac{y}{2\beta})^i}{i!} \; {}_2F_0\left(-i, -\frac{y}{2\beta} - i; -\frac{2\beta}{y}\right),$$

$$c_0 = 1, \qquad c_i = \frac{1}{i}\sum_{r=0}^{i-1} d_{i-r}c_r, \qquad d_j = \sum_{j_1=1}^{k-p}(1 - jb_{j_1}^2)(\alpha_{j_1})^{-j}, \qquad j \geq 1.$$

The matrix $\mathbf{V}$ is a $(k-p) \times (k-p)$ orthogonal matrix which diagonalizes $\lambda\boldsymbol{\Omega}$. That is,

$$\lambda\mathbf{V}^T\boldsymbol{\Omega}\mathbf{V} = \mathrm{diag}(\alpha_1, \ldots, \alpha_{k-p}), \qquad \mathbf{V}\mathbf{V}^T = \mathbf{I}_{k-p},$$

where $\alpha_1, \ldots, \alpha_{k-p}$ are eigenvalues of $\lambda\boldsymbol{\Omega}$ and $\mathbf{b} = \lambda^{-1/2}\mathbf{V}^T\boldsymbol{\Omega}^{-1/2}\boldsymbol{\nu}$.

Let $f_{\chi_n^2}$ denote the density function of the χ^2-distribution with n degrees of freedom and let $f_{N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})}$ be the density of the multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Then the density of $\mathbf{LAz}$ is expressed as in Corollary 2.

Corollary 2. *Let* $\mathbf{A} \sim W_k(n, \boldsymbol{\Sigma})$, $k > n$, $\mathbf{z} \sim N_k(\boldsymbol{\mu}, \lambda\boldsymbol{\Sigma})$ *with* $\lambda > 0$ *and* $\boldsymbol{\Sigma}$ *positive definite. Assume that* $\mathbf{A}$ *and* $\mathbf{z}$ *are independent. Let* $\mathbf{L}$ *be a* $p \times k$ *matrix of constants with* $\mathrm{rank}(\mathbf{L}) = p < k$, *and let* $\mathbf{S}_1 = (\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T)^{-1/2}\mathbf{L}\boldsymbol{\Sigma}^{1/2}$, $\mathbf{S}_2 = (\mathbf{I}_p - \mathbf{S}_1^T\mathbf{S}_1)^{1/2}$. *Then the density function of* $\mathbf{LAz}$ *is given by*

$$f_{\mathbf{LAz}}(\mathbf{x}) = \int_0^{\infty}\int_0^{\infty}\int_{\mathbb{R}^p} f_{N_p(\widetilde{\boldsymbol{\mu}},\widetilde{\boldsymbol{\Sigma}})}(\mathbf{x} \mid \xi = v, \mathbf{y}_1 = \mathbf{z}_1, \eta = z_2) f_{N_p(\breve{\boldsymbol{\mu}},\breve{\boldsymbol{\Sigma}})}(\mathbf{z}_1)$$
$$\times f_{\chi_n^2}(v) f_{\eta|\mathbf{y}_1}(z_2 \mid \mathbf{y}_1 = \mathbf{z}_1)\, \mathbf{dz}_1\, \mathbf{d}z_2\, \mathbf{d}v,$$

where $\widetilde{\boldsymbol{\mu}} = v(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T)^{1/2}\mathbf{z}_1$ *and* $\widetilde{\boldsymbol{\Sigma}} = v(\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T)^{1/2}[(\mathbf{z}_1^T\mathbf{z}_1 + z_2)\mathbf{I}_p - \mathbf{z}_1\mathbf{z}_1^T](\mathbf{L}\boldsymbol{\Sigma}\mathbf{L}^T)^{1/2}$, $\breve{\boldsymbol{\mu}} = \mathbf{S}_1\boldsymbol{\Sigma}^{1/2}\boldsymbol{\mu}$ *and* $\breve{\boldsymbol{\Sigma}} = \lambda\mathbf{S}_1\boldsymbol{\Sigma}^2\mathbf{S}_1^T$.

