Stochastic Models in Probability and Statistics
Vol. 3, No. 1, 2026, pp. 13–27. Research Paper

Multimatrix variate distributions

José A. Díaz-García[†]* **and Francisco J. Caro-Lopera**[‡]

[†] *Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua,Periférico Francisco R. Almada Km 1, Zootecnia, Chihuahua, 33820, Chihuahua, México*
[‡] *Institute of Basic Sciences and Faculty of Engineering, University of Medellin, Carrera 87 No.30-65, Medellin, Antioquia, Colombia*

Email(s): jadiaz@cimat.mx, fjcaro@udemedellin.edu.co

Abstract. A new family of distributions indexed by the class of matrix-variate elliptically contoured distributions is proposed, extending several known bimatrix-variate models. These multimatrix-variate distributions open new perspectives in classical distribution theory, which is typically grounded in probabilistic independence assumptions and frequently relies on unverified fitting laws. Most of the multimatrix models derived here are invariant under the spherical family, a property that facilitates inference and reduces the need for prior specification of the underlying distributions. This invariance clarifies the statistical methodology and addresses certain limitations present in existing approaches, such as copula-based models. The paper also presents a variety of special cases, fundamental properties, and generalizations. The proposed joint distributions allow flexible combinations across scalars, vectors, and matrices, making them adaptable to complex modeling requirements. Moreover, they are computationally tractable, enabling a wide range of practical applications. In particular, we provide a comprehensive example in molecular docking for SARS-CoV-2, illustrating the analysis of matrix-dependent samples.

Keywords: Bimatrix variate, Matricvariate, Matrix variate, Matrix variate elliptically contoured distributions, Random matrices.

1 Introduction

The modern world and its paradigms for understanding natural phenomena increasingly demand holistic solutions that involve the consensus of multiple sciences. The so-termed new way of doing science, the scientific computing, is based on the postulate that the phenomena of post modernity intertwine multiple dependencies between knowledge and variables. The paradox of the butterfly's flapping effect, for example, confirms that the independence of events in a modelable scientific phenomenon is increasingly difficult to accept in the scientific community. From the epistemological point of view, a construct of independent variables immensely distances reality from the model that seeks to explain it. Unless an expert manages to completely isolate its universe of explanatory variables of a phenomenon and has proven certainty of its independence, it is generally difficult from the usual theory to arrive at an explanation of the interrelationship in question. Chaos theory and the divergence of non-deterministic phenomena bring

*Corresponding author

Received: 19 August 2025 / Accepted: 16 February 2026
DOI: 10.22067/smps.2026.94999.1050

modern statistics closer to the elucidation of techniques increasingly in line with the divergent and integrated thinking of the expert. Initially, statistics moved away from interdisciplinary work with science, assuming theories based on probabilistic independence due to the great convenience in the computational solution. Such is the case of the well-known linear models, where the assumption of independent variables allows results as powerful, but difficult to achieve in reality, such as the Gauss-Markov theorem. But perhaps the example of greatest mathematical convenience and widely disseminated in statistics corresponds to the information from probabilistically independent samples that immensely facilitate a definition of tractable likelihood formed by the simple product of the probability densities of each variable. Experts, who deep down know of a high dependence on the samples of their experiments, have had no alternative but to use classical likelihood to approximate the highly dependent phenomena that emerge from their laboratories. A more robust modeling of the real world undoubtedly requires greater planning and statistical analysis of experiments in an interdisciplinary dialogue with the expert, who dictates the main postulates of the phenomena they are trying to decipher. In this sense, statistics does not command decisions and applications, but rather realistically guides the solution that the expert guides. Instead of establishing its classic postulates about independence, normality, estimability, linearity, it seeks to resolve the paradigms that emerge from modern modeling; methodological proposals that increasingly speak of the opposite: dependence, skewed or non-Gaussian distributions, functions of multiple local minima or maxima, nonlinearity and nonparametric statistics, etc. In response to these requirements, statistics has proposed separate theories of great complexity, but they are still not very popular and only exist in the literature that is also frequented by statisticians, not by users of other sciences. The best example speaks of normality, which, like the independence of variables, is almost never met. The normal model has been replaced by elliptical contour distributions, allowing probability distributions to be considered with heavier and lighter tails and a greater or lesser degree of kurtosis, among other significant characteristics. Although a major problem emerges when the underlying distribution is not known a priori, then the decision is left to adjustment criteria, without any explanation from expert science. In this context, invariances under complete families of distributions emerge of utmost importance, so that the results are applied regardless of knowledge of the base distribution.

