WORKING PAPER 7/2022 (STATISTICS)

Matrix Variate Generalized Laplace Distributions

Tomasz J. Kozubowski, Stepan Mazur and Krysztof Podgorski

ISSN 1403-0586

Örebro University School of Business
SE-701 82 Örebro, Sweden

MATRIX VARIATE GENERALIZED LAPLACE DISTRIBUTIONS

TOMASZ J. KOZUBOWSKI, STEPAN MAZUR, AND KRZYSZTOF PODGÓRSKI

Abstract. The generalized asymmetric Laplace (GAL) distribution, also known as the *variance/mean-gamma* model, is a popular flexible class of distributions that can account for peakedness, skewness, and heavier than normal tails, often observed in financial or other empirical data. We consider extensions of the GAL distribution to the matrix variate case, which arise as covariance mixtures of matrix variate normal distributions. Two different mixing mechanisms connected with the nature of the random scaling matrix are considered, leading to what we term matrix variate GAL distributions of Type I and II. While Type I matrix variate GAL distribution has been studied before, there is no comprehensive account of Type II in the literature, except for their rather brief treatment as a special case of matrix variate generalized hyperbolic distributions. With this work we fill this gap, and present an account for basic distributional properties of Type II matrix variate GAL distributions. In particular, we derive their probability density function and the characteristic function, as well as provide stochastic representations related to matrix variate gamma distribution. We also show that this distribution is closed under linear transformations, and study the relevant marginal distributions. In addition, we also briefly account for Type I and discuss the connections with Type II. We hope that this work will be useful in the areas where matrix variate distributions provide an appropriate probabilistic tool for three-way or, more generally, panel data sets, which can arise across different applications.

1. Introduction

Mixtures of univariate and multivariate normal distributions play a prominent role in statistical theory and practice, with numerous applications across a variety of fields. Indeed, a continuous-type normal variance-mean mixtures of the form

$$\mathbf{X} \overset{d}{=} \boldsymbol{\mu} + W\mathbf{m} + \sqrt{W}\mathbf{Z}, \tag{1.1}$$

2020 *Mathematics Subject Classification.* Primary 62H10, 60E05; Secondary 60E10.

Key words and phrases. Covariance mixture of Gaussian distributions, distribution theory, generalized Laplace distribution, MatG distribution, matrix variate distribution, matrix variate gamma distribution, matrix gamma-normal distribution, matrix variate t distribution, normal variance-mean mixture, variance gamma distribution.

Stepan Mazur acknowledges financial support from the internal research grants at Örebro University and from the project "Models for macro and financial economics after the financial crisis" (Dnr: P18-0201) funded by Jan Wallander and Tom Hedelius foundation.

Krzysztof Podgórski acknowledges financial support of the Swedish Research Council (VR) Grant DNR: 2020-05168.

where $\boldsymbol{\mu}$ and $\mathbf{m}$ are deterministic vectors in $\mathbb{R}^k$, $\mathbf{Z}$ has a zero-mean multivariate normal distribution $\mathcal{N}_k(\mathbf{0}, \boldsymbol{\Sigma})$ given by the covariance matrix $\boldsymbol{\Sigma}$, and W is a non-negative random variable independent of $\mathbf{Z}$, provide very popular modeling tools in numerous fields (see, e.g., [4]). In particular, the *generalized asymmetric Laplace* (GAL) distribution (see, e.g., [27]), also known as *variance/mean-gamma* model (see, e.g., [31]), defined by (1.1) with gamma distributed W, is a popular flexible model that can account for peakedness, skewness, and heavier than normal tails, often observed in financial or other empirical data.

In this paper, we consider extensions of the GAL distribution to the *matrix variate case*, where the $\mathbf{Z}$ in (1.1) is a *Gaussian random matrix* while W may be either a one-dimensional non-negative random variable or can itself be a random (square) non-negative definite matrix, playing the role of a *stochastic covariance matrix (Gramian).* These two different mixing mechanisms lead to what we term *matrix variate GAL distributions of Type I and II.* While Type I matrix variate GAL distribution has been studied before (see [39]), there is no comprehensive account of Type II in the literature, except for their rather brief treatment as a special case of *matrix variate generalized hyperbolic distributions*, studied in [35] (see also [23]). With this work we fill this gap, and present an account for basic distributional properties of Type II matrix variate GAL distributions. In particular, we derive their probability density function and the characteristic function, and provide stochastic representations related to the matrix variate gamma distribution. We also show that this distribution is closed under linear transformations, and study the relevant marginal distributions. In addition, we also briefly account for Type I and discuss the interconnections with Type II.

We hope that this work will be useful in the areas where matrix variate distributions provide an appropriate probabilistic tool. One of the most straightforward applications of such a model is for the panel data, which are so common in econometrics. For example, one can consider the monthly stock and bond returns of the 23 OECD countries that results in 23×2 panel data observed over time. Such data typically are regressed against some control variables, for example, exchange rates, inflation, interest rates, etc. In building such models, most of the time, the normal matrix variate errors are assumed, which are rarely justified by goodness-of-fit analysis (see, e.g., [3]). To improve the performance of such models, one can consider mixtures of matrix valued normal distributions instead. The models presented in this work are perfectly fitted for the purpose.

In other areas of application such panel structures are referred to as *three-way* data (see, e.g., [1], [16], [17], [18], [19], [33], [37], [38]). Indeed, as discussed in [37] among others, such data can arise as longitudinal data on multiple response variables, spacial multivariate data, spatio-temporal data (measurements on a unit at different time periods and locations), when objects are rated on multiple attributes by multiple experts, or symbolic data (complex information structured as multiple values for each variable). As matrix variate models that can account for skewness are particularly desirable (see, e.g, [33]), matrix variate GAL distributions (which have this feature) may provide a useful tool here as well, along with other similar mixtures of matrix normal distributions, such as matrix variate t-distribution, which arises via (1.1) where W has (univariate or matrix variate) inverse gamma distribution. The latter is indeed quite a popular model in this set-up, both with scalar scaling (see, e.g., [15], [16], [17], [18], [19]) as well as matrix scaling (see, e.g., [2], [6], [12], [20], [23], [26], [36]).

Our paper is structured as follows. In Section 2 we consider a univariate scaling mixtures of matrix variate distributions, and provide two general results for their densities and characteristic functions, which generalize particular results from the literature and lead to the special case of Type I matrix variate GAL distributions. The basic distributional properties of the latter are accounted for here as well. In Section 3, we discuss the case of matrix scaling of matrix variate normal distributions, where we provide several general results about this construction before providing specific details connected with Type II matrix variate GAL distributions. Sections 4 and 5 provide further results and illustrative examples connected with the special case of vector valued distributions, along with some open problems. Section 6 summarizes the paper.

Notation. Let us explain the notation that will be used throughout the paper. The quantity $\mathbf{I}_k$ shall denote the $k \times k$ dimensional identity matrix, while $\mathbf{0}$ shall stand for the matrix of zeros (of suitable dimension). We shall use $\mathrm{etr}\{\mathbf{C}\}$ to denote the function $\mathbf{C} \to \exp(\mathrm{tr}(\mathbf{C}))$, where $\mathrm{tr}(\mathbf{C})$ is the trace of a (square) matrix $\mathbf{C}$. The notation $\stackrel{d}{=}$ shall stand for the equality in distribution, and $\otimes$ shall denotes the Kronecker product. Further, the notation $\mathbf{C} \geq 0$ shall indicate that the square matrix $\mathbf{C}$ is non-negative definite, while for positive definite square matrices we shall write $\mathbf{C} > 0$, indicating that $\mathbf{C} \geq 0$ and $|\mathbf{C}| > 0$, where $|\mathbf{C}|$ is the determinant of the matrix $\mathbf{C}$. The set of all $k \times k$ (symmetric) positive definite matrices shall be denoted by $\mathbb{S}_k^+$. This set is closed under the addition and multiplication by a positive scalar, so it constitutes a cone. The set of all non-negative definite matrices of given size forms a cone as well, which is the closure $\overline{\mathbb{S}_k^+}$.

2. Univariate scale mixtures of matrix variate normal distributions

Consider univariate scaling mixtures of $k \times n$ matrix variate distributions

$$\mathbf{X} \stackrel{d}{=} W\mathbf{M} + \sqrt{W}\mathbf{Z} \tag{2.1}$$

with independent W and $\mathbf{Z}$, where $\mathbf{M}$ is a $k \times n$ deterministic matrix, W is a random variable on $\mathbb{R}_+$, and $\mathbf{Z} \sim \mathcal{MN}_{k,n}(\mathbf{0}, \mathbf{\Sigma} \otimes \mathbf{\Psi})$. Then, the right-hand-side in (2.1) describes the distribution of $\mathbf{Z}(W)$, where $\{\mathbf{Z}(u),\ \ u \in \mathbb{R}^+\}$ is a Lévy process built upon the matrix variate normal distribution $\mathcal{MN}_{k,n}(\mathbf{M}, \mathbf{\Sigma} \otimes \mathbf{\Psi})$. The latter describes the distribution of $\mathbf{Z}(1)$, and, in non-singular case, is given by the probability density function (PDF)

$$f(\mathbf{X}) = (2\pi)^{-\frac{nk}{2}} |\mathbf{\Sigma}|^{-\frac{n}{2}} |\mathbf{\Psi}|^{-\frac{k}{2}} \mathrm{etr}\left\{ -\frac{1}{2}\mathbf{\Sigma}^{-1}(\mathbf{X}-\mathbf{M})\mathbf{\Psi}^{-1}(\mathbf{X}-\mathbf{M})^\top \right\}, \quad \mathbf{X} \in \mathbb{R}^{k\times n}, \tag{2.2}$$

and the characteristic function (ChF)

$$\varphi(\mathbf{T}) = \mathbb{E}(\mathrm{etr}\{\iota \mathbf{Z}\mathbf{T}^\top\}) = \mathrm{etr}\left\{\iota \mathbf{T}^\top \mathbf{M} - \frac{1}{2}\mathbf{T}^\top \mathbf{\Sigma}\mathbf{T}\mathbf{\Psi}\right\}, \quad \mathbf{T} \in \mathbb{R}^{k\times n}, \tag{2.3}$$

where $\mathbf{\Sigma}$ $(k \times k)$ and $\mathbf{\Psi}$ $(n \times n)$ are positive-definite scale matrices (see, e.g., [21]).

Basic characteristics and properties of such a random matrix can be easily established via a standard conditioning argument, using the fact that, given $W = w$, $\mathbf{X}$ in (2.1) has matrix variate normal distribution $\mathcal{MN}_{k,n}(w\mathbf{M}, w\mathbf{\Sigma}\otimes\mathbf{\Psi})$, which is the same distribution as that of $\mathbf{Z}(w)$. In particular, according to (2.3), the ChF of this conditional distribution takes on the form

$$\varphi_{\mathbf{Z}(w)}(\mathbf{T}) = \left[\varphi_{\mathbf{Z}(1)}(\mathbf{T})\right]^w = \mathrm{etr}\left\{\iota \mathbf{T}^\top (w\mathbf{M}) - \frac{1}{2}\mathbf{T}^\top (w\mathbf{\Sigma})\mathbf{T}\mathbf{\Psi}\right\}, \quad \mathbf{T} \in \mathbb{R}^{k\times n}. \tag{2.4}$$

This immediately leads to the following result, which provides the ChF of $\mathbf{X}$.

Theorem 2.1. *The ChF of* $\mathbf{X}$ *in (2.1) is of the form*

$$\varphi_{\mathbf{X}}(\mathbf{T}) = \psi_W[-\log \varphi_{\mathbf{Z}}(\mathbf{T})], \tag{2.5}$$

where $\psi_W(\cdot)$ *is the Laplace transform (LT) of* W *and* $\varphi_{\mathbf{Z}}(\cdot)$ *is the ChF of* $\mathbf{Z}$*, given by the right-hand-side of (2.3).*

Similar arguments involving density functions lead to the result below, which provides a general expression for the PDF of $\mathbf{X}$.

Theorem 2.2. *If the distribution of* $\mathbf{X}$ *in (2.1) is non-singular, then its PDF is of the form*

$$f_{\mathbf{X}}(\mathbf{X}) = \frac{\mathrm{etr}\left\{\mathbf{\Sigma}^{-1}\mathbf{X}\mathbf{\Psi}^{-1}\mathbf{M}^\top\right\}}{(2\pi)^{kn/2}|\mathbf{\Sigma}|^{n/2}|\mathbf{\Psi}|^{k/2}}\mathbb{E}\left\{W^{-kn/2}e^{-\frac{1}{2}(cW + h(\mathbf{X})/W)}\right\}, \tag{2.6}$$

where

$$c = \mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{M}\mathbf{\Psi}^{-1}\mathbf{M}^\top\right) \quad \text{and} \quad h(\mathbf{X}) = \mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{X}\mathbf{\Psi}^{-1}\mathbf{X}^\top\right). \tag{2.7}$$

Proof. The proof is straightforward. □

Remark 2.1. We can obtain another compact representation of the ChF of $\mathbf{X}$, this time in terms of the product of the LT $\psi_W(\cdot)$ of W and the ChF $\varphi_{W_{\mathbf{T}}}(\cdot)$ of its *exponentially tilted* version $W_{\mathbf{T}}$,

$$\varphi_{\mathbf{X}}(\mathbf{T}) = \psi_W\left(\frac{1}{2}\mathrm{tr}(\mathbf{T}^\top\mathbf{\Sigma}\mathbf{T}\mathbf{\Psi})\right) \times \varphi_{W_{\mathbf{T}}}(\mathrm{tr}(\mathbf{M}\mathbf{T}^\top)), \tag{2.8}$$

where the laws of $W_{\mathbf{T}}$ and W, $V_{W_{\mathbf{T}}}$ and V_W, respectively, are related as follows:

$$V_{W_{\mathbf{T}}}(dw) = \frac{e^{-\frac{1}{2}\mathrm{tr}(\mathbf{T}^\top\mathbf{\Sigma}\mathbf{T}\mathbf{\Psi})w}}{\psi_W(\frac{1}{2}\mathrm{tr}(\mathbf{T}^\top\mathbf{\Sigma}\mathbf{T}\mathbf{\Psi}))}V_W(dw). \tag{2.9}$$

This result, which can be established by simple algebra, will be extended to the case of $\mathbf{X}$ obtained by matrix-scaling in (2.1) in the next section.

The following result provides the mean matrix and the covariance of $\mathbf{X}$ that follows the mixture representation (2.1). Here, the covariance matrix $\mathbb{V}ar(\mathbf{X})$ of a $k \times n$ random matrix $\mathbf{X}$ is understood as the covariance matrix of the $kn \times 1$ random vector $\mathrm{vec}(\mathbf{X})$.