Proof. Using the stochastic representation of $\mathbf{LAz}$ (see Theorem 4), the conditional density function of $\mathbf{LAz}$ is given by

$$\mathbf{LAz} \mid \xi, \mathbf{y}_1, \eta \sim N_p(\widetilde{\boldsymbol{\mu}}, \widetilde{\boldsymbol{\Sigma}}).$$

Because the densities of ξ, $\mathbf{y}_1$ and η are known and the random variables are independently distributed, we obtain the unconditional density function of $\mathbf{LAz}$ by, first, constructing the joint density of $\mathbf{LAz}$, ξ, $\mathbf{y}_1$ and η, and, then, by integrating out two random variables η and ξ as well as the random vector $\mathbf{y}_1$. $\qquad\qquad$ □

In Corollary 3, we consider the special case of Corollary 2 when $\boldsymbol{\Sigma} = \mathbf{I}_k$. In particular, the density function of $\mathbf{LAz}$ is simplified significantly because the quadratic form η times λ^{-1} has a non-central chi-squared distribution, namely $\lambda^{-1}\eta \sim \chi_{k-p;\delta^2}^2$ with $\delta^2 = \lambda^{-1}\boldsymbol{\mu}^T\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2\boldsymbol{\mu}$ where $\widetilde{\mathbf{S}}_2 = (\mathbf{I}_p - \widetilde{\mathbf{S}}_1^T\widetilde{\mathbf{S}}_1)^{1/2}$ with $\widetilde{\mathbf{S}}_1 = (\mathbf{L}\mathbf{L}^T)^{-1/2}\mathbf{L}$.

In order to prove this statement we, first, point out that the matrix $\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2$ is an idempotent matrix, since $\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2 = \widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2$ where we use that $\mathbf{Q}_1$ is idempotent and $\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2 = \mathbf{I}_k - \mathbf{Q}_1$. Furthermore, it holds that

(i) $\mathrm{tr}[(\lambda^{-1}\mathbf{I}_{k-p})(\lambda\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T)] = \mathrm{tr}(\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T) = \mathrm{tr}(\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2) = \mathrm{tr}(\mathbf{I}_k - \mathbf{Q}_1) = k - \mathrm{tr}(\mathbf{Q}_1) = k - \mathrm{tr}(\widetilde{\mathbf{S}}_1\widetilde{\mathbf{S}}_1^T) = k - \mathrm{tr}\underbrace{[(\mathbf{LL}^T)^{-1/2}\mathbf{LL}^T(\mathbf{LL}^T)^{-1/2}]}_{=\mathbf{I}_p} = k - p;$

(ii)
$$(\lambda\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T)(\lambda^{-1}\mathbf{I}_{k-p})(\lambda\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T)(\lambda^{-1}\mathbf{I}_{k-p})(\lambda\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T) = \lambda\widetilde{\mathbf{S}}_2\underbrace{\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2}_{=\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2}\widetilde{\mathbf{S}}_2^T$$
$$= (\lambda\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T)(\lambda^{-1}\mathbf{I}_{k-p})(\lambda\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T);$$

(iii)
$$(\boldsymbol{\mu}^T\widetilde{\mathbf{S}}_2^T)(\lambda^{-1}\mathbf{I}_{k-p})(\lambda\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T)(\lambda^{-1}\mathbf{I}_{k-p})(\widetilde{\mathbf{S}}_2\boldsymbol{\mu}) = \lambda^{-1}\boldsymbol{\mu}^T\underbrace{\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2}_{=\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2}\boldsymbol{\mu}$$
$$= (\boldsymbol{\mu}^T\widetilde{\mathbf{S}}_2^T)(\lambda^{-1}\mathbf{I}_{k-p})(\widetilde{\mathbf{S}}_2\boldsymbol{\mu});$$

(iv)
$$[\boldsymbol{\mu}^T\widetilde{\mathbf{S}}_2^T][(\lambda^{-1}\mathbf{I}_{k-p})(\lambda\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T)]^2 = \lambda^{-1}\boldsymbol{\mu}^T\underbrace{\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2}_{=\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2}\widetilde{\mathbf{S}}_2^T$$
$$= (\boldsymbol{\mu}^T\widetilde{\mathbf{S}}_2^T)(\lambda^{-1}\mathbf{I}_{k-p})(\lambda\widetilde{\mathbf{S}}_2\widetilde{\mathbf{S}}_2^T).$$

Now, the application of Theorem 5.1.3 by Mathai and Provost (1992) shows that $\lambda^{-1}\eta$ is $\chi^2_{k-p;\delta^2}$-distributed with $\delta^2 = \lambda^{-1}\boldsymbol{\mu}^T\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2\boldsymbol{\mu}$.