In some experiments and phenomena where more than one random variable is studied, say X e Y, it is assumed that these variables are probabilistically independent with densities $dF_X(x)$ and $dG_Y(y)$. Then, if we are interested in the joint density of X and Y, this reduces to the elementary product $dF_X(x)dG_Y(y)$. However, as we have said, modern chaos theories and divergent modeling of complex holistic problems warn against this assumption of independence, in which case finding the joint density is not an easy task. For example, in a study of climatological data it is of interest to study the random variables associated with temperature and precipitation. For practical purposes these variables are considered independent, with normal and gamma distributions, respectively. But it is known that the temperature and precipitation are not probabilistically independent random variables. Then we open the question about a joint distribution of X and Y when the variables are not probabilistically independent. Furthermore, it would be desirable that given the joint density function of the probabilistically dependent variables X and Y, the corresponding marginal distributions of the random variables X and Y were the marginal distributions that are usually assumed under independence. For the univariate case, several bivariate distributions have been studied by different authors, the models include gamma-beta, gamma-gamma, beta-beta among others. They have found applications in problems of hydrology, finance and other areas disciplines, see Libby and Novick (1982), Chen and Novick (1984), Nadarajah (2007, 2013), Olkin and Liu (2003) and Sarabia et al. (2014) and references therein. In general, in these works, joint distributions of dependent random variables are constructed, making changes of variables on independent random variables. From the matrix point of view bimatrix variate type distribution have been studied by several authors, see Olkin and Rubin (1964), Díaz-García, and Gutiérrez-Jáimez (2010a,b, 2011), and Bekker et al. (2011) among others. Ehlers (2011) presents a very complete exposition of the advances in matrix variate case up to that moment. The authors have studied this problem for scalar and vector cases in Díaz-García et al. (2022) and a matrix case in the multimatricvariate version in Díaz-García and Caro-Lopera (2022). Finally, a multimatrix theory that solves the previous problem but that cannot be computed would add to the hundreds of theoretical results in matrix analysis developed in the last 70 years. Then we expect that the joint distribution functions of several matrices are free of approximations and truncation of complex series of polynomials of matrix arguments.

The above discussion is placed in the present article by proposing a new method to establish computable joint density functions of two or more random matrices, which shall be termed *multimatrix variate distributions*, such that the corresponding marginal distributions are known. Section 1 establishes the notation to be used and some preliminary results on the calculation of Jacobians and some integrals of interest as well as the definition of the matrix variate elliptically contoured distributions. The main distributional results are collected in Section 2. Some related distributions of multimatrix variate distributions and particular generalisations are obtained in Section 3. Finally, Section 4 provides a dependent sample emergent from a problem of molecular docking into a new cavity of SARS-CoV-2. In particular, the joint distribution of the dependent sample is estimated by a simple optimisation routine.

2 Preliminary results

Matrix notations, matrix variate elliptically contoured distributions, and Jacobians are presented in this section. First, we start with notations and terminologies. $\mathbf{A} \in \Re^{n\times m}$ denotes a *matrix* with n rows and m columns; $\mathbf{A}' \in \Re^{m\times n}$ is the *transpose matrix*, and if $\mathbf{A} \in \Re^{n\times n}$ has an *inverse*, it shall be denoted by $\mathbf{A}^{-1} \in \Re^{n\times n}$. $\mathbf{A} \in \Re^{n\times n}$ is a *symmetric matrix* if $\mathbf{A} = \mathbf{A}'$. If all their eigenvalues are positive then $\mathbf{A}$ is a *positive definite matrix*, a fact denoted as $\mathbf{A} > \mathbf{0}$. An *identity matrix* shall be denoted by $\mathbf{I} \in \Re^{n\times n}$. To specify the size of the identity, we shall use $\mathbf{I}_n$. $\operatorname{tr}(\mathbf{A})$ denotes the trace of matrix $\mathbf{A} \in \Re^{m\times m}$. If $\mathbf{A} \in \Re^{n\times m}$ then by $\operatorname{vec}(\mathbf{A})$ we mean the $mn\times 1$ vector formed by stacking the columns of $\mathbf{A}$ under each other; that is, if $\mathbf{A} = [\mathbf{a}_1\mathbf{a}_2\ldots\mathbf{a}_m]$, where $\mathbf{a}_j \in \Re^{n\times 1}$ for $j = 1,2,\ldots,m$

$$\operatorname{vec}(\mathbf{A}) = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}.$$

The *Frobenius norm* of a matrix $\mathbf{A}$ shall be denoted as $||\mathbf{A}||$. Typically the Frobenius norm is denoted as $||\mathbf{A}||_F$, to differentiate it from other matrix norms. Since we shall use only the Frobenius norm, it just be denoted as $||\mathbf{A}||$. It is defined by

$$||\mathbf{A}|| = \sqrt{\operatorname{tr}(\mathbf{A}'\mathbf{A})} = \sqrt{\operatorname{vec}'(\mathbf{A})\operatorname{vec}(\mathbf{A})}.$$