Theorem 2.3. *If the relevant moments exist, then the mean matrix and the covariance of* $\mathbf{X}$ *in (2.1) are given by* $\mathbb{E}(\mathbf{X}) = \mathbb{E}(W)\mathbf{M}$ *and*

$$\mathbb{V}ar(\mathbf{X}) = \mathbb{V}ar(W)\text{vec}(\mathbf{M})[\text{vec}(\mathbf{M})]^\top + \mathbb{E}(W)\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}, \tag{2.10}$$

respectively.

Proof. Since $\mathbb{E}(\mathbf{Z}) = \mathbf{0}$ and W and $\mathbf{Z}$ are independent, we have

$$\mathbb{E}(\mathbf{X}) = \mathbb{E}(W\mathbf{M} + \sqrt{W}\mathbf{Z}) = \mathbb{E}(W)\mathbf{M} + \mathbb{E}(\sqrt{W})\mathbb{E}(\mathbf{Z}) = \mathbb{E}(W)\mathbf{M},$$

as desired. For the variance, first note that, by law of iterated variance formula, we have

$$\mathbb{V}ar(\mathbf{X}) = \mathbb{V}ar(\text{vec}(\mathbf{X})) = \mathbb{V}ar[\mathbb{E}(\text{vec}(\mathbf{X})|W)] + \mathbb{E}[\mathbb{V}ar(\text{vec}(\mathbf{X})|W)]. \tag{2.11}$$

Since given W, vec$(\mathbf{X})$ has a multivariate normal distribution on $\mathbb{R}^{kn}$ with mean vector $W\text{vec}(\mathbf{M})$ and covariance matrix $W\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$, we have $\mathbb{E}(\text{vec}(\mathbf{X})|W) = W\text{vec}(\mathbf{M})$ and $Var(\text{vec}(\mathbf{X})|W) = W\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$. When we subsitute these in (2.11), followed by simple algebra, we obtain the result. □

Models of this form with different choices of the distribution of W are very popular in the literature, and incorporate matrix variate analogs of various multivariate (and univariate) location-scale mixtures of normal distributions. A particularly convenient choice for the distribution of W is a generalized inverse Gaussian (GIG) distribution, along with its special cases of gamma, inverse gamma, and inverse Gaussian distributions, as in this case the integration in (2.6) required to calculate the PDF of $\mathbf{X}$ is straightforward due to the special structure of the GIG PDF. This has led to the development of matrix variate generalized hyperbolic distributions and their special cases of matrix variate generalized Laplace (variance gamma), matrix variate t, and matrix variate normal inverse Gaussian (NIG) distributions in this scheme (see, e.g., [15], [16], [17], [18], [19], [22], and [39]). Matrix variate analogs of Cauchy and slash distributions have also been developed along these lines (see [10] and [22], respectively). It should be noted that an additional matrix variate location parameter is often incorporated into these models as well.

Remark 2.2. As noted by several authors, we need a restriction on either $\boldsymbol{\Sigma}$ or $\boldsymbol{\Psi}$ (such as $\text{tr}(\boldsymbol{\Sigma}) = k$ or $\text{tr}(\boldsymbol{\Psi}) = n$, discussed in [16]), to ensure identifiability of the model (since for any $c \in \mathbb{R}_+$ the Kronecker product $\boldsymbol{\Sigma} \otimes \boldsymbol{\Psi}$ is unchanged when the two matrices are scaled by c and $1/c$, respectively).

2.1. Type I matrix variate generalized Laplace distributions.

If W in (2.1) has a standard gamma distribution $\mathcal{G}(\alpha)$ with scale parameter $\alpha > 0$, given by the PDF

$$f_W(w) = \frac{1}{\Gamma(\alpha)} w^{\alpha-1} e^{-w}, \;\; w \in \mathbb{R}^+, \tag{2.12}$$

and the LT

$$\varphi_W(t) = (1+t)^{-\alpha}, \;\; t \in \mathbb{R}^+, \tag{2.13}$$

we obtain a matrix variate version of generalized Laplace distribution, studied in [39]. To distinguish these distributions from those that arise via matrix scale mixtures of matrix variate normal distributions, which are treated in Section 3, we shall refer to them as *Type I matrix variate generalized Laplace distributions* and denote them by $\mathcal{MGAL}_{k,n}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{\Psi}, \alpha)$, in analogy with the notation of [27] and in the spirit of [39].

Definition 2.1. A random $k \times n$ matrix $\mathbf{X}$ is said to have MGAL distribution with shape parameter $\alpha > 0$, $k \times n$ matrix skewness parameter $\mathbf{M}$, and matrix scaling parameters $\mathbf{\Sigma} \in \mathbb{S}_k^+$, $\mathbf{\Psi} \in \mathbb{S}_n^+$, denoted by $\mathcal{MGAL}_{k,n}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{\Psi}, \alpha)$, if $\mathbf{X}$ admits the stochastic representation (2.1), where $W \sim \mathcal{G}(\alpha)$ is independent of $\mathbf{Z} \sim \mathcal{MN}_{k,n}(\mathbf{0}, \mathbf{\Sigma} \otimes \mathbf{\Psi})$.

A straightforward application of Theorem 2.1 produces the ChF of this distribution,

$$\varphi_{\mathbf{X}}(\mathbf{T}) = \left(1 - \iota \mathrm{tr}(\mathbf{T}^\top \mathbf{M}) + \frac{1}{2}\mathrm{tr}(\mathbf{T}^\top \mathbf{\Sigma} \mathbf{T} \mathbf{\Psi})\right)^{-\alpha}, \quad \mathbf{T} \in \mathbb{R}^{k \times n}, \tag{2.14}$$

which in [39] was derived by different arguments. To obtain the corresponding PDF, we use Theorem 2.2 with $W \sim \mathcal{G}(\alpha)$. The expectation in (2.6) reduces to

$$\frac{1}{\Gamma(\alpha)} \int_0^\infty w^{\alpha - kn/2 - 1} e^{-\frac{1}{2}[(2+c)w + h(\mathbf{X})/w]} dw \tag{2.15}$$

with c and $h(\mathbf{X})$ as in (2.7). By relating the integrand in (2.15) to the PDF of the GIG distribution $\mathcal{GIG}(\lambda, a, b)$ with parameters $\lambda = \alpha - kn/2$, $a = 2 + c$, and $b = h(\mathbf{X})$, which is given by

$$f(w) = \frac{(a/b)^{\lambda/2} w^{\lambda - 1}}{2K_\lambda(\sqrt{ab})} e^{-\frac{1}{2}[aw + b/w]}, \quad w \in \mathbb{R}^+, \tag{2.16}$$

followed by routine algebra, we obtain the PDF of $\mathbf{X} \sim \mathcal{MGAL}_{k,n}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{\Psi}, \alpha)$,

$$\begin{aligned} f(\mathbf{X}) \quad &= \quad \frac{2\mathrm{etr}\left\{\mathbf{\Sigma}^{-1}\mathbf{X}\mathbf{\Psi}^{-1}\mathbf{M}^\top\right\}}{(2\pi)^{kn/2}\Gamma(\alpha)|\mathbf{\Sigma}|^{n/2}|\mathbf{\Psi}|^{k/2}} \left(\frac{\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{X}\mathbf{\Psi}^{-1}\mathbf{X}^\top)}{2 + \mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{M}\mathbf{\Psi}^{-1}\mathbf{M}^\top)}\right)^{\frac{1}{2}(\alpha - kn/2)} \\ &\times K_{\alpha - kn/2}\left(\sqrt{(2 + \mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{M}\mathbf{\Psi}^{-1}\mathbf{M}^\top))\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{X}\mathbf{\Psi}^{-1}\mathbf{X}^\top)}\right). \end{aligned}$$

Remark 2.3. Since $\mathbf{Z}$ is matrix variate normal $\mathcal{MN}_{k,n}(\boldsymbol{M}, \mathbf{\Sigma} \otimes \mathbf{\Psi})$ if and only if $\mathrm{vec}(\mathbf{Z})$ is multivariate normal $\mathcal{N}_{kn}(\mathrm{vec}(\mathbf{M}), \mathbf{\Psi} \otimes \mathbf{\Sigma})$, it is clear from (2.1) that $\mathbf{X}$ is $\mathcal{MGAL}_{k,n}(\boldsymbol{M}, \mathbf{\Sigma}, \mathbf{\Psi}, \alpha)$ if and only if $\mathrm{vec}(\mathbf{X})$ is a multivariate $\mathcal{GAL}_{kn}(\mathrm{vec}(\mathbf{M}), \mathbf{\Psi} \otimes \mathbf{\Sigma})$ random vector in the notation of [27] (cf., [39]). Similar interpretation applies to any matrix variate model defined via (2.1), which can be seen as a matrix re-formulation of a multivariate normal mean-variance mixture distribution on $\mathbb{R}^{kn}$. In particular, if $n = 1$, where the matrix $\mathbf{M}$ in (2.1) reduces to a column vector $\mathbf{m}$ of size k and the scale matrix $\mathbf{\Psi}$ reduces to a scalar (which we set to 1 to avoid identifiability issues), we obtain a multivariate GAL (column) random vector on $\mathbb{R}^k$ that follows $\mathcal{GAL}_k(\mathbf{m}, \mathbf{\Sigma}, \alpha)$ distribution in the notation of [27]. Here, the ChF (2.14) reduces to that given in eq. (6.2.1) in [27] while the corresponding PDF turns into the GAL PDF given by eq. (6.9.3) in [27]. Similarly, for $k = 1$ we obtain an n-dimensional GAL (row) random vector that follows $\mathcal{GAL}_n(\mathbf{m}, \mathbf{\Psi}, \alpha)$ distribution (where we set the scale 1×1 matrix $\mathbf{\Sigma}$ to 1 to avoid identifiability issues).

Remark 2.4. A matrix-variate analog of multivariate asymmetric Laplace (AL) distribution, studied in [39], arises in this set-up when we take $\alpha = 1$, so that the variable

W in (2.1) has a standard exponential distribution. Basic properties of both, matrix variate Laplace and generalized Laplace distributions, can be obtained in a straightforward way, either directly from their PDF/ChF provided above, or from the properties of matrix normal distribution and standard conditioning arguments via (2.1). In particular, since for $W \sim \mathcal{G}(\alpha)$ we have $\mathbb{E}(W) = \mathbb{V}ar(W) = \alpha$, in view of Theorem 2.3 we obtain the expectation and the variance of $\mathbf{X} \sim \mathcal{MGAL}_{k,n}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{\Psi}, \alpha)$, which are given by $\mathbb{E}(\mathbf{X}) = \alpha\mathbf{M}$ and $\mathbb{V}ar(\mathbf{X}) = \alpha(\mathbf{\Psi} \otimes \mathbf{\Sigma} + [\text{vec}(\mathbf{M})][\text{vec}(\mathbf{M})]^\top)$, respectively, in complete analogy with the GAL distribution studied in [27]. In addition, the distribution of $\mathbf{X}$ remains in the same class under either left or right (or both, left and right) matrix multiplications. More details can be found in [39].

Further properties of MGAL distributions, extending those of univariate and multivariate Laplace and generalized Laplace distributions, can be developed as well. In particular, using the special structure of the MGAL ChF with $\alpha = 1$ one can show that the matrix variate Laplace distributions are *geometric infinitely divisible* (GID). In analogy with the vector case, we say that a matrix variate distribution of $\mathbf{X}$ is GID if for each $p \in (0, 1)$ we have the following equality in distribution:

$$\mathbf{X} \stackrel{d}{=} \mathbf{X}_1^{(p)} + \cdots + \mathbf{X}_{N_p}^{(p)}, \tag{2.17}$$

where N_p is a geometric random variable with parameter p (representing the number of trials) and the $\{\mathbf{X}_i^{(p)}\}$ are some IID random matrices, independent of N_p. A straightforward calculation involving MGAL ChF and conditioning on N_p shows that matrix variate Laplace distributions are indeed GIG, and the new result below provides the distribution of the $\{\mathbf{X}_i^{(p)}\}$ in (2.17).

Theorem 2.4. *Let $\mathbf{X} \sim \mathcal{MGAL}_{k,n}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{\Psi}, \alpha)$ where $\alpha = 1$. Then the distribution of $\mathbf{X}$ is GID and (2.17) holds with $\mathbf{X}_i^{(p)} \sim \mathcal{MGAL}_{k,n}(p\mathbf{M}, p\mathbf{\Sigma}, \mathbf{\Psi}, \alpha)$ where $\alpha = 1$.*

Proof. The proof is straightforward. □

One can also show that MGAL random matrices $\mathbf{X}$ where either $\mathbf{M}$ or $\mathbf{Z}$ in (2.1) is equal to $\mathbf{0}$, are strictly *matrix geometric stable* (MGS), so that we have the following equality in distribution for each $p \in (0, 1)$:

$$\mathbf{X} \stackrel{d}{=} a_p(\mathbf{X}_1 + \cdots + \mathbf{X}_{N_p}), \tag{2.18}$$

where N_p is a geometric variable with mean $1/p$, $a_p \in \mathbb{R}_+$, and the $\{\mathbf{X}_i\}$ are IID copies of $\mathbf{X}$, independent of N_p. The following result, which is an extension of the well-known properties of stability of multivariate Laplace distributions with respect to geometric summation (see, e.g., [27], Theorem 6.10.2, p. 259), provides more details regarding this.

Theorem 2.5. *Let $\mathbf{X} \sim \mathcal{MGAL}_{k,n}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{\Psi}, \alpha)$ where $\alpha = 1$. If either $\mathbf{M} = \mathbf{0}$ or one of $\mathbf{\Sigma}$, $\mathbf{\Psi}$ is $\mathbf{0}$, then the distribution of $\mathbf{X}$ is MGS and (2.18) holds for each $p \in (0, 1)$, with either $a_p = \sqrt{p}$ (if $\mathbf{M} = \mathbf{0}$) or $a_p = p$ (if $\mathbf{\Sigma} = \mathbf{0}$ or $\mathbf{\Psi} = \mathbf{0}$).*

Proof. The proof is straightforward. □

3. Matrix scale mixtures of matrix normal distributions

The random matrix models obtained in the previous section have one structural deficiency. Namely, their matrix form is simply obtained by rearrangement of vector variables. This is because the matrix valued normal variables are such and scaling them by a random value does not impose any truly advanced matrix structure. The model discussed in this section is intrinsically matrix variate, i.e. it cannot be simply explained by its vectorization. As seen in what follows, it is just opposite, i.e. analyzing its purely matrix variate character makes all the properties easy to understand.

Consider a model analogous to (2.1) but with a *matrix scaling*, where the univariate random variable W in (2.1) is replaced by a symmetric random matrix $\mathbf{\Gamma} \in \mathbb{S}_n^+$, leading to a $k \times n$ random matrix $\mathbf{X}$ with the stochastic representation

$$\mathbf{X} \stackrel{d}{=} \mathbf{\Gamma}\mathbf{A} + \mathbf{\Gamma}^{1/2}\mathbf{Z}\mathbf{\Sigma}^{1/2}, \tag{3.1}$$

where $\mathbf{Z} \sim \mathcal{MN}_{k,n}(\mathbf{0}, \mathbf{I}_k \otimes \mathbf{I}_n)$, $\mathbf{A}$ is a $k \times n$ matrix skewness parameter, and $\mathbf{\Sigma} \in \mathbb{S}_n^+$ is a matrix scale parameter. As in the case of univariate scaling, the basic characteristics of this random matrix can be established through simple conditioning as well.