Corollary 3. *Let* $\mathbf{A} \sim W_k(n, \mathbf{I}_k)$, $k > n$ *and* $\mathbf{z} \sim N_k(\boldsymbol{\mu}, \lambda\mathbf{I}_k)$ *with* $\lambda > 0$. *Assume that* $\mathbf{A}$ *and* $\mathbf{z}$ *are independent. Let* $\mathbf{L}$ *be a* $p \times k$ *matrix of constants, with* $\mathrm{rank}(\mathbf{L}) = p < k$, *and let* $\widetilde{\mathbf{S}}_1 = (\mathbf{LL}^T)^{-1/2}\mathbf{L}$, $\widetilde{\mathbf{S}}_2 = (\mathbf{I}_p - \widetilde{\mathbf{S}}_1^T\widetilde{\mathbf{S}}_1)^{1/2}$. *Then the density of* $\mathbf{LAz}$ *is given by*

$$f_{\mathbf{LAz}}(\mathbf{x}) = \frac{1}{\lambda}\int_{\mathbb{R}^p}\int_0^\infty\int_0^\infty f_{N_p(\widetilde{\boldsymbol{\mu}}_1,\widetilde{\boldsymbol{\Sigma}}_1)}(\mathbf{x} \mid \xi = v, \mathbf{y}_1 = \mathbf{z}_1, \eta = z_2)f_{N_p(\widetilde{\mathbf{S}}_1\boldsymbol{\mu},\lambda\widetilde{\mathbf{S}}_1^T\widetilde{\mathbf{S}}_1)}(\mathbf{z}_1)$$
$$\times f_{\chi^2_n}(v)f_{\chi^2_{k-p;\delta^2}}\left(\lambda^{-1}z_2 \mid \mathbf{y}_1 = \mathbf{z}_1\right)\,\mathbf{dz}_1\,\mathbf{d}z_2\,\mathbf{d}v,$$

where $\widetilde{\boldsymbol{\mu}}_1 = v(\mathbf{LL}^T)^{1/2}\mathbf{z}_1$, $\widetilde{\boldsymbol{\Sigma}}_1 = v(\mathbf{LL}^T)^{1/2}\left[(\mathbf{z}_1^T\mathbf{z}_1 + \eta)\mathbf{I}_p - \mathbf{z}_1\mathbf{z}_1^T\right](\mathbf{LL})^{1/2}$, *and* $\delta^2 = \lambda^{-1}\boldsymbol{\mu}^T\widetilde{\mathbf{S}}_2^T\widetilde{\mathbf{S}}_2\boldsymbol{\mu}$.

Proof. From Theorem 4 the stochastic representation of $\mathbf{LAz}$ is expressed as

$$\mathbf{LAz} \overset{d}{=} \xi\left(\mathbf{LL}^T\right)^{1/2}\mathbf{y}_1 + \sqrt{\xi}\left(\mathbf{LL}^T\right)^{1/2}\left[\sqrt{\mathbf{y}_1^T\mathbf{y}_1 + \eta}\mathbf{I}_p - \frac{\sqrt{\mathbf{y}_1^T\mathbf{y}_1 + \eta} - \sqrt{\eta}}{\mathbf{y}_1^T\mathbf{y}_1}\mathbf{y}_1\mathbf{y}_1^T\right]\mathbf{z}_0,$$

where $\mathbf{z}_0 \sim N_p(\mathbf{0}, \mathbf{I}_p)$, $\xi \sim \chi^2_n$,

$$\mathbf{y} = \begin{pmatrix}\mathbf{y}_1 \\ \mathbf{y}_2\end{pmatrix} \sim N_k\left(\begin{pmatrix}\mathbf{S}_1\boldsymbol{\mu} \\ \mathbf{S}_2\boldsymbol{\mu}\end{pmatrix}, \lambda\begin{pmatrix}\mathbf{S}_1\mathbf{S}_1^T & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_2\mathbf{S}_2^T\end{pmatrix}\right) \text{ with } \eta = \mathbf{y}_2^T\mathbf{y}_2;$$