Finally, $\mathscr{V}_{m,n}$ denotes the *Stiefel manifold*, the space of all matrices $\mathbf{H}_1 \in \Re^{n\times m}$ $(n \geq m)$ with orthogonal columns, that is, $\mathscr{V}_{m,n} = \{\mathbf{H}_1 \in \Re^{n\times m}; \mathbf{H}_1'\mathbf{H}_1 = \mathbf{I}_m\}$. In addition, if $(\mathbf{H}_1' d\mathbf{H}_1)$ defines an *invariant measure on the Stiefel manifold* $\mathscr{V}_{m,n}$, from Theorem 2.1.15, p. 70 in Muirhead (2005),

$$\int_{\mathscr{V}_{m,n}} (\mathbf{H}_1' d\mathbf{H}_1) = \frac{2^m \pi^{mn/2}}{\Gamma_m[n/2]}, \tag{1}$$

where $\Gamma_m[a]$ denotes the multivariate gamma function, see (Muirhead, 2005, Definiton 2.1.10, p. 61)

Next a result about Jacobians is derived.

Goodall and Mardia (1993) proposed the term *matrix multivariate* instead of *matrix variate*, commonly used in multivariate analysis literature, Gupta and Varga (1993), among many other authors. The term *matricvariate* was introducing by Dickey (1967) to distinguish between the two possible expressions of the density matrix multivariate T-distribution. A detail explanation about this topic is given in (Díaz-García, and Gutiérrez-Jáimez, 2006, see Section 3).

Theorem 1. *Suppose that* $\mathbf{X} \in \Re^{n\times m}$ *and* $\mathbf{Y} \in \Re^{n\times m}$ *are matrices with mathematically independent elements.*

i) *Let* $\mathbf{Y} = (1-\operatorname{tr}\mathbf{X}'\mathbf{X})^{-1/2}\mathbf{X}$. *Then*

$$(d\mathbf{Y}) = (1-\operatorname{tr}\mathbf{X}'\mathbf{X})^{-(nm/2+1)}(d\mathbf{X}).$$

ii) *If* $\mathbf{X} = (1+\operatorname{tr}\mathbf{Y}'\mathbf{Y})^{-1/2}\mathbf{Y}$, *we have*

$$(d\mathbf{X}) = (1+\operatorname{tr}\mathbf{Y}'\mathbf{Y})^{-(nm/2+1)}(d\mathbf{Y}).$$

Proof. **i**) The proof follows from Theorem 1 in Díaz-García et al. (2022), observing that $\mathbf{Y} = (1-\operatorname{tr}\mathbf{X}'\mathbf{X})^{-1/2}\mathbf{X}$ can be written as $\mathbf{y} = (1-\mathbf{x}'\mathbf{x})^{-1/2}\mathbf{x}$, where $\mathbf{x} = \operatorname{vec}\mathbf{X} \in \Re^{mn}$ and $\mathbf{y} = \operatorname{vec}\mathbf{Y} \in \Re^{mn}$. **ii**). Its proof is analogous to one given for **i**). □

Now, let $\mathbf{V} \in \Re^{N\times m}$ random matrix with a matrix variate elliptically contoured distribution with respect to the Lebesgue measure $(d\mathbf{V})$. Therefore its density function is given by

$$dF_{\mathbf{V}}(\mathbf{V}) = |\boldsymbol{\Sigma}|^{-N/2}|\boldsymbol{\Theta}|^{-m/2}h\left\{\operatorname{tr}\left[(\mathbf{V}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{V}-\boldsymbol{\mu})\boldsymbol{\Theta}^{-1}\right]\right\}(d\mathbf{V}). \tag{2}$$

The location parameter is $\boldsymbol{\mu} \in \Re^{N\times m}$; and the scale parameters $\boldsymbol{\Sigma} \in \Re^{N\times N}$ and $\boldsymbol{\Theta} \in \Re^{m\times m}$, are positive definite matrices. The distribution shall be denoted by $\mathbf{V} \sim \mathscr{E}_{N\times m}(\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Theta};h)$, and indexed by the kernel function h: $\Re \to [0,\infty)$, where

$$\int_0^\infty u^{Nm/2-1}h(u)du < \infty.$$

When $\boldsymbol{\mu} = \mathbf{0}$, $\boldsymbol{\Sigma} = \mathbf{I}_N$ and $\boldsymbol{\Theta} = \mathbf{I}_m$ a special case of a matrix variate elliptically contoured distribution appears, in this case it is said that $\mathbf{V}$ has a *matrix variate spherical distribution.*

Note that for constant $a \in \Re$, the substitution $v = u/a$ and $(du) = a(dv)$ in (Fang et al., 1990, Equation 2.21, p. 26) provides that

$$\int_{v>0} v^{Nm/2-1}h(av)(dv) = \frac{a^{-Nm/2}\Gamma_1[Nm/2]}{\pi^{Nm/2}}. \tag{3}$$