Let us start with the ChF. Clearly, since $\mathbf{\Gamma}$ and $\mathbf{Z}$ in (3.1) are independently distributed, we have

$$\mathbf{X}|\mathbf{\Gamma} \sim \mathcal{MN}_{k,n}\left(\mathbf{\Gamma}\mathbf{A}, \mathbf{\Gamma} \otimes \mathbf{\Sigma}\right), \tag{3.2}$$

so that, by (2.3) and standard properties of the trace operator, we have

$$\varphi_{\mathbf{X}|\mathbf{\Gamma}}(\mathbf{T}) = \operatorname{etr}\left\{\iota\mathbf{A}\mathbf{T}^\top\mathbf{\Gamma} - \frac{1}{2}\mathbf{T}\mathbf{\Sigma}\mathbf{T}^\top\mathbf{\Gamma}\right\}.$$

Taking the expectation of the above expression with respect to the distribution of $\mathbf{\Gamma}$ leads the unconditional ChF of $\mathbf{X}$, which can be expressed in terms of the LT of $\mathbf{\Gamma}$ and the ChF of a certain exponentially tilted version of the distribution of $\mathbf{\Gamma}$, in analogy with (2.8) - (2.9). The details are given in the result below, which can be proven by simple algebra.

Theorem 3.1. *Let $\mathbf{X}$ be given by (3.1), where $\mathbf{Z} \sim \mathcal{MN}_{k,n}(\mathbf{0}, \mathbf{I}_k \otimes \mathbf{I}_n)$ and $\mathbf{\Gamma} > 0$ is a random square matrix with law $V_{\mathbf{\Gamma}}$, independent of $\mathbf{Z}$. Then the ChF of $\mathbf{X}$ is of the form*

$$\varphi_{\mathbf{X}}(\mathbf{T}) = \psi_{\mathbf{\Gamma}}\left(\frac{1}{2}\mathbf{T}\mathbf{\Sigma}\mathbf{T}^\top\right)\varphi_{\mathbf{\Gamma}_{\mathbf{T}}}(\mathbf{A}\mathbf{T}^\top), \tag{3.3}$$

where $\psi_{\mathbf{\Gamma}}(\cdot)$ is the LT of $\mathbf{\Gamma}$ and $\varphi_{\mathbf{\Gamma}_{\mathbf{T}}}(\cdot)$ is the ChF of an exponentially tilted version of $\mathbf{\Gamma}$, whose law $V_{\mathbf{\Gamma}_{\mathbf{T}}}$ is given by

$$V_{\mathbf{\Gamma}_{\mathbf{T}}}(d\mathbf{Y}) = \frac{\operatorname{etr}\left\{-\frac{1}{2}\mathbf{T}\mathbf{\Sigma}\mathbf{T}^\top\mathbf{Y}\right\}}{\psi_{\mathbf{\Gamma}}\left(\frac{1}{2}\mathbf{T}\mathbf{\Sigma}\mathbf{T}^\top\right)}V_{\mathbf{\Gamma}}(d\mathbf{Y}). \tag{3.4}$$

Proof. The proof is straightforward. □

Similar arguments involving the PDFs lead to the result below, which is an analog of Theorem 2.2 for matrix scaling.

Theorem 3.2. *If the distribution of* $\mathbf{X}$ *in (3.1) is non-singular, then its PDF is of the form*

$$f_{\mathbf{X}}(\mathbf{X}) = \frac{\mathrm{etr}\{\mathbf{X}\mathbf{\Sigma}^{-1}\mathbf{A}^\top\}}{(2\pi)^{kn/2}|\mathbf{\Sigma}|^{k/2}}\mathbb{E}\left(|\mathbf{\Gamma}|^{-n/2}\mathrm{etr}\left\{-\frac{1}{2}\left(\mathbf{A}\mathbf{\Sigma}^{-1}\mathbf{A}^\top\mathbf{\Gamma} + \mathbf{X}\mathbf{\Sigma}^{-1}\mathbf{X}^\top\mathbf{\Gamma}^{-1}\right)\right\}\right). \tag{3.5}$$

Proof. The result follows by standard conditioning arguments and properties of the trace operator. □

Let us also account for the mean and the variance of this distribution.

Theorem 3.3. *If the relevant moments exist, then the mean matrix and the covariance of* $\mathbf{X}$ *in (3.1) are given by* $\mathbb{E}(\mathbf{X}) = \mathbb{E}(\mathbf{\Gamma})\mathbf{A}$ *and*

$$\mathbb{V}ar(\mathbf{X}) = (\mathbf{A}^\top \otimes \mathbf{I}_k)\mathbb{V}ar(\mathbf{\Gamma})(\mathbf{A} \otimes \mathbf{I}_k) + \mathbf{\Sigma} \otimes \mathbb{E}(\mathbf{\Gamma}), \tag{3.6}$$

respectively.

Proof. Since $\mathbb{E}(\mathbf{Z}) = \mathbf{0}$ and $\mathbf{\Gamma}$ and $\mathbf{Z}$ are independent, we have

$$\mathbb{E}(\mathbf{X}) = \mathbb{E}(\mathbf{\Gamma}\mathbf{A} + \mathbf{\Gamma}^{1/2}\mathbf{Z}\mathbf{\Sigma}^{1/2}) = \mathbb{E}(\mathbf{\Gamma})\mathbf{A} + \mathbb{E}(\mathbf{\Gamma}^{1/2})\mathbb{E}(\mathbf{Z})\mathbf{\Sigma}^{1/2} = \mathbb{E}(\mathbf{\Gamma})\mathbf{A},$$

as desired. Next, by law of iterated variance formula, we have

$$\mathbb{V}ar(\mathbf{X}) = \mathbb{V}ar(\mathrm{vec}(\mathbf{X})) = \mathbb{V}ar[\mathbb{E}(\mathrm{vec}(\mathbf{X})|\mathbf{\Gamma})] + \mathbb{E}[\mathbb{V}ar(\mathrm{vec}(\mathbf{X})|\mathbf{\Gamma})]. \tag{3.7}$$

By (3.2), given $\mathbf{\Gamma}$ the random vector $\mathrm{vec}(\mathbf{X})$ has a multivariate normal distribution on $\mathbb{R}^{kn}$ with mean vector $\mathrm{vec}(\mathbf{\Gamma}\mathbf{A})$ and covariance matrix $\mathbf{\Sigma} \otimes \mathbf{\Gamma}$, leading to

$$\mathbb{E}(\mathrm{vec}(\mathbf{X})|\mathbf{\Gamma}) = \mathrm{vec}(\mathbf{\Gamma}\mathbf{A}) \ \text{ and } \ \mathbb{V}ar(\mathrm{vec}(\mathbf{X})|\mathbf{\Gamma}) = \mathbf{\Sigma} \otimes \mathbf{\Gamma}.$$

Thus, in view of (3.7), we have

$$\mathbb{V}ar(\mathbf{X}) = \mathbb{V}ar[\mathrm{vec}(\mathbf{\Gamma}\mathbf{A})] + \mathbb{E}[\mathbf{\Sigma} \otimes \mathbf{\Gamma}] = \mathbb{V}ar[\mathrm{vec}(\mathbf{\Gamma}\mathbf{A})] + \mathbf{\Sigma} \otimes \mathbb{E}(\mathbf{\Gamma}). \tag{3.8}$$

To conclude the proof, note that by the algebraic property $\mathrm{vec}(\mathbf{ABC}) = \left(\mathbf{C}^\top \otimes \mathbf{A}\right)\mathrm{vec}\left(\mathbf{B}\right)$, we have that

$$\mathrm{vec}(\mathbf{\Gamma}\mathbf{A}) = \mathrm{vec}(\mathbf{I}_k\mathbf{\Gamma}\mathbf{A}) = (\mathbf{A}^\top \otimes \mathbf{I}_k)\mathrm{vec}(\mathbf{\Gamma}),$$

so that

$$\mathbb{V}ar[\mathrm{vec}(\mathbf{\Gamma}\mathbf{A})] = \mathbb{V}ar[(\mathbf{A}^\top \otimes \mathbf{I}_k)\mathrm{vec}(\mathbf{\Gamma})] = (\mathbf{A}^\top \otimes \mathbf{I}_k)\mathbb{V}ar[\mathrm{vec}(\mathbf{\Gamma})](\mathbf{A}^\top \otimes \mathbf{I}_k)^\top.$$

Since $\mathbb{V}ar[\mathrm{vec}(\mathbf{\Gamma})] = \mathbb{V}ar(\mathbf{\Gamma})$ and $(\mathbf{A}^\top \otimes \mathbf{I}_k)^\top = \mathbf{A} \otimes \mathbf{I}_k$, we obtain (3.6). □

It is clear from the form of the ChF of $\mathbf{X}$ in (3.3) that in the symmetric case (where $\mathbf{A} = \mathbf{0}$) the distribution of $\mathbf{X}$ is infinitely divisible (ID) whenever that of the random scaling matrix $\mathbf{\Gamma}$ is. Infinitely divisible Gaussian covariance mixture models in the same spirit, with the random scaling (and not necessarily squared) matrix on the right-hand-side of the $\mathbf{Z}$ in (3.1) rather than on the left side, were studied in [5], where they were termed *MatG random matrices*, in analogy with *Type G* random variables and vectors defined in similar manner (see the references in [5]). On the other hand, upon close examination of the relevant PDF in (3.5), it is also clear that a a convenient choice of $\mathbf{\Gamma}$ in this set-up is one where the PDF of the latter is compatible with the expression under the expectation in (3.5). A rather obvious choice here is a *matrix variate generalized* *inverse Gaussian* (MGIG) distribution, denoted by $\mathcal{MGIG}_k(\lambda, \mathbf{\Psi}, \mathbf{\Theta})$ and given by the PDF

$$f(\mathbf{\Gamma}) = \frac{2^{k\lambda}|\mathbf{\Psi}|^{-\lambda}}{B_\lambda\left(\frac{1}{4}\mathbf{\Theta}\mathbf{\Psi}\right)}|\mathbf{\Gamma}|^{\lambda-(k+1)/2}\mathrm{etr}\left\{-\frac{1}{2}\left(\mathbf{\Psi}\mathbf{\Gamma}^{-1}+\mathbf{\Theta}\mathbf{\Gamma}\right)\right\}, \tag{3.9}$$

where $\mathbf{\Psi}$ and $\mathbf{\Theta}$ are $k \times k$ symmetric non-negative definite matrices, $\lambda \in \mathbb{R}$, and $B_\lambda(\cdot)$ is the type-2 Bessel function of Herz of matrix argument (see [25]). Indeed, the PDF of this matrix variate analog of the univariate GIG distribution (2.16) has exactly the same functional form as the expression under the expectation in (3.5), which aids in evaluation of its expectation. In analogy with its univariate and multivariate analogs, matrix variate distributions that arise in this scheme with MGIG random scaling matrices were termed *matrix variate generalized hyperbolic* distributions in [35], who considered the distribution of $\mathbf{X}^\top$ with MGIG distributed scaling matrix $\mathbf{\Gamma}$ in (3.1) and an additional matrix location parameter. A symmetric version of this distribution (with $\mathbf{A} = \mathbf{0}$) was also studied in [23]. These papers also briefly considered the special cases of MGIG where either $\mathbf{\Psi} = \mathbf{0}$ or $\mathbf{\Theta} = \mathbf{0}$, corresponding to matrix variate gamma and inverse matrix variate gamma distributions, respectively. While the latter special case has received a considerable attention in the literature, as it generates a matrix variate analog of t distribution in this scheme (see, e.g., [2], [6], [12], [20], [23], [26], [36]), the former, leading to a matrix variate version of the classical Laplace distribution, has not been studied much. Below we fill this gap and provide its important distributional properties. In the section below, we first recall the classical matrix variate gamma distribution (see, e.g, [21]), which is needed in this construction.

3.1. **Matrix variate gamma distribution.** A $k \times k$ positive definite random matrix $\mathbf{X}$ has the classical matrix variate gamma (MG) distribution with shape parameter $\alpha > (k-1)/2$ and scale matrix parameter $\mathbf{D} \in \mathbb{S}_k^+$, denoted by $\mathcal{MG}_k(\alpha, \mathbf{D})$, if its PDF is of the form

$$f(\mathbf{X}) = \frac{1}{\Gamma_k(\alpha)|\mathbf{D}|^\alpha}|\mathbf{X}|^{\alpha-(k+1)/2}\mathrm{etr}\{-\mathbf{D}^{-1}\mathbf{X}\}, \quad \mathbf{X} > 0, \tag{3.10}$$

where

$$\Gamma_k(\alpha) = \pi^{k(k-1)/4}\prod_{i=1}^{k}\Gamma\left(\alpha - \frac{i-1}{2}\right)$$

is the generalized gamma function (see [21, Chapter 3.6]). The LT of $\mathbf{X} \sim \mathcal{MG}_k(\alpha, \mathbf{D})$ is of the form

$$\psi_{\mathbf{X}}(\mathbf{T}) = \mathbb{E}\left(\mathrm{etr}\{-\mathbf{T}^\top\mathbf{X}\}\right) = |\mathbf{I}_k + \mathbf{T}\mathbf{D}|^{-\alpha}, \tag{3.11}$$

where $\mathbf{T}$ is a $k \times k$ symmetric matrix such that $\mathbf{I}_k + \mathbf{T}\mathbf{D}$ is of the full rank, while the ChF of $\mathbf{X} \sim \mathcal{MG}_k(\alpha, \mathbf{D})$ is given by

$$\varphi_{\mathbf{X}}(\mathbf{T}) = \mathbb{E}\left(\mathrm{etr}\{\iota\mathbf{T}^\top\mathbf{X}\}\right) = \left|\mathbf{I}_k - \frac{1}{2}\iota(\mathbf{T}+\mathbf{T}^\top)\mathbf{D}\right|^{-\alpha}, \tag{3.12}$$

where $\mathbf{T}$ is an arbitrary $k \times k$ real matrix. In addition, as shown in [28], the mean and the variance of $\mathbf{X} \sim \mathcal{MG}_k(\alpha, \mathbf{D})$ are given by

$$\mathbb{E}(\mathbf{X}) = \alpha\mathbf{D}, \ \ \mathbb{V}ar(\mathbf{X}) = \mathbb{V}ar[\text{vec}(\mathbf{X})] = \frac{\alpha}{2}\left(\mathbf{D}\otimes\mathbf{D}\right)\left(\mathbf{I}_{k^2} + \mathbf{K}_k\right), \tag{3.13}$$

where the $k^2 \times k^2$ commutation matrix $\mathbf{K}_k$ consisting of $k \times k$-blocks $\mathbf{S}_{ij}$ is defined through its entries,

$$S_{ij;rs} = \begin{cases} 1; & i = s, j = r, \\ 0; & \text{otherwise.} \end{cases} \tag{3.14}$$

Remark 3.1. If $\mathbf{D} = \mathbf{I}_k$ we get the *standard matrix variate gamma* case, denoted by $\mathcal{MG}_k(\alpha)$, which is also a special case $\lambda = \alpha$, $\mathbf{\Psi} = \mathbf{0}$, and $\mathbf{\Theta} = 2\mathbf{I}_k$ of MGIG distribution (3.9). In this case, we have $\mathbb{E}(\mathbf{X}) = \alpha\mathbf{I}_k$ and $\mathbb{V}ar(\mathbf{X}) = (\alpha/2)(\mathbf{I}_{k^2} + \mathbf{K}_k)$. If $k = 1$, we obtain standard gamma distribution (2.12).