$\mathbf{z}_0, \xi,$ and $\mathbf{y}$ are independently distributed. Using a fact that the covariance matrix of $\mathbf{y}$ is block diagonal it holds that $\mathbf{y}_1$ and $\mathbf{y}_2$ are independently distributed. Moreover, η is a function of $\mathbf{y}_2$ only and, hence, $\mathbf{y}_1$ and η are independent as well. Finally, using $\lambda^{-1}\eta \sim \chi^2_{k-p;\delta^2}$ and the stochastic representation (17) we get the statement of Corollary 2. $\square$

In the next corollary, we consider a special case of Theorem 4 and Corollary 2 with $p = 1$ and $\mathbf{L} = \mathbf{l}^T$. As a result we obtain that the density function of $\mathbf{l}^T\mathbf{Az}$ is given

by three integrals only as well as known univariate density functions. This stochastic representation simplifies it significantly.

Corollary 4. *Let* $\mathbf{A} \sim W_k(n, \boldsymbol{\Sigma})$, $k > n$, *and* $\mathbf{z} \sim N_k(\boldsymbol{\mu}, \lambda \boldsymbol{\Sigma})$ *with* $\lambda > 0$. *Let* $\boldsymbol{\Sigma}$ *be positive definite. Assume that* $\mathbf{A}$ *and* $\mathbf{z}$ *are independent and* $p + 1 > n$. *Let* $\mathbf{l}$ *be a k-dimensional vector of constants and let* $\mathbf{S}_1 = (\mathbf{l}^T \boldsymbol{\Sigma} \mathbf{l})^{-1/2} \mathbf{l}^T \boldsymbol{\Sigma}^{1/2}$, $\mathbf{S}_2 = (\mathbf{I} - \mathbf{S}_1^T \mathbf{S}_1)^{1/2}$. *Then*

(a) *the stochastic representation of* $\mathbf{l}^T \mathbf{A} \mathbf{z}$ *is given by*

$$\mathbf{l}^T \mathbf{A} \mathbf{z} \stackrel{d}{=} \left(\mathbf{l}^T \boldsymbol{\Sigma} \mathbf{l} \right)^{1/2} \left[\xi y_1 - \sqrt{\eta \xi} z_0 \right],$$

where $\xi \sim \chi_n^2$, $z_0 \sim N(0, 1)$,

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} \sim N_k \left(\begin{pmatrix} \mathbf{S}_1 \boldsymbol{\Sigma}^{1/2} \boldsymbol{\mu} \\ \mathbf{S}_2 \boldsymbol{\Sigma}^{1/2} \boldsymbol{\mu} \end{pmatrix}, \lambda \begin{pmatrix} \mathbf{S}_1 \boldsymbol{\Sigma}^2 \mathbf{S}_1^T & \mathbf{S}_1 \boldsymbol{\Sigma}^2 \mathbf{S}_2^T \\ \mathbf{S}_2 \boldsymbol{\Sigma}^2 \mathbf{S}_1^T & \mathbf{S}_2 \boldsymbol{\Sigma}^2 \mathbf{S}_2^T \end{pmatrix} \right) \quad \text{with } \eta = \mathbf{y}_2^T \mathbf{y}_2,$$

and ξ, z_0 *are independent of* $\mathbf{y}$.