3 Multimatrix variate distributions

Assume that $\mathbf{X} \sim \mathscr{E}_{N\times m}(\mathbf{0},\mathbf{I}_N,\mathbf{I}_m;h)$, such that $n_0+n_1+\cdots+n_k = N$, and $\mathbf{X} = \left(\mathbf{X}_0',\mathbf{X}_1',\dots,\mathbf{X}_k'\right)'$. Then (2) can be written as

$$dF_{\mathbf{X}_0,\mathbf{X}_1,\dots,\mathbf{X}_k}(\mathbf{X}_1,\dots,\mathbf{X}_k) = h[\operatorname{tr}(\mathbf{X}_0'\mathbf{X}_0+\mathbf{X}_1'\mathbf{X}_1+\cdots+\mathbf{X}_k'\mathbf{X}_k)]\bigwedge_{i=0}^{k}(d\mathbf{X}_i), \tag{4}$$

or in terms of the Frobenius norm, we can rewrite (4) as

$$dF_{\mathbf{X}_0,\mathbf{X}_1,\dots,\mathbf{X}_k}(\mathbf{X}_1,\dots,\mathbf{X}_k) = h\left(||\mathbf{X}_0||^2+||\mathbf{X}_1||^2+\cdots+||\mathbf{X}_k||^2\right)\bigwedge_{i=0}^{k}(d\mathbf{X}_i),$$

where $\mathbf{X}_i \in \Re^{n_i\times m}$, $n_i \geq m \geq 1$, $i = 0,1,\dots,k$. Note that the random matrices $\mathbf{X}_0,\mathbf{X}_1,\dots,\mathbf{X}_k$ are probabilistically dependent. Furthermore, only under a matrix variate normal distribution, the random matrices are independent, see Fang and Zhang (1990).

Analogously to (Díaz-García et al., 2022, Equation 4.2) but now defining $V_i = ||\mathbf{X}_i||^2$, $i = 0, 1, \dots, k$, it is obtained

$$dF_{V_0,\dots,V_k}(v_0,\dots,v_k) = \frac{\pi^{Nm/2}}{\prod_{i=0}^{k}\Gamma[n_i m/2]} h\left(\sum_{i=0}^{k} v_i\right)\prod_{i=0}^{k} v_i^{n_i m/2-1}\bigwedge_{i=0}^{k}(du_i),$$

where $N = n_0 + \cdots + n_k$, this distribution shall be termed *multivariate generalised gamma distribution* and its univariate case ($i = 0$) shall be termed *generalised gamma distribution.*

Theorem 2. *Suppose that* $\mathbf{X} = \left(\mathbf{X}_0', \dots, \mathbf{X}_k'\right)'$ *has a matrix variate spherical distribution, with* $\mathbf{X}_i \in \Re^{n_i \times m}$, $n_i \geq m$, $i = 0, 1, \dots, k$. *Define* $V = ||\mathbf{X}_0||^2$. *Then, the joint density* $dF_{V,\mathbf{X}_1,\dots,\mathbf{X}_k}(v, \mathbf{X}_1, \dots, \mathbf{X}_k)$ *is given by*

$$\frac{\pi^{n_0 m/2}}{\Gamma_1[n_0 m/2]} h\left[v + \sum_{i=1}^{k}||\mathbf{X}_i||^2\right] v^{n_0 m/2-1}(dv)\bigwedge_{i=1}^{k}(d\mathbf{X}_i), \tag{5}$$

where $V > 0$. *This distribution shall be termed* multimatrix variate generalised gamma-elliptical distribution.

Proof. We have that

$$dF_{\mathbf{X}_0,\mathbf{X}_1,\dots,\mathbf{X}_k}(\mathbf{X}_1,\dots,\mathbf{X}_k) = h\left(||\mathbf{X}_0||^2 + ||\mathbf{X}_1||^2 + \cdots + ||\mathbf{X}_k||^2\right)\bigwedge_{i=0}^{k}(d\mathbf{X}_i).$$

Define $V = ||\mathbf{X}_0||^2$ and from (Muirhead, 2005, Theorem 2.1.14, p. 65), we have that

$$(d\mathbf{X}_0) = 2^{-1} v^{n_0 m/2-1}(dv)\wedge(\mathbf{h}_1 d\mathbf{h}_1),$$

where $\mathbf{h}_1 \in \mathscr{V}_{1,n_0 m}$. Thus, the joint density of $V, \mathbf{h}_1, \mathbf{X}_1, \dots, \mathbf{X}_k$ is

$$\frac{v^{n_0 m/2-1}}{2} h\left(v + ||\mathbf{X}_1||^2 + \cdots + ||\mathbf{X}_k||^2\right)(dv)\wedge(\mathbf{h}_1 d\mathbf{h}_1)\bigwedge_{i=1}^{k}(d\mathbf{X}_i).$$

Integration over $\mathbf{h}_1 \in \mathscr{V}_{1,n_0 m}$ by using (1) gives the required result. □