Remark 3.2. The condition $\alpha > (k-1)/2$ guarantees that the PDF is properly defined and integrates to one over the cone $\mathbb{S}_k^+$. The case where $\alpha \le (k-1)/2$ is complicated as $\mathbf{X}$ follows singular matrix variate gamma distribution or singular Wishart distribution (see, for example, [8, 9, 7, 28]).

3.2. **Type II matrix variate generalized Laplace distributions.** In analogy with Type I matrix variate GAL distributions discussed in Section 2.1, we now define the Type II version of this distribution, which arises as a covariance mixture of Gaussian random matrix in (3.1) were the scale matrix $\mathbf{\Gamma}$ has a standard matrix variate gamma distribution.

Definition 3.1. A random $k \times n$ matrix $\mathbf{X}$ is said to have MAL distribution with shape parameter $\alpha > (k-1)/2$, $k \times n$ matrix skewness parameter $\mathbf{A}$, and $k \times k$ matrix scaling parameter $\mathbf{\Sigma} \in \mathbb{S}_n^+$, denoted by $\mathcal{MAL}_{k,n}(\alpha; \mathbf{A}, \mathbf{\Sigma})$, if $\mathbf{X}$ admits the stochastic representation (3.1) with $\mathbf{\Gamma} \sim \mathcal{MG}_k(\alpha)$ independent of $\mathbf{Z} \sim \mathcal{MN}_{k,n}(\mathbf{0}, \mathbf{I}_k \otimes \mathbf{I}_n)$. The distribution of $\mathbf{X}^\top$ is referred to as the MAL$^\top$ distribution, denoted by $\mathcal{MAL}_{n,k}^\top(\alpha; \mathbf{A}, \mathbf{\Sigma})$.

Basic properties of this distribution can be obtained by standard conditioning arguments via (3.2). The result below provides the relevant ChF, obtained by specializing the general result given in Theorem 3.1 to the MAL case.

Theorem 3.4. *The ChF of* $\mathbf{X} \sim \mathcal{MAL}_{k,n}(\alpha; \mathbf{A}, \mathbf{\Sigma})$ *is given by*

$$\varphi_{\mathbf{X}}(\mathbf{T}) = \left| \mathbf{I}_k - \frac{\iota}{2}(\mathbf{A}\mathbf{T}^\top + \mathbf{T}\mathbf{A}^\top) + \frac{1}{2}\mathbf{T}\mathbf{\Sigma}\mathbf{T}^\top \right|^{-\alpha}, \tag{3.15}$$

where $\mathbf{T}$ *is an arbitrary* $k \times n$ *matrix.*

Proof. When we specialize the result in Theorem 3.1 to the MAL case, where the PDF and the LT of $\mathbf{\Gamma}$ are given by (3.10) and (3.11) with $\mathbf{D} = \mathbf{I}_k$, respectively, after some algebra we arrive at

$$\varphi_{\mathbf{X}}(\mathbf{T}) = \int_{\mathbf{\Gamma}>0} \text{etr}\left\{\iota(\mathbf{A}\mathbf{T}^\top)^\top\mathbf{\Gamma}\right\} \frac{1}{\Gamma_k(\alpha)}|\mathbf{\Gamma}|^{\alpha-(k+1)/2}\text{etr}\left\{-\left(\mathbf{I}_k + \frac{1}{2}\mathbf{T}\mathbf{\Sigma}\mathbf{T}^\top\right)\mathbf{\Gamma}\right\} d\mathbf{\Gamma}. \tag{3.16}$$

Upon setting $\mathbf{D}^{-1} = \mathbf{I}_k + \frac{1}{2}\mathbf{T\Sigma T}^\top$, the above can be written as

$$\varphi_{\mathbf{X}}(\mathbf{T}) = |\mathbf{D}|^\alpha \int_{\mathbf{\Gamma}>0} \mathrm{etr}\left\{\iota(\mathbf{AT}^\top)^\top\mathbf{\Gamma}\right\} \frac{|\mathbf{D}|^{-\alpha}}{\Gamma_k(\alpha)}|\mathbf{\Gamma}|^{\alpha-(k+1)/2}\mathrm{etr}\left\{-\mathbf{D}^{-1}\mathbf{\Gamma}\right\} d\mathbf{\Gamma}. \tag{3.17}$$

The integral above can be recognized as the ChF of MG distribution $\mathcal{MG}_k(\alpha, \mathbf{D})$ evaluated at $\mathbf{AT}^\top$, which, according to (3.12) with $\mathbf{D} = \mathbf{I}_k$, is equal to $|\mathbf{I}_k - \iota(1/2)(\mathbf{AT}^\top + \mathbf{TA}^\top)\mathbf{D}|^{-\alpha}$. Therefore, according to (3.17) where the order of the multiplication is reversed, we can conclude that

$$\varphi_{\mathbf{X}}(\mathbf{T}) = |\mathbf{I}_k - \iota(1/2)(\mathbf{AT}^\top + \mathbf{TA}^\top)\mathbf{D}|^{-\alpha}|\mathbf{D}|^\alpha = |\mathbf{D}^{-1} - \iota(1/2)(\mathbf{AT}^\top + \mathbf{TA}^\top)|^{-\alpha}, \tag{3.18}$$

which produces the result since $\mathbf{D}^{-1} = \mathbf{I}_k + \frac{1}{2}\mathbf{T\Sigma T}^\top$. The theorem is proved. □

Our next result provides the PDF of the MAL distribution, which easily follows from the general result in Theorem 3.2.

Theorem 3.5. *Let $\mathbf{X} \sim \mathcal{MAL}_{k,n}(\alpha; \mathbf{A}, \mathbf{\Sigma})$, where $\alpha > (k-1)/2$ and $\mathbf{\Sigma} \in \mathbb{S}_n^+$. Then the PDF of $\mathbf{X}$ is given by*

$$f(\mathbf{X}) = \frac{\left|\mathbf{X\Sigma}^{-1}\mathbf{X}^\top\right|^{\alpha-n/2} \mathrm{etr}\left\{\mathbf{X\Sigma}^{-1}\mathbf{A}^\top\right\}}{2^{\alpha k}\pi^{kn/2}\Gamma_k(\alpha)|\mathbf{\Sigma}|^{k/2}} \times B_{\alpha-n/2}\left(\frac{1}{4}\left(2\mathbf{I}_k + \mathbf{A\Sigma}^{-1}\mathbf{A}^\top\right)\mathbf{X\Sigma}^{-1}\mathbf{X}^\top\right), \tag{3.19}$$

where $B_{\alpha-n/2}(\cdot)$ stands for the Bessel function of matrix argument of the second kind as in [25].

Proof. Theorem 3.2 specialized to the MAL case, where $\mathbf{\Gamma}$ in (3.1) has the standard $\mathcal{MG}_k(\alpha)$ distribution, immediately leads to

$$\begin{aligned} f(\mathbf{X}) &= \frac{1}{(2\pi)^{kn/2}\Gamma_k(\alpha)|\mathbf{\Sigma}|^{k/2}}\mathrm{etr}\left\{\mathbf{X\Sigma}^{-1}\mathbf{A}^\top\right\} \\ &\quad \times \int_{\mathbf{\Gamma}>0} |\mathbf{\Gamma}|^{\lambda-(k+1)/2}\,\mathrm{etr}\left\{-\frac{1}{2}\left[\left(2\mathbf{I}_k + \mathbf{A\Sigma}^{-1}\mathbf{A}^\top\right)\mathbf{\Gamma} + \mathbf{\Gamma}^{-1}\mathbf{X\Sigma}^{-1}\mathbf{X}^\top\right]\right\} d\mathbf{\Gamma} \end{aligned}$$

with $\lambda = \alpha - n/2$. By denoting $\mathbf{\Psi} = \mathbf{X\Sigma}^{-1}\mathbf{X}^\top$ and $\mathbf{\Theta} = 2\mathbf{I}_k + \mathbf{A\Sigma}^{-1}\mathbf{A}^\top$, we can write the above integral as

$$\int_{\mathbf{\Gamma}>0} |\mathbf{\Gamma}|^{\lambda-(k+1)/2}\,\mathrm{etr}\left\{-\frac{1}{2}\left(\mathbf{\Psi\Gamma}^{-1} + \mathbf{\Theta\Gamma}\right)\right\} d\mathbf{\Gamma},$$

and relate it to the $\mathcal{MGIG}_k(\lambda, \mathbf{\Psi}, \mathbf{\Theta})$ PDF (3.9), leading to

$$\int_{\mathbf{\Gamma}>0} |\mathbf{\Gamma}|^{\lambda-(k+1)/2}\,\mathrm{etr}\left\{-\frac{1}{2}\left(\mathbf{\Psi\Gamma}^{-1} + \mathbf{\Theta\Gamma}\right)\right\} d\mathbf{\Gamma} = 2^{-k\lambda}|\mathbf{\Psi}|^\lambda B_\lambda\left(\frac{1}{4}\mathbf{\Theta\Psi}\right).$$

When we combine the above steps, followed by simple algebra, we obtain the result. □

Remark 3.3. We note that in the symmetric case (where $\mathbf{A}$ is a matrix of zeros) the above formulas for the ChF and the PDF of matrix GAL distribution reduce to those given in [23], who studied this symmetric case (which was termed *matrix variate variance gamma*) along with another special case of matrix GIG scaling, involving the inverse matrix variate gamma distribution and leading to (symmetric) matrix variate t distribution (see eq. (3.8)-(3.9) in [23]). [Due to notational differences, to see the equivalence one should replace the quantities λ, n, p, and $\mathbf{\Theta}$ in these equations by α, k, n, and $2\mathbf{I}_k$, respectively, and for the PDF use the relation $B_\lambda(\mathbf{Z}) = B_{-\lambda}(\mathbf{Z})|\mathbf{Z}|^{-\lambda}$ that appears on p. 1470 in [23].] It was shown in [23] that these two matrix variate distributions provide an example of a matrix variate *dual pair*, in the sense that the ChF of one has the same functional form as the PDF of the other.

Remark 3.4. Let us note that if $k = 1$, so that the matrix variate gamma distribution reduces to a univariate standard gamma variable given by the PDF (2.12), the matrix variate MAL turns into a (row) vector of size n with multivariate GAL distribution studied in [27] and [30], and we recover the ChF and the PDF of the GAL distribution (see eq. (3) and (4) in [30]). In fact, this is where the two versions of the matrix variate generalized Laplace distributions coincide. Note that in the GAL PDF we have the modified Bessel function of the third kind with index λ (of scalar argument), denoted by $K_\lambda(\cdot)$, instead of the Bessel function of matrix argument $B_\lambda(\cdot)$, due to the relation $K_\lambda(z) = (1/2)B_\lambda\left(z^2/4\right)(z/2)^\lambda$. In the symmetric (and elliptically contoured) case with $k = 1$ and $\mathbf{A} = \mathbf{0}$ the PDF of the $1 \times n$ MAL random vector becomes:

$$f(\mathbf{X}) = \frac{2^{1-\alpha}\left(\sqrt{2\mathbf{X}\mathbf{\Sigma}^{-1}\mathbf{X}^\top}\right)^{\alpha-n/2}}{\pi^{n/2}\Gamma(\alpha)|\mathbf{\Sigma}|^{1/2}} \times K_{\alpha-n/2}\left(\sqrt{2\mathbf{X}\mathbf{\Sigma}^{-1}\mathbf{X}^\top}\right). \tag{3.20}$$

If we also have $n = 1$, then we recover the univariate GAL distribution studied in [27], which for $\alpha = 1$ coincides with the classical Laplace distribution.

Remark 3.5. As we have seen above, there is a strong connection between matrix variate MAL distributions and matrix variate gamma distributions through the stochastic representation (3.1), where $\mathbf{X} \sim \mathcal{MAL}_{k,n}(\alpha; \mathbf{A}, \mathbf{\Sigma})$ and $\mathbf{\Gamma} \sim \mathcal{MG}_k(\alpha, \mathbf{I}_k)$. However, these distributions are also connected in a different way, as the matrix variate MAL distribution arises as the distribution of the off diagonal blocks of a random MG matrix. Indeed, let $\mathbf{\Gamma} \sim \mathcal{MG}_k(\alpha, \mathbf{D})$, and consider a partition of $\mathbf{\Gamma}$ and its dispersion matrix $\mathbf{D} \in \mathbb{S}_k^+$ into the blocks

$$\mathbf{\Gamma} = \begin{pmatrix} \mathbf{\Gamma}_{11} & \mathbf{\Gamma}_{12} \\ \mathbf{\Gamma}_{21} & \mathbf{\Gamma}_{22} \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{pmatrix} \tag{3.21}$$

with $\dim(\mathbf{\Gamma}_{11}) = \dim(\mathbf{D}_{11}) = r \times r$, $r = 1, \ldots, k-1$. Then, it is well-known that $\mathbf{\Gamma}_{11} \sim \mathcal{MG}_r(\alpha, \mathbf{D}_{11})$ and, conditionally on $\mathbf{\Gamma}_{11}$, the distribution of $\mathbf{\Gamma}_{12}$ is matrix variate normal, $\mathbf{\Gamma}_{12}|\mathbf{\Gamma}_{11} \sim \mathcal{MN}_{r,k-r}\left(\mathbf{\Gamma}_{11}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}, \mathbf{\Gamma}_{11} \otimes \frac{1}{2}\mathbf{D}_{22\cdot 1}\right)$, where $\mathbf{D}_{22\cdot 1} = \mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$ is the Schur complement of $\mathbf{D}_{11}$ (see, e.g., Proposition 2.2 in [28]). Therefore, the random matrix $\mathbf{\Gamma}_{12}$ has the stochastic representation given by the right-hand-side in (3.1), where $\mathbf{A} = \mathbf{D}_{11}^{-1}\mathbf{D}_{12}$, $\mathbf{\Sigma} = \frac{1}{2}\mathbf{D}_{22\cdot 1}$, $\mathbf{Z} \sim \mathcal{MN}_{r,k-r}\left(\mathbf{0}, \mathbf{I}_r \otimes \mathbf{I}_{k-r}\right)$, and $\mathbf{\Gamma}$ is replaced by $\mathbf{\Gamma}_{11}$. Thus, if $\mathbf{D} = \mathbf{I}_k$, we will have $\mathbf{\Gamma}_{12} \sim \mathcal{MAL}(\alpha; \mathbf{0}, (1/2)\mathbf{I}_{k-r})$.