(b) *the density of* $\mathbf{l}^T \mathbf{A} \mathbf{z}$ *is given by*

$$f_{\mathbf{l}^T \mathbf{A} \mathbf{z}}(x) = \int_{-\infty}^{\infty} \int_0^{\infty} \int_0^{\infty} f_{N((\mathbf{l}^T \boldsymbol{\Sigma} \mathbf{l})^{1/2} v z_1, (\mathbf{l}^T \boldsymbol{\Sigma} \mathbf{l}) v z_2)}(x \mid \xi = v, y_1 = z_1, \eta = z_2)$$

$$\times f_{N(\mathbf{S}_1 \boldsymbol{\Sigma}^{1/2} \boldsymbol{\mu}, \lambda \mathbf{S}_1 \boldsymbol{\Sigma}^2 \mathbf{S}_1^T)}(z_1) f_{\chi_1^2}(v) f_\eta(z_2 \mid y_1 = z_1) \, dz_1 \, dz_2 \, dv.$$

There are several important applications of the results presented in Corollary 4. One of them is the derivation of the characteristic function of $\mathbf{A} \mathbf{z}$ whose analytical expression is provided in Theorem 5. Another application is discussed in Section 4.

Theorem 5. *Let* $\mathbf{A} \sim W_k(n, \boldsymbol{\Sigma})$, $k > n$, *and* $\mathbf{z} \sim N_k(\boldsymbol{\mu}, \lambda \boldsymbol{\Sigma})$ *with* $\lambda > 0$ *and* $\boldsymbol{\Sigma}$ *is positive definite. Assume that* $\mathbf{A}$ *and* $\mathbf{z}$ *are independent. Let* $\mathbf{S}_1 = (\mathbf{L} \boldsymbol{\Sigma} \mathbf{L}^T)^{-1/2} \mathbf{L} \boldsymbol{\Sigma}^{1/2}$, $\mathbf{S}_2 = (\mathbf{I}_p - \mathbf{S}_1^T \mathbf{S}_1)^{1/2}$. *Then the characteristic function of* $\mathbf{A} \mathbf{z}$ *is given by*

$$\phi_{\mathbf{A} \mathbf{z}}(\mathbf{t}) = \int_{-\infty}^{\infty} \int_0^{\infty} \int_0^{\infty} \left(1 + \left(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t} \right) z_2 - 2i \left(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t} \right)^{1/2} z_1 \right)^{-k/2}$$

$$\times f_{N(\mathbf{S}_1 \boldsymbol{\Sigma}^{1/2} \boldsymbol{\mu}, \lambda \mathbf{S}_1 \boldsymbol{\Sigma}^2 \mathbf{S}_1^T)}(z_1) f_\eta(z_2 \mid y_1 = z_1) \, dz_1 \, dz_2.$$

Proof. The characteristic function of $\mathbf{A} \mathbf{z}$ for $\mathbf{t} \in \mathbb{R}^k$ is given by

$$\phi_{\mathbf{A} \mathbf{z}}(\mathbf{t}) = \mathsf{E} \left[\exp \left(i \mathbf{t}^T \mathbf{A} \mathbf{z} \right) \right].$$

Let $\zeta = \mathbf{t}^T \mathbf{A} \mathbf{z}$. Then applying Corollary 4 and integrating out v and ξ we get

$$\phi_{\mathbf{A} \mathbf{z}}(\mathbf{t}) = \mathsf{E} \left[\exp \left(i \mathbf{t}^T \mathbf{A} \mathbf{z} \right) \right] = \mathsf{E}[\exp(i\zeta)]$$

$$= \int_{-\infty}^{\infty} \int_0^{\infty} \int_0^{\infty} \int_{-\infty}^{\infty} \exp(iw) f_{N((\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t})^{1/2} v z_1, (\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}) v z_2)}(w \mid \xi = v, y_1 = z_1, \eta = z_2)$$

$$\times f_{N(\mathbf{S}_1 \boldsymbol{\Sigma}^{1/2} \boldsymbol{\mu}, \lambda \mathbf{S}_1 \boldsymbol{\Sigma}^2 \mathbf{S}_1^T)}(z_1) f_{\chi_1^2}(v) f_\eta(z_2 \mid y_1 = z_1) \, dz_1 \, dz_2 \, dv \, dw$$

$$= \int_{-\infty}^{\infty} \int_0^{\infty} \int_0^{\infty} \exp \left\{ i \left(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t} \right)^{1/2} v z_1 - \frac{1}{2} \left(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t} \right) v z_2 \right\} f_{\chi_1^2}(v)$$