Theorem 3. *Assume that* $\mathbf{X} = \left(\mathbf{X}_0', \dots, \mathbf{X}_k'\right)'$ *has a matrix variate spherical distribution, with* $\mathbf{X}_i \in \Re^{n_i \times m}$, $n_i \geq m$, $i = 0, 1, \dots, k$. *Define* $V = ||\mathbf{X}_0||^2$ *and* $\mathbf{T}_i = V^{-1/2}\mathbf{X}_i$, $i = 1, \dots, k$. *The joint density* $dF_{V,\mathbf{T}_1,\dots,\mathbf{T}_k}(v, \mathbf{T}_1, \dots, \mathbf{T}_k)$ *is given by*

$$\frac{\pi^{n_0 m/2}}{\Gamma_1[n_0 m/2]} h\left[v\left(1 + \sum_{i=1}^{k}||\mathbf{T}_i||^2\right)\right] v^{Nm/2-1}(dv)\bigwedge_{i=1}^{k}(d\mathbf{T}_i), \tag{6}$$

where $N = n_0 + n_1 + \cdots + n_k$, $V > 0$ *and* $\mathbf{T}_i \in \Re^{n_i \times m}$, $i = 1, \dots, k$. *This distribution shall be termed* multimatrix variate generalised gamma-Pearson type VII distribution. *Moreover, the termed* multimatrix variate Pearson type VII *is the marginal density* $dF_{\mathbf{T}_1,\dots,\mathbf{T}_k}(\mathbf{T}_1, \dots, \mathbf{T}_k)$ *of* $\mathbf{T}_1, \dots, \mathbf{T}_k$ *and is given by*

$$\frac{\Gamma_1[Nm/2]}{\pi^{(N-n_0)m/2}\Gamma_1[n_0 m/2]}\left(1 + \sum_{i=1}^{k}||\mathbf{T}_i||^2\right)^{-Nm/2}\bigwedge_{i=1}^{k}(d\mathbf{T}_i), \tag{7}$$

where $\mathbf{T}_i \in \Re^{n_i \times m}$, $n_i \geq m$ *and* $N = n_0 + n_1 + \cdots + n_k$.

Proof. The density (6) follows from (5) by defining $\mathbf{T}_i = V^{-1/2}\mathbf{X}_i$, $i = 1,\dots,k$. Hence by (Muirhead, 2005, Theorem 2.1.4, p. 57), we have

$$\bigwedge_{i=1}^{k}(d\mathbf{X}_i) = \prod_{i=1}^{k} v^{n_i m/2} \bigwedge_{i=1}^{k}(d\mathbf{T}_i).$$

And then the required result follows. Now, integrating (6) over $V > 0$, by using (3), provides the density (7). □

Theorem 4. *Suppose that* $\mathbf{X} = \left(\mathbf{X}'_0,\dots,\mathbf{X}'_k\right)'$ *has a matrix variate spherical distribution, with* $\mathbf{X}_i \in \Re^{n_i\times m}$, $n_i \geq m$, $i = 0,1,\dots,k$. *Define* $V = ||\mathbf{X}_0||^2$ *and* $\mathbf{R}_i = \left(V + ||\mathbf{X}_i||^2\right)^{-1/2}\mathbf{X}_i$, $i = 1,\dots,k$. *Then the joint density of* $V,\mathbf{R}_1,\dots,\mathbf{R}_k$, *denoted as* $dF_{V,\mathbf{R}_1,\dots,\mathbf{R}_k}(v,\mathbf{R}_1,\dots,\mathbf{R}_k)$, *is given by*

$$\frac{\pi^{n_0 m/2}}{\Gamma_1[n_0 m/2]} v^{Nm/2-1} h\left[v\left(1+\sum_{i=1}^{k}\frac{||\mathbf{R}_i||^2}{(1-||\mathbf{R}_i||^2)}\right)\right]\prod_{i=1}^{k}\left(1-||\mathbf{R}_i||^2\right)^{-n_i m/2-1}(dv)\bigwedge_{i=1}^{k}(d\mathbf{R}_i), \tag{8}$$

where $N = n_0 + n_1 + \cdots + n_k$, $V > 0$ *and* $\mathbf{R}_i \in \Re^{n_i\times m}$, *and* $||\mathbf{R}_i||^2 \leq 1$ $i = 1,\dots,k$. *This distribution shall be termed* multimatrix variate generalised gamma-Pearson type II distribution. *Also, the marginal density* $dF_{\mathbf{R}_1,\dots,\mathbf{R}_k}(\mathbf{R}_1,\dots,\mathbf{R}_k)$ *is*

$$\frac{\Gamma_1[Nm/2]}{\pi^{(N-n_0)m/2}\Gamma_1[n_0 m/2]}\left[1+\sum_{i=1}^{k}\frac{||\mathbf{R}_i||^2}{(1-||\mathbf{R}_i||^2)}\right]^{-Nm/2}\prod_{i=1}^{k}\left(1-||\mathbf{R}_i||^2\right)^{-n_i m/2-1}\bigwedge_{i=1}^{k}(d\mathbf{R}_i), \tag{9}$$

which shall be termed multimatrix variate Pearson type II distribution.