Remark 3.6. Let us note that an alternative matrix variate analog of multivariate generalized asymmetric Laplace distribution was briefly treated in [34]. That model, termed *matrix gamma-normal distribution*, was also defined via (3.1) with $\mathbf{A} = \mathbf{0}$ and $\mathbf{\Sigma} = \mathbf{I}_n$, but another matrix variate analog of gamma distribution was used for a random scale matrix $\mathbf{\Gamma}$, defined through the ChF to ensure its infinite divisibility (which is lacking in the case of the classical matrix variate distribution used here). However, the lack of convenient expression for the PDF connected with that model is a disadvantage, which may limit its practical utility.

It should be noted that scale mixtures of matrix variate normal distributions can be defined in two ways, by mixing either one of the two covariance matrices $\mathbf{\Sigma}$ and $\mathbf{\Psi}$ of $\mathbf{Z} \sim \mathcal{MN}_{k,n}(\mathbf{0}, \mathbf{\Sigma} \otimes \mathbf{\Psi})$. While the resulting distributions are related through the operation of transposition, both schemes generally lead to different probability distributions. For example, the matrix variate variance gamma model considered in [35], obtained as a special case of matrix GIG scaling of normal random matrices, is connected with mixing the covariance matrix $\mathbf{\Psi}$ of the normal matrix rather than the matrix $\mathbf{\Sigma}$, as done here. Consequently, care is needed when translating the properties across the two models. However, in case of standard distributions with $\mathbf{A} = \mathbf{0}$, these two mixing schemes lead to one and the same distribution, as shown below.

Theorem 3.6. *Let $\mathbf{Z} \sim \mathcal{MN}_{k,n}(\mathbf{0}, \mathbf{I}_k \otimes \mathbf{I}_n)$, $\mathbf{\Gamma}_k \sim \mathcal{MG}_k(\alpha)$, and $\mathbf{\Gamma}_n \sim \mathcal{MG}_n(\alpha)$, where $\alpha > (\max\{k,n\} - 1)/2$, be mutually independent. Then, we have the following equality in distribution:*

$$\mathbf{\Gamma}_k^{1/2}\mathbf{Z} \stackrel{d}{=} \mathbf{Z}\mathbf{\Gamma}_n^{1/2}. \tag{3.22}$$

Proof. The result can be easily established via the ChF of the variables on each side in (3.22). Indeed, since $\mathbf{\Gamma}_k^{1/2}\mathbf{Z} \sim \mathcal{MAL}_{k,n}(\alpha; \mathbf{I}_k, \mathbf{I}_n)$, Theorem 3.4 shows that the ChF of $\mathbf{\Gamma}_k^{1/2}\mathbf{Z}$ is given by

$$\varphi_1(\mathbf{T}) = \left|\mathbf{I}_k + \frac{1}{2}\mathbf{T}\mathbf{T}^\top\right|^{-\alpha}. \tag{3.23}$$

On the other hand, since $(\mathbf{Z}\mathbf{\Gamma}_n^{1/2})^\top \sim \mathcal{MAL}_{n,k}(\alpha; \mathbf{I}_n, \mathbf{I}_k)$, another application of Theorem 3.4 shows that the ChF of $(\mathbf{Z}\mathbf{\Gamma}_n^{1/2})^\top$ is given by

$$\varphi_2(\mathbf{T}) = \left|\mathbf{I}_n + \frac{1}{2}\mathbf{T}\mathbf{T}^\top\right|^{-\alpha}. \tag{3.24}$$

Consequently, the ChF of $\mathbf{Z}\mathbf{\Gamma}_n^{1/2}$, denoted by $\varphi_3(\cdot)$, will be

$$\varphi_3(\mathbf{T}) = \varphi_2(\mathbf{T}^\top) = \left|\mathbf{I}_n + \frac{1}{2}\mathbf{T}^\top\mathbf{T}\right|^{-\alpha}. \tag{3.25}$$

Since the ChFs in (3.23) and (3.25) coincide by Sylvester's determinant theorem (see, for example, Corollary 18.1.2 in [24]), we obtain the result. □

Remark 3.7. The above result is related to the fact that the probability distributions of the variables on each side in (3.22) are spherical, which means invariance with respect to linear orthogonal transformations. Indeed, if in the notation of the above theorem we have $\mathbf{X} \stackrel{d}{=} \mathbf{Z}\mathbf{\Gamma}_n^{1/2}$, then conditionally on $\mathbf{\Gamma}_n$ we have that $\mathbf{X} \sim \mathcal{MN}_{k,n}(\mathbf{0}, \mathbf{I}_k \otimes \mathbf{\Gamma}_n)$. Since for any deterministic orthogonal matrix $\mathbf{P}$ we have $\mathbf{P}\mathbf{X} \sim \mathcal{MN}_{k,n}(\mathbf{0}, \mathbf{P}\mathbf{I}_k\mathbf{P}^\top \otimes \mathbf{\Gamma}_n)$ and $\mathbf{P}\mathbf{I}_k\mathbf{P}^\top = \mathbf{I}_k$, the conditional distributions of $\mathbf{P}\mathbf{X}$ and $\mathbf{X}$ coincide. Thus, their unconditional distributions must be the same as well. This shows that $\mathbf{X}$ is left-spherical. Similar argument shows that $\mathbf{X}$ is also right-spherical. This is because given $\mathbf{\Gamma}_n$, $\mathbf{X}\mathbf{P}$ is $\mathcal{MN}_{k,n}(\mathbf{0}, \mathbf{I}_k \otimes \mathbf{P}^\top\mathbf{\Gamma}_n\mathbf{P})$. However, by orthogonality of $\mathbf{P}$, the distribution of $\mathbf{P}^\top\mathbf{\Gamma}_n\mathbf{P}$ is the same as that of $\mathbf{\Gamma}$, which can be established by comparing their ChFs (a matrix with this property is said to be *rotatable* in [11]). This shows that the unconditional distributions of $\mathbf{XP}$ and $\mathbf{X}$ coincide. Since $\mathbf{X}$ is both, left and right spherical, its distribution is spherical. Thus, if $k = n$, we would have that $\mathbf{X} \stackrel{d}{=} \mathbf{X}^\top$ (see Theorem 1 in [11]), leading to the result in Theorem 3.6 for square matrices. An extension to the general case follows by standard arguments (see [11] for more details).

Below we consider several further properties of MAL distributions. The following result concerning the mean and the variance of MAL random matrix is a simple consequence of Theorem 3.3 and the relevant properties of matrix variate gamma distribution, see also equation (3.13), the definition of the commutation matrix $\mathbf{K}_k$ given in (3.14), and the remark following it.

Corollary 3.1. If $\mathbf{X} \sim \mathcal{MAL}_{k,n}(\alpha; \mathbf{A}, \mathbf{\Sigma})$ then $\mathbb{E}(\mathbf{X}) = \alpha\mathbf{A}$ and

$$\mathbb{V}ar(\mathbf{X}) = \frac{\alpha}{2}\left\{\mathbf{A}^\top\mathbf{A} \otimes \mathbf{I}_k + \mathbf{K}_{n,k}(\mathbf{A} \otimes \mathbf{A}^\top)\right\} + \alpha\mathbf{\Sigma} \otimes \mathbf{I}_k, \tag{3.26}$$

where the $nk \times nk$ commutation matrix $\mathbf{K}_{n,k}$ is a $n \times k$ matrix of $k \times n$-blocks $\mathbf{S}_{ij}$, $i = 1, \ldots, n$, $j = 1, \ldots, k$ defined through their entries as in (3.14).

Proof. First, notice that $(\mathbf{A}^\top \otimes \mathbf{I}_k)(\mathbf{A} \otimes \mathbf{I}_k) = \mathbf{A}^\top\mathbf{A} \otimes \mathbf{I}_k$, which holds because of the general relation

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}. \tag{3.27}$$

This, combined with Theorem 3.3 and the expressions for the mean and the variance of $\mathbf{\Gamma} \sim \mathcal{MG}_k(\alpha)$ given in (3.13), produce

$$\mathbb{V}ar(\mathbf{X}) = \frac{\alpha}{2}\left\{\mathbf{A}^\top\mathbf{A} \otimes \mathbf{I}_k + (\mathbf{A}^\top \otimes \mathbf{I}_k)\mathbf{K}_k(\mathbf{A} \otimes \mathbf{I}_k)\right\} + \alpha\mathbf{\Sigma} \otimes \mathbf{I}_k.$$

Next, apply the relation

$$(\mathbf{C} \otimes \mathbf{D})\mathbf{K}_{r,k} = \mathbf{K}_{s,n}(\mathbf{D} \otimes \mathbf{C})$$

which is valid for any $s \times r$ matrix $\mathbf{C}$ and $n \times k$ matrix $\mathbf{D}$ (see Theorem 16.3.2 on p. 337 in [24]), along with (3.27), to obtain

$$(\mathbf{A}^\top \otimes \mathbf{I}_k)\mathbf{K}_k(\mathbf{A} \otimes \mathbf{I}_k) = \mathbf{K}_{n,k}(\mathbf{I}_k \otimes \mathbf{A}^\top)(\mathbf{A} \otimes \mathbf{I}_k) = \mathbf{K}_{n,k}(\mathbf{A} \otimes \mathbf{A}^\top).$$

□

The result below shows that the class of $\mathcal{MAL}_{k,n}(\alpha; \mathbf{A}, \mathbf{\Sigma})$ distributions with fixed $\mathbf{A}$ and $\mathbf{\Sigma}$ is closed under convolutions.

Corollary 3.2. Let $\mathbf{X}_i \sim \mathcal{MAL}_{k,n}(\alpha_i; \mathbf{A}, \mathbf{\Sigma})$, where $i = \{1, \ldots, m\}$, be mutually independent, with $\min\{\alpha_1, \ldots, \alpha_m\} > (k-1)/2$. Then,

$$\mathbf{X}_1 + \cdots + \mathbf{X}_m \sim \mathcal{MAL}_{k,n}(\alpha; \mathbf{A}, \mathbf{\Sigma}),$$

where $\alpha = \alpha_1 + \cdots + \alpha_m$.

Proof. This is a simple consequence of the formula for the MAL ChF given in Theorem 3.4. □

Our next result shows yet another distributional connection between the matrix variate generalized Laplace and gamma distributions. The result below concerns the distribution of a *symmetrized* version of $\mathbf{X} \sim \mathcal{MAL}_{k,k}(\alpha; \mathbf{0}, \mathbf{I}_k)$, $\mathbf{Y} \stackrel{d}{=} \mathbf{X} + \mathbf{X}^\top$, so that the ChFs of $\mathbf{X}$ and $\mathbf{Y}$ are related as

$$\varphi_{\mathbf{Y}}(\mathbf{T}) = \mathbb{E}(\mathrm{etr}\{\iota \mathbf{T}^\top \mathbf{Y}\}) = \mathbb{E}(\mathrm{etr}\{\iota [\mathbf{T}^\top + \mathbf{T}]\mathbf{X}\}) = \varphi_{\mathbf{X}}(\mathbf{T}^\top + \mathbf{T}), \tag{3.28}$$

where $\mathbf{T}$ is an arbitrary square matrix.

Proposition 1. *If $\mathbf{X} \sim \mathcal{MAL}_{k,k}(\alpha; \mathbf{0}, \mathbf{I}_k)$ then we have the stochastic representation*

$$\mathbf{X} + \mathbf{X}^\top \stackrel{d}{=} \mathbf{\Gamma}_1 - \mathbf{\Gamma}_2, \tag{3.29}$$

where the $\{\mathbf{\Gamma}_i\}$ are IID with $\mathcal{MG}_k(\alpha, \sqrt{2}\mathbf{I}_k)$ distributions.

Proof. By Theorem 3.4 and (3.28), the ChF of $\mathbf{Y} = \mathbf{X} + \mathbf{X}^\top$ evaluated at an arbitrary $k \times k$ matrix $\mathbf{T}$ is given by

$$\varphi_{\mathbf{Y}}(\mathbf{T}) = \left| \mathbf{I}_k + \frac{1}{2}(\mathbf{T} + \mathbf{T}^\top)(\mathbf{T} + \mathbf{T}^\top) \right|^{-\alpha}. \tag{3.30}$$

At the same time, in view of (3.12), the ChF of $\mathbf{\Gamma}_1 - \mathbf{\Gamma}_2$ evaluated at the same $\mathbf{T}$ is given by

$$\varphi_{\mathbf{\Gamma}_1 - \mathbf{\Gamma}_2}(\mathbf{T}) = \left| \mathbf{I}_k - \frac{1}{2}\iota(\mathbf{T} + \mathbf{T}^\top)\sqrt{2}\mathbf{I}_k \right|^{-\alpha} \left| \mathbf{I}_k + \frac{1}{2}\iota(\mathbf{T} + \mathbf{T}^\top)\sqrt{2}\mathbf{I}_k \right|^{-\alpha}, \tag{3.31}$$

which, after routine algebra, turns into the right-hand-side of (3.30). This proves the result. □

Remark 3.8. When we take $k = 1$ in Proposition 1, so that all the matrix variate distributions turn into univariate distributions, the distributional equality in (3.29) turns into

$$X \stackrel{d}{=} \frac{1}{\sqrt{2}}(G_1 - G_2), \tag{3.32}$$

where X has a symmetric GAL distribution with the ChF $\varphi_X(t) = (1 - t^2/2)^{-\alpha}$ and the $\{G_i\}$ are IID standard gamma variables with the PDF (2.12). While in the one dimensional case an analog of (3.32) holds for asymmetric GAL distributions as well (sse Proposition 4.1.3 in [27]), no such extension appears to be true in the matrix variate asymmetric generalized Laplace case. It is also worth mentioning that the property in (3.32) is a generalization of the classical result for the Laplace random variable X (corresponding to $\alpha = 1$ in this setting), recently discussed in [14] (see also [29]).

Our next result shows that all linear combinations of the components of MAL distribution are jointly MAL distributed as well (under scaling the MAL random matrix on the right-hand side).

Corollary 3.3. Let $\mathbf{X} \sim \mathcal{MAL}_{k,n}(\alpha; \mathbf{A}, \mathbf{\Sigma})$ with $\alpha > (k-1)/2$, and let $\mathbf{L}$ be a $n \times q$ non-singular matrix of constants such that $rank(\mathbf{L}) = q \leq n$. Then, we have

$$\mathbf{XL} \sim \mathcal{MAL}_{k,q}(\alpha; \mathbf{AL}, \mathbf{\Sigma_L}),$$

where $\mathbf{\Sigma_L} = \mathbf{L}^\top \mathbf{\Sigma} \mathbf{L}$.