$$\times f_{N(\mathbf{S}_1 \boldsymbol{\Sigma}^{1/2} \boldsymbol{\mu}, \lambda \mathbf{S}_1 \boldsymbol{\Sigma}^2 \mathbf{S}_1^T)}(z_1) f_\eta(z_2 \mid y_1 = z_1) \, dz_1 \, dz_2 \, dv$$

$$= \frac{1}{2^{k/2} \Gamma(k/2)} \int_{-\infty}^{\infty} \int_0^{\infty} \int_0^{\infty} v^{k/2 - 1} \exp \left\{ i \left(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t} \right)^{1/2} v z_1 - \frac{v}{2} \left(1 + z_2 \mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t} \right) \right\}$$

$$\times f_{N(\mathbf{S}_1 \boldsymbol{\Sigma}^{1/2} \boldsymbol{\mu}, \lambda \mathbf{S}_1 \boldsymbol{\Sigma}^2 \mathbf{S}_1^T)}(z_1) f_\eta(z_2 \mid y_1 = z_1) \, dz_1 \, dz_2 \, dv$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(1 + \left(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t} \right) z_2 \right)^{-k/2} \left(1 - 2i \frac{(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t})^{1/2} z_1}{(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}) z_2 + 1} \right)^{-k/2}$$

$$\times f_{N(\mathbf{S}_1 \boldsymbol{\Sigma}^{1/2} \boldsymbol{\mu}, \lambda \mathbf{S}_1 \boldsymbol{\Sigma}^2 \mathbf{S}_1^T)}(z_1) f_\eta(z_2 \mid y_1 = z_1) \, dz_1 \, dz_2,$$

where the last integral is obtained from the expression of the characteristic function for the gamma distribution with shape $k/2$ and scale $(1 + (\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}) z_2)/2$ at $(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t})^{1/2} z_1$. Finally, simplifying the expression under the integral we get the expression provided in the statement of Theorem 5. $\qquad\square$

It is remarkable to note that the characteristic function for $\mathbf{Az}$ is given as a two-dimensional integral only with the known univariate density functions.

4. Numerical illustration

In this section we present the results of a simulation study. The aim is to compare the kernel density estimators with each other calculated for two data sets where the first one is obtained by generating singular Wishart matrices and normal vectors, whereas the elements from the second data are obtained by using the stochastic representation of Corollary 4. We put $p = 1$, $\mathbf{l}^T = (1/n, 0, \ldots, 0)$, $\boldsymbol{\mu}^T = (1, \ldots, k)$, $\boldsymbol{\Sigma} = \mathbf{I}_k$, $\lambda = 1/n$ and compare the result for several values of $k \in \{50, 100\}$ and $n \in \{1, 20, 40\}$. Each of the simulated data consists of $N = 10^4$ independent realizations which are used to fit the corresponding kernel density estimators with the Gaussian kernel. The bandwidth parameters are determined via cross-validation for every sample.

The first data set is generated directly from the product $\mathbf{l}^T \mathbf{Az}$. It corresponds to the abbreviation 'data 1' in the notations of Figure 1 and it is obtained in the following way

    a) generate independently $\mathbf{A} \sim W_k(n, \mathbf{I}_k)$, $k > n$ and $\mathbf{z} \sim N_k(\boldsymbol{\mu}, \lambda \mathbf{I}_k)$;
    b) compute $\mathbf{l}^T \mathbf{Az}$;
    c) repeat a)–b) N times.

The second data set is simulated using the results of Corollary 4 and it is denoted by 'data 2' in this section. The corresponding algorithm is given next:

    a) generate independently $z_0 \sim N(0, 1)$, $\xi \sim \chi_n^2$, $y_1 \sim N\big(\mathbf{l}^T \boldsymbol{\mu}/\sqrt{\mathbf{l}^T \mathbf{l}}, \lambda\big)$, and $\eta \sim \lambda \chi_{k-1;\delta^2}^2$ with $\delta^2 = \lambda^{-1} \big(\boldsymbol{\mu}^T \boldsymbol{\mu} - (\mathbf{l}^T \boldsymbol{\mu})^2/\mathbf{l}^T \mathbf{l}\big)$;
    b) compute

$$\mathbf{l}^T \mathbf{Az} = \sqrt{\mathbf{l}^T \mathbf{l}} \left(\xi y_1 - \sqrt{\eta \xi} z_0 \right); \tag{18}$$

    c) repeat a)–b) N times.