Proof. Consider the substitution $\mathbf{T}_i = (1-||\mathbf{R}_i||^2)^{-1/2}\mathbf{R}_i$, $i = 1,\dots,k$ in (6), hence by Theorem 1,

$$\bigwedge_{i=1}^{k}(d\mathbf{T}_i) = \prod_{i=1}^{k}\left(1-||\mathbf{R}_i||^2\right)^{-n_i m/2-1}\bigwedge_{i=1}^{k}(d\mathbf{R}_i),$$

and the desired result follows. Now (9) follows by integration of (8) over $V > 0$ via (3). □

Theorem 5. *Consider the hypotheses of Theorem 3 and define* $\mathbf{F}_i = \mathbf{T}'_i\mathbf{T}_i > \mathbf{0}$, $i = 1,\dots,k$. *Then the joint density* $dF_{V,\mathbf{F}_1,\dots,\mathbf{F}_k}(v,\mathbf{F}_1,\dots,\mathbf{F}_k)$ *is given by*

$$\frac{\pi^{Nm/2}v^{Nm/2-1}}{\Gamma_1[n_0 m/2]}\prod_{i=1}^{k}\left(\frac{|\mathbf{F}_i|^{(n_i-m-1)/2}}{\Gamma_m[n_i/2]}\right)h\left[v\left(1+\sum_{i=1}^{k}\operatorname{tr}\mathbf{F}_i\right)\right](dv)\bigwedge_{i=1}^{k}(d\mathbf{F}_i). \tag{10}$$

This distribution shall be termed multimatrix variate generalised gamma-beta type II distribution. *The corresponding marginal density function* $dF_{\mathbf{F}_1,\dots,\mathbf{F}_k}(\mathbf{F}_1,\dots,\mathbf{F}_k)$ *is given by*

$$\frac{\Gamma_1[Nm/2]}{\Gamma_1[n_0 m/2]\prod_{i=1}^{k}\Gamma_m[n_i/2]}\prod_{i=1}^{k}|\mathbf{F}_i|^{(n_i-m-1)/2}\left(1+\sum_{i=1}^{k}\operatorname{tr}\mathbf{F}_i\right)^{-Nm/2}\bigwedge_{i=1}^{k}(d\mathbf{F}_i), \tag{11}$$

and shall be termed multimatrix variate beta type II distribution.

Proof. Joint density functions (10) and (11) follow from substitutions $\mathbf{F}_i = \mathbf{T}'_i\mathbf{T}_i$, $i = 1,\dots,k$ in (6) and (7), respectively. First, note that (Muirhead, 2005, Theorem 2.1.14, p. 66) provides

$$\bigwedge_{i=1}^{k}(d\mathbf{T}_i) = 2^{-mk}\prod_{i=1}^{k}|\mathbf{F}_i|^{(n_i-m-1)/2}\bigwedge_{i=1}^{k}(d\mathbf{F}_i)\bigwedge_{i=1}^{k}\left(\mathbf{H}'_{1_i}d\mathbf{H}_{1_i}\right),$$

where $\mathbf{H}_{1_i} \in \mathscr{V}_{n_i,m}$, $i = 1, \ldots, k$. Hence, integrating over $\mathbf{H}_{1_i} \in \mathscr{V}_{n_i,m}$ $i = 1, \ldots, k$ by (1) turns

$$\int_{\mathbf{H}_{1_1}} \cdots \int_{\mathbf{H}_{1_k}} \bigwedge_{i=1}^{k} \left(\mathbf{H}'_{1_i} d\mathbf{H}_{1_i}\right) = \frac{2^{mk} \pi^{(N-n_0)m/2}}{\displaystyle\prod_{i=1}^{k} \Gamma_m[n_i/2]},$$

and the required result is derived. □

A particular case of the density function (11) is proposed in (Muirhead, 2005, Problem 3.18, p.118).