Proof. By the assumption and in view of the stochastic representation (3.1) of $\mathbf{X}$, we have

$$\mathbf{XL} \stackrel{d}{=} \mathbf{\Gamma AL} + \mathbf{\Gamma}^{1/2}\mathbf{Z\Sigma}^{1/2}\mathbf{L}. \tag{3.33}$$

This shows that, conditionally on $\mathbf{\Gamma}$, we have $\mathbf{\Gamma}^{1/2}\mathbf{Z\Sigma}^{1/2}|\mathbf{\Gamma} \sim \mathcal{MN}_{k,n}(\mathbf{0}, \mathbf{\Gamma} \otimes \mathbf{\Sigma})$, so that by well-known properties of matrix variate normal distributions (see, e.g., Theorem 2.3.10 in [21]), we also have $\mathbf{\Gamma}^{1/2}\mathbf{Z\Sigma}^{1/2}\mathbf{L}|\mathbf{\Gamma} \sim \mathcal{MN}_{k,q}\left(\mathbf{0}, \mathbf{\Gamma} \otimes \mathbf{L}^\top\mathbf{\Sigma L}\right)$. Since this leads to the conclusion that $\mathbf{XL}|\mathbf{\Gamma} \sim \mathcal{MN}_{k,q}(\mathbf{\Gamma AL}, \mathbf{\Gamma} \otimes \mathbf{\Sigma_L})$, the result follows. □

In the special case where the $\mathbf{L}$ in Corollary 3.3 is of the form $\mathbf{L}^\top = [\mathbf{I}_q\ \mathbf{0}]$, where $\mathbf{0}$ is a $q \times n - q$ matrix of zeros, we obtain the distribution of the $k \times q$ random matrix $\mathbf{X}_1$ that appears in the partition $\mathbf{X} = [\mathbf{X}_1\ \mathbf{X}_2]$ of $\mathbf{X} \sim \mathcal{MAL}_{k,n}(\alpha; \mathbf{A}, \mathbf{\Sigma})$.

Corollary 3.4. Let $\mathbf{X} \sim \mathcal{MAL}_{k,n}(\alpha; \mathbf{A}, \mathbf{\Sigma})$ with $\alpha > (k-1)/2$, and let $\mathbf{X} = [\mathbf{X}_1\ \mathbf{X}_2]$ and $\mathbf{A} = [\mathbf{A}_1\ \mathbf{A}_2]$, where $\mathbf{X}_1$ and $\mathbf{A}_1$ are $k \times q$ sub-matrices of $\mathbf{X}$ and $\mathbf{A}$, respectively. Then

$$\mathbf{X}_1 \quad \sim \quad \mathcal{MAL}_{k,q}(\alpha; \mathbf{A}_1, \mathbf{\Sigma}_{11}),$$

where the matrix $\mathbf{\Sigma}_{11}$ is defined in (3.34).

Proof. In the context of Corollary 3.3, take $\mathbf{L}$ to be an $n \times q$ matrix such that $\mathbf{L}^\top = [\mathbf{I}_q\ \mathbf{0}]$, where $\mathbf{0}$ is a $q \times n - q$ matrix of zeros, and note that $\mathbf{XL} = \mathbf{X}_1$, $\mathbf{AL} = \mathbf{A}_1$, and $\mathbf{\Sigma_L} = \mathbf{\Sigma}_{11}$. □

Remark 3.9. The results in Corollaries 3.3 and 3.4 where discussed in Sections 3.3 and 3.4 in [35] in the context of matrix variate generalized hyperbolic distributions. Note that their formulations is somewhat different due to different definitions of scaling by the covariance matrix (in our definition the scaling random matrix is on the left-hand-side of $\mathbf{Z}$, while in [35] it was placed on the right-hand-side of $\mathbf{Z}$).

Finally, we consider the distributions of its diagonal blocks in case where a MAL random matrix $\mathbf{X}$ is a $k \times k$ square matrix and both, $\mathbf{X}$ and $\mathbf{\Sigma}$, are partitioned as

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \end{bmatrix} \quad \text{and} \quad \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}, \tag{3.34}$$

where $\mathbf{X}_{11}$ and $\mathbf{\Sigma}_{11}$ are $q \times q$ matrices with $1 \le q \le k-1$. The result below provides the details.

Corollary 3.5. Let $\mathbf{X} \sim \mathcal{MAL}_{k,k}(\alpha; \mathbf{0}, \mathbf{\Sigma})$ where $k \ge 2$. Then

$$\begin{aligned} \mathbf{X}_{11} &\sim \mathcal{MAL}_{q,q}(\alpha; \mathbf{0}, \mathbf{\Sigma}_{11}), \\ \mathbf{X}_{22} &\sim \mathcal{MAL}_{k-q,k-q}(\alpha; \mathbf{0}, \mathbf{\Sigma}_{22}), \end{aligned}$$

where $\mathbf{X}_{11}$, $\mathbf{X}_{22}$, $\mathbf{\Sigma}_{11}$ and $\mathbf{\Sigma}_{22}$ are defined in (3.34).

Proof. The result can be proven by showing that the ChFs of $\mathbf{X}_{11}$ and $\mathbf{X}_{22}$ have the structure given in Theorem 3.4. Indeed, observe that the ChF of $\mathbf{X}_{11}$ evaluated at an arbitrary $q \times q$ square matrix $\mathbf{T}_{11}$ is the same as the ChF of $\mathbf{X}$ evaluated at

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{3.35}$$

Upon substitution of this $\mathbf{T}$ into the ChF (3.15) of $\mathbf{X} \sim \mathcal{MAL}_{k,k}(\alpha; \mathbf{0}, \mathbf{\Sigma})$, followed by straightforward algebra, we arrive at

$$\varphi_{\mathbf{X}}(\mathbf{T}) = \left| \mathbf{I}_k + \frac{1}{2} \begin{bmatrix} \mathbf{T}_{11}\mathbf{\Sigma}_{11}\mathbf{T}_{11}^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right|^{-\alpha}.$$

However, we have

$$\left| \mathbf{I}_k + \frac{1}{2} \begin{bmatrix} \mathbf{T}_{11}\mathbf{\Sigma}_{11}\mathbf{T}_{11}^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right| = \left| \begin{bmatrix} \mathbf{I}_q + \frac{1}{2}\mathbf{T}_{11}\mathbf{\Sigma}_{11}\mathbf{T}_{11}^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{k-q} \end{bmatrix} \right| = \left| \mathbf{I}_q + \frac{1}{2}\mathbf{T}_{11}\mathbf{\Sigma}_{11}\mathbf{T}_{11}^\top \right|.$$

Thus, the ChF of $\mathbf{X}_{11}$ is of the form

$$\varphi_{\mathbf{X}_{11}}(\mathbf{T}_{11}) = \left| \mathbf{I}_q + \frac{1}{2}\mathbf{T}_{11}\mathbf{\Sigma}_{11}\mathbf{T}_{11}^\top \right|^{-\alpha},$$

so that, in view of Theorem 3.4, we conclude that $\mathbf{X}_{11} \sim \mathcal{MAL}_{q,q}(\alpha; \mathbf{0}, \mathbf{\Sigma}_{11})$, as desired. The proof of the result for $\mathbf{X}_{22}$ is similar. □

4. Vector-valued MAL distributions

Below we provide a few further properties and examples related to vector-valued MGAL and MAL distributions, which shed light on the inter-relations between these two classes of multivariate distributions. We consider the case where these correspond to column vectors of size $k \times 1$.

4.1. Type I.

In case of Type I, the ChF of $\mathbf{X} \sim \mathcal{MGAL}_{k,1}(\mathbf{A}, \mathbf{\Sigma}, 1, \alpha)$, where $\mathbf{A}$ is a $k \times 1$ deterministic vector and $\mathbf{\Psi}$ is taken to be a 1×1 matrix (taken to be 1), is of the form

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \left(1 - \imath\mathbf{t}^\top\mathbf{A} + \frac{1}{2}\mathbf{t}^\top\mathbf{\Sigma}\mathbf{t}\right)^{-\alpha}, \quad \mathbf{t} \in \mathbb{R}^k. \tag{4.1}$$

This distribution coincides with the multivariate generalized asymmetric Laplace (GAL) distribution, denoted by $\mathcal{GAL}_k(\mathbf{A}, \mathbf{\Sigma}, \alpha)$ in the notation of [27]. The mean vector and the covariance matrix of such $\mathbf{X} \sim \mathcal{GAL}_k(\mathbf{A}, \mathbf{\Sigma}, \alpha)$ are given by

$$\mathbb{E}(\mathbf{X}) = \alpha\mathbf{A}, \quad \mathbb{V}ar(\mathbf{X}) = \alpha\left(\mathbf{A}\mathbf{A}^\top + \mathbf{\Sigma}\right). \tag{4.2}$$

In particular, the components of $\mathbf{X}$ are uncorrelated (but not independent) if $\mathbf{A} = \mathbf{0}$ and $\mathbf{\Sigma} = \sigma^2\mathbf{I}_k$ for some $\sigma^2 > 0$. Basic properties of this distribution can be found in [27].

4.2. **Type II.** For Type II MAL random matrix of size $k \times 1$ the scaling matrix $\mathbf{\Sigma}$ reduces to a scalar, which we shall denote by σ^2. Such a random vector $\mathbf{X} \sim \mathcal{MAL}_{k,1}(\alpha; \mathbf{A}, \sigma^2)$ has the stochastic representation

$$\mathbf{X} \stackrel{d}{=} \mathbf{\Gamma A} + \sigma \mathbf{\Gamma}^{1/2} \mathbf{Z}, \tag{4.3}$$

where $\mathbf{A}$ is as before, $\mathbf{\Gamma} \sim \mathcal{MG}_k(\alpha)$ with $\alpha > (k-1)/2$, and $\mathbf{Z}$ is a $k \times 1$ column vector with IID standard normal components, independent of $\mathbf{\Gamma}$. The ChF of this $\mathbf{X}$ is of the form

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \left| \mathbf{I}_k - \frac{\iota}{2}(\mathbf{t}\mathbf{A}^\top + \mathbf{A}\mathbf{t}^\top) + \frac{\sigma^2}{2}\mathbf{t}\mathbf{t}^\top \right|^{-\alpha}, \ \ \mathbf{t} \in \mathbb{R}^k. \tag{4.4}$$

Remark 4.1. We note a curious similarity between the above ChF and the one for the Type I random vector in the special case where the scale matrix $\mathbf{\Sigma}$ in (4.1) is an identity matrix multiplied by σ^2, so that (4.1) takes on the form

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \left| \mathbf{I}_1 - \frac{\iota}{2}(\mathbf{t}^\top\mathbf{A} + \mathbf{A}^\top\mathbf{t}) + \frac{\sigma^2}{2}\mathbf{t}^\top\mathbf{t} \right|^{-\alpha}, \ \ \mathbf{t} \in \mathbb{R}^k. \tag{4.5}$$

[Note that we used a determinant in (4.5), as the expression inside is a scalar.] The only difference between (4.4) and (4.5) is the positioning of the transposition operator.

We now consider the mean vector and the covariance matrix of $\mathbf{X} \sim \mathcal{MAL}_{k,1}(\alpha; \mathbf{A}, \sigma^2)$. By Corollary 3.1, we have $\mathbb{E}(\mathbf{X}) = \alpha\mathbf{A}$, which coincides with the mean of Type I matrix variate Laplace considered above. The same result shows that the variance of $\mathbf{X}$ is given by (3.26) where now $\mathbf{A}^\top\mathbf{A}$ is a scalar and $\mathbf{\Sigma} = \sigma^2$, so that the variance of $\mathbf{X}$ in (3.26) becomes

$$\mathbb{V}ar(\mathbf{X}) = \frac{\alpha}{2}\left\{\mathbf{A}^\top\mathbf{A}\mathbf{I}_k + \mathbf{K}_{1,k}(\mathbf{A} \otimes \mathbf{A}^\top)\right\} + \alpha\sigma^2\mathbf{I}_k. \tag{4.6}$$

By standard properties of the Kronecker product and commuting operator (see, e.g., Corollary 16.3.3 in [24]), we have $\mathbf{K}_{1,k}(\mathbf{A} \otimes \mathbf{A}^\top) = \mathbf{A}^\top\mathbf{A}$, so that

$$\mathbb{V}ar(\mathbf{X}) = \frac{\alpha}{2}\left\{\mathbf{A}^\top\mathbf{A}\mathbf{I}_k + \mathbf{A}\mathbf{A}^\top\right\} + \alpha\sigma^2\mathbf{I}_k. \tag{4.7}$$

4.3. **Further notes and examples.** Based on the discussion above, it appears that the two types of MAL distributions have different covariance structures. We illustrate the effect on the distributions through an example involving three random vectors of size $k = 2$, $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$, corresponding to Type I $\mathcal{M}GAL_{2,1}(\mathbf{A}, \sigma^2\mathbf{I}_2, 1, \alpha)$, Type II $\mathcal{M}AL_{2,1}(\alpha; \mathbf{A}, \sigma^2)$, and Type I $\mathcal{M}GAL_{2,1}(\mathbf{A}, \mathbf{\Sigma}, 1, \alpha)$, respectively. Here, the scale matrix $\mathbf{\Sigma}$ of $\mathbf{X}_3$ is chosen so that the mean and the covariance of the last two distributions are matched,

$$\mathbf{\Sigma} = \frac{1}{2}\mathbf{A}^\perp\left(\mathbf{A}^\perp\right)^\top + \sigma^2\mathbf{I}_2 = \frac{1}{2}\begin{pmatrix} a_2^2 + 2\sigma^2 & -a_1a_2 \\ -a_1a_2 & a_1^2 + 2\sigma^2 \end{pmatrix}, \tag{4.8}$$

where $\mathbf{A} = (a_1 \ a_2)^\top$ and $\mathbf{A}^\perp = (-a_2 \ a_1)^\top$. Note that this is a genuine covariance matrix, as for any $\mathbf{x} = (x_1 \ x_2)^\top$ we have

$$\mathbf{x}^\top\mathbf{\Sigma}\mathbf{x} = \frac{1}{2}(a_1x_2 - a_2x_1)^2 + \sigma^2(x_1^2 + x_2^2) \geq 0.$$