It is noted that the second algorithm is more computationally efficient than the first one since only four random variables instead of $k + k(k+1)/2$ are generated within each repetition. Moreover, the number of the simulated random variables is independent of k and, hence, even in large dimensions, for example in case of $k = 100$, only four random variables have to be simulated. Finally, since no matrix appears in (18), the expression for $\mathbf{l}^T \mathbf{Az}$ can be easily vectorized which provides an additional increase in the efficiency of the second procedure with respect to the first one.

In Figure 1, we present the kernel density estimators computed from both data sets for several values of k and n. It is noted that the obtained density functions almost coincide for all considered values of k and n. Some minor differences are present only around the picks of the densities. Furthermore, we observe that all density functions are slightly skewed to the right.

5. Summary

Wishart and normal distributions are the most frequently applied families of probability distributions in both statistics and probability theory with a number of important applications in finance, economics, biology, etc. Although many results have already been established for each of these distributions separately as well as for quadratic forms involving these distributional classes, other types of combinations are not deeply studied
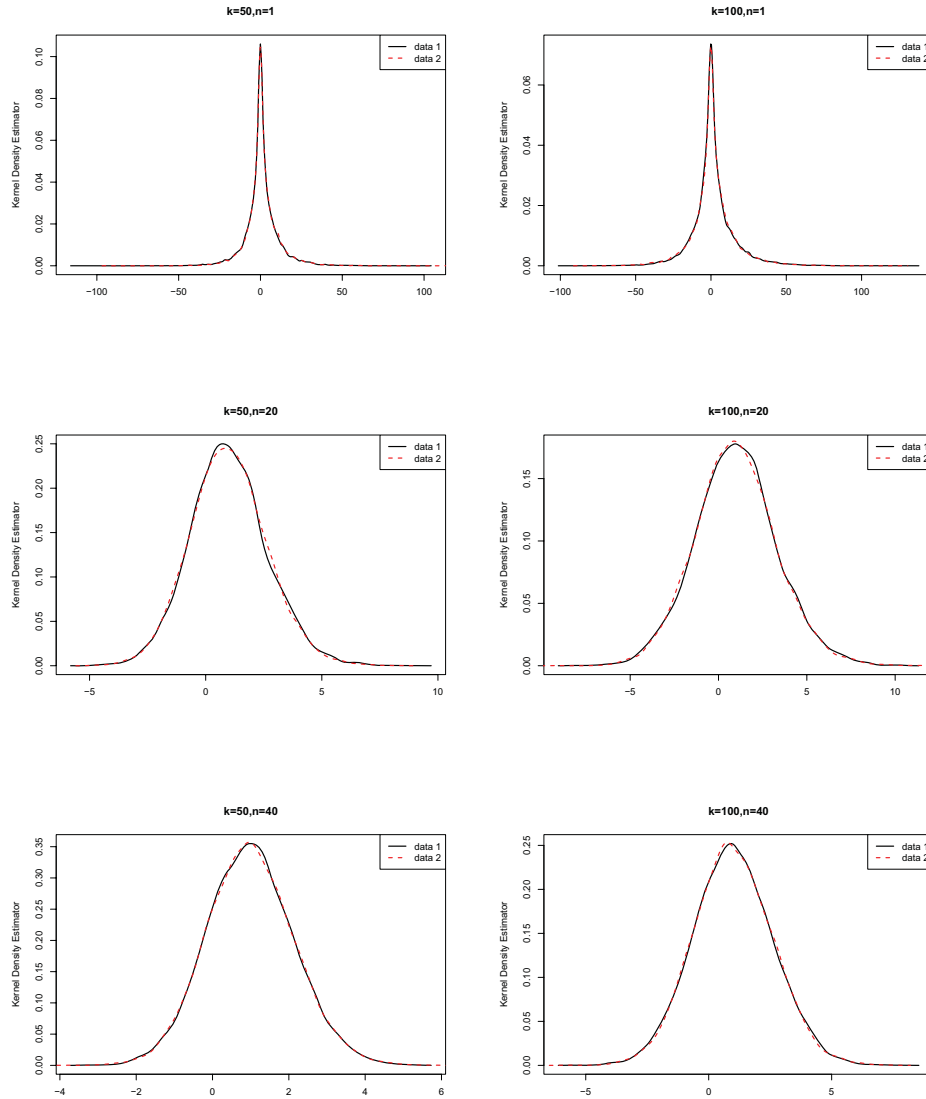
FIGURE 1. Kernel density estimators for the two considered simulation procedures with $k \in \{50, 100\}$ and $n \in \{1, 20, 40\}$.