Theorem 6. *Assuming that* $\mathbf{B}_i = \mathbf{R}'_i \mathbf{R}_i > \mathbf{0}$ *and* $\operatorname{tr} \mathbf{B}_i \leq 1$ *with* $i = 1, \ldots, k$, *in Theorem 4. Then the joint density* $dF_{V,\mathbf{B}_1,\ldots,\mathbf{B}_k}(v, \mathbf{B}_1, \ldots, \mathbf{B}_k)$ *is*

$$\frac{\pi^{Nm/2} \ v^{Nm/2-1}}{\Gamma_1[n_0 m/2]} \prod_{i=1}^{k} \left(\frac{|\mathbf{B}_i|^{(ni-m-1)/2}}{\Gamma_m[n_i/2]}\right) h\left[v\left(1 + \sum_{i=1}^{k} \frac{\operatorname{tr} \mathbf{B}_i}{(1 - \operatorname{tr} \mathbf{B}_i)}\right)\right] \prod_{i=1}^{k} (1 - \operatorname{tr} \mathbf{B}_i)^{-n_i m/2-1} (dv) \bigwedge_{i=1}^{k} (d\mathbf{B}_i).$$

This distribution shall be termed multimatrix variate generalised gamma-beta type I distribution*. Moreover, the density function* $dF_{\mathbf{B}_1,\ldots,\mathbf{B}_k}(\mathbf{B}_1, \ldots, \mathbf{B}_k)$ *can written as*

$$\frac{\Gamma_1[Nm/2]}{\Gamma_1[n_0 m/2]} \prod_{i=1}^{k} \left(\frac{|\mathbf{B}_i|^{(ni-m-1)/2}}{\Gamma_m[n_i/2]}\right) \left[1 + \sum_{i=1}^{k} \frac{\operatorname{tr} \mathbf{B}_i}{(1 - \operatorname{tr} \mathbf{B}_i)}\right]^{-Nm/2} \prod_{i=1}^{k} (1 - \operatorname{tr} \mathbf{B}_i)^{-n_i m/2-1} \bigwedge_{i=1}^{k} (d\mathbf{B}_i),$$

and this marginal distribution shall be called multimatrix variate beta type I distribution*.*

Proof. A similar procedure to the proof of Theorem 5 can be applied here. Just consider the joint density functions (8) and (9) under the change of variable $\mathbf{B}_i = \mathbf{R}'_i \mathbf{R}_i$ with $i = 1, \ldots, k$, and the derivation is complete. □

Finally, we have the following result.

Theorem 7. *Assume that* $\mathbf{X} = \left(\mathbf{X}'_0, \ldots, \mathbf{X}'_k\right)'$ *has a matrix variate spherical distribution, with* $\mathbf{X}_i \in \Re^{n_i \times m}$, $n_i \geq m$, $i = 0, 1, \ldots, k$. *Define* $V = ||\mathbf{X}_0||^2$ *and* $\mathbf{W}_i = \mathbf{X}'_i \mathbf{X}_i > \mathbf{0}$, $i = 1, \ldots, k$.
Then, the joint density $dF_{V,\mathbf{W}_1,\ldots,\mathbf{W}_k}(v, \mathbf{W}_1, \ldots, \mathbf{W}_k)$ *is given by*

$$\frac{\pi^{Nm/2} v^{n_0 m/2-1}}{\Gamma_1[n_0 m/2]} \prod_{i=1}^{k} \left(\frac{|\mathbf{W}_i|^{(ni-m-1)/2}}{\Gamma_m[n_i/2]}\right) h\left[v + \sum_{i=1}^{k} \operatorname{tr} \mathbf{W}_i\right] (dv) \bigwedge_{i=1}^{k} (d\mathbf{W}_i), \tag{12}$$

where $V > 0$. *This distribution shall be termed* multimatrix variate generalised gamma - generalised Wishart distribution*.*
Similarly, let $\mathbf{W}_i = \mathbf{X}'_i \mathbf{X}_i > \mathbf{0}$, $i = 1, \ldots, k$. *Therefore, the joint density* $dF_{\mathbf{W}_1,\ldots,\mathbf{W}_k}(\mathbf{W}_1, \ldots, \mathbf{W}_k)$ *is*

$$\pi^{(N-n_0)m/2} \left(\prod_{i=1}^{k} \frac{|\mathbf{W}_i|^{(ni-m-1)/2}}{\Gamma_m[n_i/2]}\right) h\left(\operatorname{tr} \sum_{i=1}^{k} \mathbf{W}_i\right) \bigwedge_{i=1}^{k} (d\mathbf{W}_i). \tag{13}$$

Proof. From Theorem 2 we have

$$\frac{\pi^{n_0 m/2}}{\Gamma_1[n_0 m/2]} h\left[v + \sum_{i=1}^{k} ||\mathbf{X}_i||^2\right] v^{n_0 m/2-1} (dv) \bigwedge_{i=1}^{k} (d\mathbf{X}_i).$$

Defining $\mathbf{W}_i = \mathbf{X}'_i \mathbf{X}_i$ with $i = 1, \ldots, k$ and proceeding as in the proof of Theorem 5, the result (12) is immediate. (13) is obtained in Díaz-García and Caro-Lopera (2022). □

Several multivariate density functions studied in the statistical literature are obtained as particular cases of the distributions proposed in this section. For example the distribution proposed in (Muirhead, 2005, Problem 3.18, p.118) and the distributions in (Gupta and Varga, 1993, Subsections 2.7.3.2 and 2.7.3.3, p. 54).