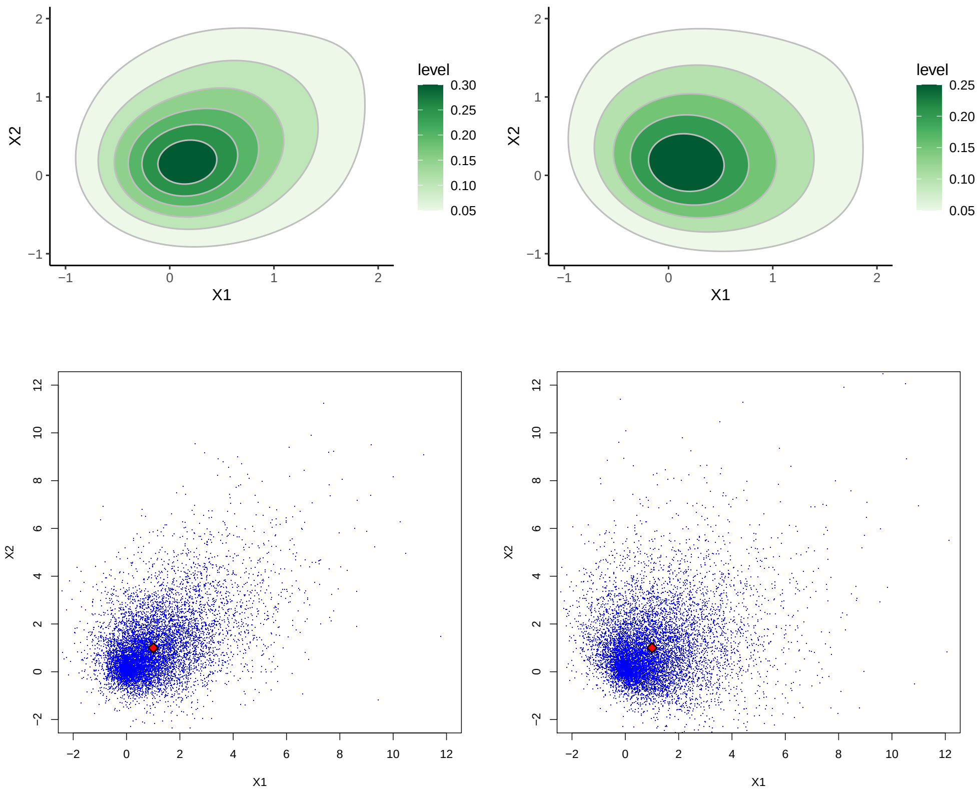

FIGURE 1. Contour plots of bivariate MGAL and MAL distributions of $\mathbf{X}_1$ and $\mathbf{X}_2$ in Example 4.1 (top) estimated from 10000 simulated pairs (bottom). *Left Panels*: the MGAL distribution, constructed via standard univariate gamma scaling with $\alpha = 1$, $\mathbf{A} = (1\ 1)^\top$, and $\mathbf{\Sigma} = \mathbf{I}_2$. *Right Panels*: the MAL distribution, constructed via matrix variate gamma scaling with $\alpha = 1$, $\mathbf{A} = (1\ 1)^\top$, and $\sigma^2 = 1$. The diamond in the bottom figures marks the mean vector. Its location relative to the mode illustrates skewness of the distributions.

Example 4.1. We consider a special case where $\alpha = \sigma = 1$ and $\mathbf{A} = (1\ 1)^\top$, which yields

$$\mathbf{\Sigma} = \frac{1}{2}\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}. \tag{4.9}$$

The three distributions have the same mean vector of $(1\ 1)^\top$ and the following covariance matrices:

$$\mathbb{V}ar(\mathbf{X}_1) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},\ \mathbb{V}ar(\mathbf{X}_2) = \mathbb{V}ar(\mathbf{X}_3) = \begin{pmatrix} 2.5 & 0.5 \\ 0.5 & 2.5 \end{pmatrix}.$$

The sample estimate of the densities from these variables based on 10000 simulated values and the simulated values are shown in Figure 1. The figure only shows the distributions of $\mathbf{X}_1$ and $\mathbf{X}_2$, as the distributions of $\mathbf{X}_2$ and $\mathbf{X}_3$ are actually the same. Indeed, straightforward algebra, based on (4.1) and (4.4), shows that

$$\varphi_{\mathbf{X}_2}(t_1, t_2) = \varphi_{\mathbf{X}_3}(t_1, t_2) = \left(1 - \iota(t_1 + t_2) - \frac{t_1 t_2}{2} + 3\frac{t_1^2 + t_2^2}{4}\right)^{-\alpha}.$$

This observation leads us to further questions concerning the relation between the two classes of distributions, discussed in the next section.

5. Are MAL distributions also MGAL?

The MAL distributions appear to be more complex than the MGAL since they are obtained by mean-variance *matrix-gamma* mixtures of the normal matrices, while the MGAL are 'only' *scalar-gamma* mixtures. However, Example 4.1 suggests that this intuition may be wrong, as it shows that at least some of the MAL distributions are also MGAL. We demonstrate that all $k \times 1$ MAL distributions are actually MGAL.

First, we observe that in the symmetric case where $\mathbf{A} = \mathbf{0}$, for any dimension $k \in \mathbb{N}$ we have $\mathcal{MAL}_{k,1}(\alpha; \mathbf{A}, \sigma^2) = \mathcal{MGAL}_{k,1}(\mathbf{A}, \sigma^2\mathbf{I}_k, 1, \alpha)$.

This is because, by Corollary 18.1.3 in [24], the corresponding ChFs coincide

$$\left|\mathbf{I}_1 + \frac{\sigma^2}{2}\mathbf{t}^\top\mathbf{t}\right|^{-\alpha} = \left|\mathbf{I}_k + \frac{\sigma^2}{2}\mathbf{t}\mathbf{t}^\top\right|^{-\alpha}, \quad \mathbf{t} \in \mathbb{R}^k. \tag{5.1}$$

It is quite remarkable that either matrix or scalar random scaling of a random vector with IID standard normal components both lead to one and the same distribution. More precisely, we have the following stochastic identity

$$W^{1/2}\mathbf{Z} \stackrel{d}{=} \mathbf{\Gamma}^{1/2}\mathbf{Z}, \tag{5.2}$$

where W has standard (univariate) gamma distribution with shape parameter α, $\mathbf{\Gamma} \sim \mathcal{MG}_k(\alpha)$ with $\alpha > (k-1)/2$, and $\mathbf{Z}$ is a $k \times 1$ column vector with IID standard normal components, independent of $\mathbf{\Gamma}$ and W. It follows from this relation that the MAL random vector on the right-hand-side in (5.2) is infinitely divisible (since the left-hand-side is), even though the distribution of the random gamma matrix $\mathbf{\Gamma}$ is not infinitely divisible.

Remark 5.1. A question arises whether the relation in (5.2) may be true if $\mathbf{\Gamma}^{1/2}$ is replaced by another $k \times k$ random matrix such $\mathbf{Y}$ that $\mathbf{Y}\mathbf{Y}^\top = \mathbf{\Gamma}$. In one particular case where $\alpha = k/2$ the answer is affirmative. Indeed, consider $\mathbf{Y} = \mathbf{X}/\sqrt{2}$, where $\mathbf{X}$ is a $k \times k$ random matrix with IID standard normal components. Then, the product $\mathbf{X}\mathbf{X}^\top$ has Wishart distribution with $k/2$ degrees of freedom, which is the same as $\mathcal{MG}_k(k/2, 2\mathbf{I}_k)$ distribution, showing that $\mathbf{Y}\mathbf{Y}^\top \sim \mathcal{MG}_k(\alpha)$. Then, the relation $W^{1/2}\mathbf{Z} \stackrel{d}{=} \mathbf{Y}\mathbf{Z}$, or, equivalently, $(2W)^{1/2}\mathbf{Z} \stackrel{d}{=} \mathbf{X}\mathbf{Z}$ holds by Corollary 1 in [32], since $(2W)^{1/2}\mathbf{Z} \sim \mathcal{GAL}_k(\mathbf{0}, 2\mathbf{I}_k, \alpha)$.

It is even more remarkable that for any, not necessarily symmetric MAL random vector, we can find a corresponding MGAL random vector with the same mean and the covariance matrix, and the two actually have the same probability distribution! In other words, in the vector-valued case ($n = 1$) the class of Type II matrix variate generalized Laplace distributions is a subset of the class of Type I matrix variate generalized Laplace distributions, as shown in the result below.

Before we state the result, let us note that in order for $\mathbf{X}_1 \sim \mathcal{MAL}_{k,1}(\alpha; \mathbf{A}_1, \sigma^2)$ and $\mathbf{X}_2 \sim \mathcal{MGAL}_{k,1}(\mathbf{A}_2, \mathbf{\Sigma}, 1, \alpha)$ to have the same means and covariance matrices, we must have $\mathbf{A}_1 = \mathbf{A}_2$ and

$$\mathbf{\Sigma} = \frac{1}{2}\left\{\mathbf{A}^\top\mathbf{A}\mathbf{I}_k - \mathbf{A}\mathbf{A}^\top\right\} + \sigma^2\mathbf{I}_k, \tag{5.3}$$

where $\mathbf{A} = \mathbf{A}_1 = \mathbf{A}_2$. In the special case $k = 2$, the above matrix reduces to that given by (4.8), and just like the latter it is always non-negative definite (and positive definite whenever $\sigma^2 > 0$), so that it is a valid covariance matrix. To see this, note that, by the matrix determinant lemma, the characteristic polynomial of the matrix $\mathbf{C} = \mathbf{A}^\top\mathbf{A}\mathbf{I}_k - \mathbf{A}\mathbf{A}^\top$ is given by

$$|\lambda\mathbf{I}_k - \mathbf{A}^\top\mathbf{A}\mathbf{I}_k + \mathbf{A}\mathbf{A}^\top| = \left(\lambda - \mathbf{A}^\top\mathbf{A}\right)^k\left(1 + \frac{\mathbf{A}^\top\mathbf{A}}{\lambda - \mathbf{A}^\top\mathbf{A}}\right) = (\lambda - \mathbf{A}^\top\mathbf{A})^{k-1}\lambda,$$

so that $\mathbf{C}$ is singular with one-dimensional kernel and the remaining eigenvalues are non-negative (positive unless $\mathbf{A} = \mathbf{0}$), and all equal to $\mathbf{A}^\top\mathbf{A}$.

Theorem 5.1. *Let $\mathbf{X} \sim \mathcal{MAL}_{k,1}(\alpha; \mathbf{A}, \sigma^2)$ with $\mathbf{A} = (a_1 \dots a_k)^\top$, $\alpha > (k-1)/2$, $\sigma^2 \geq 0$, and the ChF given by (4.4). Then, we also have $\mathbf{X} \sim \mathcal{MGAL}_{k,1}(\mathbf{A}, \mathbf{\Sigma}, 1, \alpha)$ with $\mathbf{\Sigma}$ given by (5.3).*

Observe in view of the forms of the ChFs of the Type I and Type II distributions, the above theorem is equivalent to the algebraic identity stated in the lemma below, which is of independent interest. Indeed, by taking $\mathbf{a} = -\mathbf{A}\iota/2$, $\mathbf{b} = \mathbf{t}$, and $s = \sigma^2$ in the result below, the identity in (5.4) is equivalent to the equality of the ChFs of $\mathbf{X} \sim \mathcal{MAL}_{k,1}(\alpha; \mathbf{A}, \sigma^2)$ and $\mathbf{X} \sim \mathcal{MGAL}_{k,1}(\mathbf{A}, \mathbf{\Sigma}, 1, \alpha)$ evaluated at $\mathbf{t} \in \mathbb{R}^{\mathbf{k}}$.

Lemma 5.1. *For each $\mathbf{a}, \mathbf{b} \in \mathbb{C}^k, s \in \mathbb{C}$, we have*

$$|\mathbf{I}_k + \mathbf{a}\mathbf{b}^\top + \mathbf{b}\mathbf{a}^\top + s\mathbf{b}\mathbf{b}^\top| = \left(1 + \mathbf{b}^\top\mathbf{a}\right)^2 + \mathbf{b}^\top\mathbf{b}\left(s - \mathbf{a}^\top\mathbf{a}\right). \tag{5.4}$$

Proof. Since the functions of $(\mathbf{a}, \mathbf{b})$ on both sides of (5.4) are continuous it is enough to show the identity for $1 + \mathbf{a}^\top\mathbf{b} \neq 0$.

An application of Lemma 1.1 of [13] and Corollary 18.1.3 of [24] produce

$$\begin{aligned} \left|\mathbf{I}_k + \mathbf{a}\mathbf{b}^\top + \mathbf{b}\mathbf{a}^\top + s\mathbf{b}\mathbf{b}^\top\right| &= \left(1 + \mathbf{b}^\top(\mathbf{I}_k + \mathbf{b}\mathbf{a}^\top)^{-1}(\mathbf{a} + s\mathbf{b})\right)\left|\mathbf{I}_k + \mathbf{b}\mathbf{a}^\top\right| \\ &= \left(1 + \mathbf{b}^\top(\mathbf{I}_k + \mathbf{b}\mathbf{a}^\top)^{-1}(\mathbf{a} + s\mathbf{b})\right)\left(1 + \mathbf{b}^\top\mathbf{a}\right). \end{aligned}$$

Applying the Sherman–Morrison formula (see Corollary 18.2.10 of [24]), we obtain

$$(\mathbf{I}_r + \mathbf{b}\mathbf{a}^\top)^{-1} = \mathbf{I}_r - \frac{\mathbf{b}\mathbf{a}^\top}{1 + \mathbf{a}^\top\mathbf{b}}.$$

By putting the above together we obtain the result. □

Next, we demonstrate that the identity in the above lemma is equivalent to a seemingly stronger one.