in the literature. This topic is even less analyzed when the singular Wishart distribution in the combination with the normal distribution is the goal of investigation.

In this paper, we derive several distributional properties of the product of a singular Wishart matrix with a normal vector. A very useful stochastic representation of the product is established which is then used in the derivation of the density function as well as of the characteristic function. Moreover, the obtained results increases significantly the efficiency of the numerical studies involving such a product by reducing the computational time substantially. Finally, several interesting an important properties of the singular Wishart distribution are derived.

References

1. S. A. Andersson and G. G. Wojnar, *Wishart distribution on homogeneous cones*, Journal of Theoretical Probability **17** (2004), 781–818.
2. T. Bodnar and Y. Okhrin, *Properties of the partitioned singular, inverse and generalized Wishart distributions*, Journal of Multivariate Analysis **99** (2008), 2389–2405.
3. T. Bodnar and Y. Okhrin, *On the product of inverse Wishart and normal distributions with applications to discriminant analysis and portfolio theory*, Scandinavian Journal of Statistics **38** (2011), 311–331.
4. T. Bodnar, S. Mazur, and Y. Okhrin, *On the exact and approximate distributions of the product of a Wishart matrix with a normal vector*, Journal of Multivariate Analysis **122** (2013), 70–81.
5. M. Drton, H. Massam, and I. Olkin, *Moments of minors of Wishart matrices*, Annals of Statistics **36** (2008), 2261–2283.
6. D. Geiger and D. Heckerman, *A characterization of the bivariate Wishart distribution*, Probability and Mathemathical Statistics **18** (1998), 119–131.
7. D. Geiger and D. Heckerman, *Parameter priors for directed acyclic graphical models and the characterization of several probability distributions*, Annals of Statistics **30** (2002), 1412–1440.
8. A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*, Chapman and Hall/CRC, Boca Raton, 2000.
9. A. K. Gupta, T. Varga, and T. Bodnar, *Elliptically Contoured Models in Statistics and Portfolio Theory*, Springer, New York, 2013.
10. D. A. Harville, *Matrix Algebra from Statistician's Perspective*, Springer, New York, 1997.
11. C. G. Khatri, *On the mutual independence of certain statistics*, Annals of Mathematical Statistics **30** (1959), 1258–1262.
12. H. Massam and J. Weselowski, *The Matsumoto–Yor property and the structure of the Wishart distribution*, Journal of Multivariate Analysis **97** (2006), 103–123.
13. A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables*, Marcel Dekker, Inc, New York, 1992.
14. I. Olkin and S. N. Roy, *On multivariate distribution theory*, Annals of Mathematical Statistics **25** (1954), 329–339.
15. I. Olkin and H. Rubin, *Multivariate beta distribution and independency properties of Wishart distribution*, Annals of Mathematical Statistics **35** (1964), 261–269.
16. A. C. Rencher, *Methods of Multivariate Analysis*, A Wiley-Interscience publication, 2002.
17. M. S. Srivastava, *Singular Wishart and multivariate beta distributions*, Annals of Statistics **31** (2003), 1537–1560.

Department of Mathematics, Humboldt-University of Berlin, Unter den Linden 6, D-10099 Berlin, Germany
  E-mail address: `bodnar@math.hu-berlin.de`

Department of Statistics, Lund University, PO Box 743, SE-22007 Lund, Sweden
  E-mail address: `stepan.mazur@stat.lu.se`

Department of Statistics, University of Augsburg, Universitaetsstrasse 16, D-86159 Augsburg, Germany
  E-mail address: `yarema.okhrin@wiwi.uni-augsburg.de`