Corollary 5.1. For $s, t, v \in \mathbb{C}$ we have

$$\left|\mathbf{I}_r + v\mathbf{a}\mathbf{a}^\top + t\left(\mathbf{a}\mathbf{b}^\top + \mathbf{b}\mathbf{a}^\top\right) + s\mathbf{b}\mathbf{b}^\top\right| = \\ = 1 + v\mathbf{a}^\top\mathbf{a} \; + 2t\mathbf{b}^\top\mathbf{a} + \left(t^2 - vs\right)\left(\left(\mathbf{b}^T\mathbf{a}\right)^2 - \mathbf{a}^\top\mathbf{a}\mathbf{b}^\top\mathbf{b}\right) + s\mathbf{b}^T\mathbf{b}. \tag{5.5}$$

Proof. To see that the above is implied by (5.4), we first substitute inthe latter $u\mathbf{a}$ for $\mathbf{a}$, where $u \in \mathbb{C}$, leading to

$$\left|\mathbf{I}_r + u\mathbf{a}\mathbf{b}^\top + \mathbf{b}\mathbf{a}^\top + s\mathbf{b}\mathbf{b}^\top\right| = \left(1 + u\mathbf{b}^\top\mathbf{a}\right)^2 + \mathbf{b}^\top\mathbf{b}\left(s - u^2\mathbf{a}^\top\mathbf{a}\right).$$

Then, we substitute $\mathbf{a}+\mathbf{b}$ for $\mathbf{b}$ in the above expression, so that the left-hand-side becomes

$$\left|\mathbf{I}_r + u\left(2\mathbf{a}\mathbf{a}^\top + \mathbf{a}\mathbf{b}^\top + \mathbf{b}\mathbf{a}^\top\right) + s\left(\mathbf{a}\mathbf{a}^\top + \mathbf{a}\mathbf{b}^\top + \mathbf{b}\mathbf{a}^\top + \mathbf{b}\mathbf{b}^\top\right)\right| = \\ = \left|\mathbf{I}_r + (2u+s)\,\mathbf{a}\mathbf{a}^\top + (u+s)\left(\mathbf{a}\mathbf{b}^\top + \mathbf{b}\mathbf{a}^\top\right) + s\mathbf{b}\mathbf{b}^\top\right|.$$

The same substitution in the right-hand-side yields

$$1+(2u+s)\,\mathbf{a}^\top\mathbf{a}+2(u+s)\mathbf{b}^\top\mathbf{a}+u^2\left(\mathbf{a}^\top\mathbf{a}+\mathbf{b}^\top\mathbf{a}\right)^2+s\mathbf{b}^\top\mathbf{b}-u^2\mathbf{a}^\top\mathbf{a}\left(\mathbf{a}^\top\mathbf{a}+2\mathbf{a}^\top\mathbf{b}+\mathbf{b}^\top\mathbf{b}\right) \\ = 1 + (2u+s)\,\mathbf{a}^\top\mathbf{a} + 2(u+s)\mathbf{b}^\top\mathbf{a} + u^2\left(\mathbf{a}^\top\mathbf{a}\right)^2 + 2u^2\mathbf{a}^\top\mathbf{a}\mathbf{b}^\top\mathbf{a} + u^2\left(\mathbf{b}^\top\mathbf{a}\right)^2 + s\mathbf{b}^\top\mathbf{b} - \\ - u^2\left(\mathbf{a}^\top\mathbf{a}\right)^2 - 2u^2\mathbf{a}^\top\mathbf{a}\mathbf{b}^\top\mathbf{a} - u^2\mathbf{a}^\top\mathbf{a}\mathbf{b}^\top\mathbf{b} = \\ = 1 + (2u+s)\,\mathbf{a}^\top\mathbf{a} + 2(u+s)\mathbf{b}^\top\mathbf{a} + u^2\left(\mathbf{b}^\top\mathbf{a}\right)^2 + s\mathbf{b}^\top\mathbf{b} - u^2\mathbf{a}^\top\mathbf{a}\mathbf{b}^\top\mathbf{b}.$$

Then 5.5 follows by first replacing $\mathbf{a}$ by $\mathbf{a}/(u+s)$, $\tilde{v} = (2u+s)/(u+s)^2$, then $\mathbf{a}$ by $t\mathbf{a}$ and $v = \tilde{v}/t^2$. □

Remark 5.2. It should be noted that there are MGAL distributions that are not MAL, so these two classes of distributions do not coincide. This is because of the special structure of the covariance matrix $\mathbf{\Sigma}$ in (5.3). We can further infer from the above result which MGAL distributions are MAL and which ones are not. For example, if $k = 2$ and $\mathbf{X} \sim \mathcal{MGAL}_{2,1}(\mathbf{A}, \mathbf{\Sigma}, 1, \alpha)$ is also MAL, where $\mathbf{\Sigma} = (\sigma_{ij})$, then by (4.8) we must have that $\mathbf{X} \sim \mathcal{MAL}_{2,1}(\alpha; \mathbf{A}, \sigma^2)$ and

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} a_2^2 + 2\sigma^2 & -a_1a_2 \\ -a_1a_2 & a_1^2 + 2\sigma^2 \end{pmatrix}. \tag{5.6}$$

This shows that the entries of the vector $\mathbf{A} = (a_1\ a_2)^\top$ and the matrix $\mathbf{\Sigma} = (\sigma_{ij})$ must satisfy the following relations:

$$\sigma_{11} = \frac{1}{2}a_2^2 + \sigma^2, \;\; \sigma_{22} = \frac{1}{2}a_1^2 + \sigma^2, \;\; \sigma_{12} = -\frac{1}{2}a_1a_2. \tag{5.7}$$

Careful analysis of the equations in (5.7) shows that if the matrix $\mathbf{\Sigma} = (\sigma_{ij})$ is diagonal, then we must have either $a_1 = a_2 = 0$ (for $\sigma_{11} = \sigma_{22}$), $a_1 = 0$ and $a_2 = \pm\sqrt{2(\sigma_{11} - \sigma_{22})}$ (for $\sigma_{11} > \sigma_{22}$), or $a_2 = 0$ and $a_1 = \pm\sqrt{2(\sigma_{22} - \sigma_{11})}$ (for $\sigma_{22} > \sigma_{11}$). In all these cases we also have $\sigma^2 = \min(\sigma_{11}, \sigma_{22})$. On the other hand, if $\sigma_{12} \neq 0$, then the equations in

(5.7) produce

$$
\begin{aligned}
a_1^2 &= \sigma_{22} - \sigma_{11} + \sqrt{(\sigma_{11} - \sigma_{22})^2 + 4\sigma_{12}} \\
a_2^2 &= \sigma_{11} - \sigma_{22} + \sqrt{(\sigma_{11} - \sigma_{22})^2 + 4\sigma_{12}} \\
\sigma^2 &= \frac{1}{2}\left\{\sigma_{11} + \sigma_{22} - \sqrt{(\sigma_{11} - \sigma_{22})^2 + 4\sigma_{12}}\right\}.
\end{aligned}
$$

The signs of a_1 and a_2 are determined from the last equation in (5.7): if $\sigma_{12} < 0$ then a_1 and a_2 must have the same sign and their signs must be opposite if $\sigma_{12} > 0$. See Figure 1 for an example of the graphical illustration of such a distribution.

We conclude with a few additional comments and open questions. First, let us note that all Type I MGAL distributions are infinitely divisible (ID), since their ChFs (2.14) are valid for any positive α. However, it is unclear if this property is shared by Type II distributions, since their ChF (3.15) has been defined only for $\alpha > (k-1)/2$. It is unclear whether the expression in (3.15) is a genuine ChF for all $\alpha \in (0, (k-1)/2]$. Our results show that this is the case when $n = 1$, $k \in \mathbb{N}$, and arbitrary $\mathbf{A}$. In both cases the vector-valued MAL distributions are ID. However, the ID of general Type II matrix valued generalized Laplace distributions is an important open problem. The relation between the two classes of distributions in the general case of $k \times n$ matrices where $n > 1$ is, at the moment, a completely open question as well.

6. Summary

We considered extensions of the GAL distribution to the matrix variate case, which arise by mixing matrix variate normal distribution with respect to one of its covariance matrix parameters. When the distribution of the latter has to do with one-dimensional scaling of the original normal covariance matrix, and the univariate random scaling factor is gamma distributed, we obtain what we call *Type I matrix variate GAL distribution.* On the other hand, *Type II matrix variate GAL* is obtained by mixing the normal covariance with respect to *matrix variate* gamma distribution. While Type I matrix variate GAL distribution has been studied before (see [39]), there is no comprehensive account of Type II in the literature, except for their rather brief treatment as a special case of matrix variate generalized hyperbolic distributions, studied in [35] and [23]. With this work we filled this gap, and provided an account for basic distributional properties of Type II matrix variate GAL distributions. In particular, we derived their probability density function and the characteristic function, and provided stochastic representations related to matrix variate gamma distribution. We also showed that this distribution is closed under linear transformations, and studied the relevant marginal distributions. In addition, we also briefly accouned for Type I and discussed the interconnections with Type II. We hope that this work will be useful in the areas where matrix variate distributions provide an appropriate probabilistic tool for *three-way* data sets, which can arise in a variety of ways across different applications.

Bibliography

1. L. Anderlucci and C. Viroli, *Covariance pattern mixture models for the analysis of multivariate heterogeneous longitudinal data*, The Annals of Applied Statistics **9** (2015), no. 2, 777–800.
2. I. Anis, M. Arashi, and S. M. M. Tabatabaey, *On conditional applications of matrix variate normal distribution*, Iranian Journal of Mathematical Sciences and Informatics **5** (2010), no. 2, 33–43.
3. H. Asgharian, K. Podgórski, F. N. Shariati, and L. Liu, *Structural Multivariate Spatial Econometrics: Application to Cross-Country Interdependence of Stock and Bond markets*, Available at SSRN: https://ssrn.com/abstract=3527346, 2018.
4. O. E. Barndorff-Nielsen, J. Kent, and M. Sørensen, *Normal variance-mean mixtures and z distributions*, International Statistical Review **50** (1982), no. 2, 145–159.
5. O. E. Barndorff-Nielsen, V. Pérez-Abreu, and A. Rocha-Arteaga, *MatG random matrices*, Stochastic Models **22** (2006), 723–734.
6. A. Bekker and J. J. J. Roux, *Bayesian multivariate normal analysis with a Wishart prior*, Communications in Statistics - Theory and Methods **24** (1995), no. 10, 2485–2497.
7. T. Bodnar, S. Mazur, S. Muhinyuza, and N. Parolya, *On the product of a singular Wishart matrix and a singular Gaussian vector in high dimension*, Theory of Probability and Mathematical Statistics **99** (2018), no. 2, 37–50.
8. T. Bodnar, S. Mazur, and Y. Okhrin, *Distribution of the product of singular Wishart matrix and normal vector*, Theory of Probability and Mathematical Statistics **91** (2014), 1–15.
9. T. Bodnar, S. Mazur, and K. Podgórski, *Singular inverse Wishart distribution and its application to portfolio theory*, Journal of Multivariate Analysis **143** (2016), 314–326.
10. Y. M. Bulut and O. Arslan, *Matrix variate slash distribution*, Journal of Multivariate Analysis **137** (2015), 173–178.
11. A. P. Dawid, *Some matrix-variate distribution theory: Notational considerations and a Bayesian application*, Biometrika **68** (1981), no. 1, 265–274.
12. J. M. Dickey, *Matricvariate generalizations of the multivariate t distribution and the inverted multivariate t distribution*, The Annals of Mathematical Statistics **38** (1967), no. 2, 511–518.
13. J. Ding and A. Zhou, *Eigenvalues of rank-one updated matrices with some applications*, Applied Mathematics Letters **20** (2007), no. 12, 1223–1226.
14. P. Ding and J. K. Blitzstein, *On the Gaussian mixture representation of the Laplace distribution*, The American Statistician **72** (2018), no. 2, 172–174.
15. F. Z. Doğru, Y. M. Bulut, and O. Arslan, *Finite mixtures of matrix variate t distributions*, Gazi University Journal of Science **29** (2016), no. 2, 335–341.
16. M. P. B. Gallaugher and P. D. McNicholas, *A matrix variate skew-t distribution*, Stat **6** (2017), no. 1, 160–170.
17. ———, *Finite mixtures of skewed matrix variate distributions*, Pattern Recognition **80** (2018), 83–93.
18. ———, *Three skewed matrix variate distributions*, Statistics & Probability Letters **145** (2019), 103–109.
19. ———, *Mixtures of skewed matrix variate bilinear factor analyzers*, Advances in Data Analysis and Classification **14** (2020), no. 2, 415–434.
20. S. Geisser, *Bayesian estimation in multivariate analysis*, The Annals of Mathematical Statistics **36** (1965), no. 1, 150–159.
21. A. K. Gupta and D. K. Nagar, *Matrix variate distributions*, Chapman and Hall/CRC, Boca Raton, 2000.
22. A. K. Gupta and T. Varga, *Normal mixture representations of matrix variate elliptically contoured distributions*, Sankhyā: The Indian Journal of Statistics, Series A (1961-2002) **57** (1995), no. 1, 68–78.

23. S. W. Harrar, E. Seneta, and A. K. Gupta, *Duality between matrix variate t and matrix variate V.G. distributions*, Journal of Multivariate Analysis **97** (2006), no. 6, 1467–1475.
24. D. A. Harville, *Matrix algebra from a statistician perspective*, Springer, 1997.
25. C. S. Herz, *Bessel functions of matrix argument*, Annals of Mathematics **61** (1955), no. 3, 474–523.
26. G. B. M. Kibria, *The matrix-t distribution and its applications in predictive inference*, Journal of Multivariate Analysis **97** (2006), no. 3, 785–795.
27. S. Kotz, T. J. Kozubowski, and K. Podgórski, *The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering and finance*, Birkhauser, Boston, 2001.
28. T. J. Kozubowski, S. Mazur, and K. Podgórski, *Singular matrix variate gamma distributions and their convolution properties*, Working paper, Örebro University, 2022.
29. T. J. Kozubowski and K. Podgórski, *Gaussian mixture representation of the Laplace distribution revisited: Bibliographical connections and extensions*, The American Statistician **74** (2020), no. 4, 407–412.
30. T. J. Kozubowski, K. Podgórski, and I. Rychlik, *Multivariate generalized Laplace distribution and related random fields*, Journal of Multivariate Analysis **113** (2013), 59–72, Special Issue on Multivariate Distribution Theory in Memory of Samuel Kotz.
31. D. B. Madan and E. Seneta, *The variance gamma (V.G.) model for share market returns*, The Journal of Business **63** (1990), no. 4, 511–524.
32. P.-A. Mattei, *Multiplying a Gaussian matrix by a Gaussian vector*, Statistics & Probability Letters **128** (2017), 67–70.
33. V. Melnykov and X. Zhu, *On model-based clustering of skewed matrix data*, Journal of Multivariate Analysis **167** (2018), 181–194.
34. V. Pérez-Abreu and R. Stelzer, *Infinitely divisible multivariate and matrix gamma distributions*, Journal of Multivariate Analysis **130** (2014), 155–175.
35. L. Thabane and M. S. Haq, *On the matrix-variate generalized hyperbolic distribution and its Bayesian applications*, Statistics **38** (2004), no. 6, 511–526.
36. G. Z. Thompson, R. Maitra, W. Q. Meeker, and A. F. Bastawros, *Classification with the matrix-variate-t distribution*, Journal of Computational and Graphical Statistics **29** (2020), no. 3, 668–674, PMID: 33716477.
37. C. Viroli, *Finite mixtures of matrix normal distributions for classifying three-way data*, Statistics and Computing **21** (2011), no. 4, 511–522.
38. ______, *Model based clustering for three-way data structures*, Bayesian Analysis **6** (2011), no. 4, 573–602.
39. Y. Yurchenko, *Matrix variate and tensor variate laplace distributions*, (2021).

Department of Mathematics and Statistics, University of Nevada, NV-89557 RENO, USA
Email address: tkozubow@unr.edu

Department of Statistics, Örebro University, SE-70182 Örebro, Sweden,
Department of Economics and Statistics, Linnaeus University, SE-35195 Växjö, Sweden
Email address: Stepan.Mazur@oru.se

Department of Statistics, Lund University, SE-22007 Lund, Sweden
Email address: Krzysztof.Podgorski@stat.lu.se