4 Some properties and generalization

This section focus on multimatrix variate distributions for two and three matrix arguments. Additionally, if the matrix arguments are non-singular, the corresponding inverse distributions are also obtained. Despite a laborious task, the proposed novel method can consider several marginal distributions. The content of the section is an useful combination for multiple possible uses in real phenomena guide by experts with the corresponding previous knowledge of the marginal distributions.

Theorem 8. *Assume that* $\mathbf{X} = (\mathbf{X}_0', \mathbf{X}_1', \mathbf{X}_2')'$ *has a matrix variate spherical distribution, with* $\mathbf{X}_i \in \Re^{n_i \times m}$, $n_i \geq m$, $i = 0,1,2$. *Define* $V = ||\mathbf{X}_0||^2$, $\mathbf{T} = V^{-1/2}\mathbf{X}_1$, *and* $\mathbf{R} = (V + ||\mathbf{X}_2||)^{-1/2}\mathbf{X}_2$. *The joint density* $dF_{V,\mathbf{T},\mathbf{R}}(v,\mathbf{T},\mathbf{R})$ *is given by*

$$\frac{\pi^{n_0 m/2} v^{Nm/2-1}}{\Gamma_1[n_0 m/2]} h\left\{v\left[1 + \left(1 - ||\mathbf{R}||^2\right)||\mathbf{T}||^2\right]\right\}\left(1 - ||\mathbf{R}||^2\right)^{(n_0+n_1)m/2-1}(dv) \wedge d\mathbf{T}) \wedge d\mathbf{R}). \tag{14}$$

where $N = n_0 + n_1 + n_2$, $V > 0$, $\mathbf{T} \in \Re^{n_1 \times m}$, $\mathbf{R} \in \Re^{n_2 \times m}$ *and* $||\mathbf{R}||^2 \leq 1$. *This distribution shall be termed* threematrix variate generalized gamma-Pearson type VII - Pearson type II distribution. *Moreover, the termed* bimatrix variate Pearson type VII - Pearson type II distribution *is the marginal density* $dF_{\mathbf{T},\mathbf{R}}(\mathbf{T},\mathbf{R})$ *of* $\mathbf{T},\mathbf{R}$ *and is given by*

$$\frac{\Gamma_1[Nm/2]}{\pi^{(N-n_0)m/2}\Gamma_1[n_0 m/2]}\left[(1 + \left(1 - ||\mathbf{R}||^2\right)||\mathbf{T}||^2\right]^{-Nm/2}\left(1 - ||\mathbf{R}||^2\right)^{(n_0+n_1)m/2-1}(d\mathbf{T}) \wedge d\mathbf{R}). \tag{15}$$

Proof. By Theorem 2

$$\frac{\pi^{n_0 m/2}}{\Gamma_1[n_0 m/2]} h\left[v + ||\mathbf{X}_1||^2 + ||\mathbf{X}_2||^2\right] v^{n_0 m/2-1}(dv) \wedge (d\mathbf{X}_1) \wedge (d\mathbf{X}_2),$$

where $V = ||\mathbf{X}_0||^2 > 0$. Let $V_0 = V + ||\mathbf{X}_2||^2$, $\mathbf{T} = V^{-1/2}\mathbf{X}_1$, and $\mathbf{R} = (V + ||\mathbf{X}_2||)^{-1/2}\mathbf{X}_2$. Hence $V = V_0 - ||\mathbf{X}_2||^2 = V_0 - ||V_0\mathbf{R}_2||^2 = V_0(1 - ||\mathbf{R}||^2)$, $\mathbf{T} = [V_0(1 - ||\mathbf{R}||^2)]^{-1/2}\mathbf{X}_1$, and $\mathbf{R} = V_0^{-1/2}\mathbf{X}_2$. Therefore, $(dv) = (dv_0)$ and

$$(dv) \wedge (d\mathbf{X}_1) \wedge (d\mathbf{X}_2) = v_0^{(n_1+n_2)m/2}(1 - ||\mathbf{R}||^2)^{n_1 m/2}(dv_0) \wedge (d\mathbf{T}) \wedge (d\mathbf{R}),$$

thus the desired result is obtained. The density (15) is achieved by integrating (14) over $V > 0$ via (3). □

Corollary 1. *Under the hypotheses of Theorem 8, define* $\mathbf{F} = \mathbf{T}'\mathbf{T} > \mathbf{0}$ *and* $\mathbf{B} = \mathbf{R}'\mathbf{R} > \mathbf{0}$. *Then the joint density function* $dF_{v,\mathbf{F},\mathbf{B}}(V,\mathbf{F},\mathbf{B})$ *can be written as*

$$\frac{\pi^{Nm/2} v^{Nm/2-1}|\mathbf{F}|^{(n_1-m-1)/2}|\mathbf{B}|^{(n_2-m-1)/2}}{\Gamma_1[n_0 m/2]\Gamma_m[n_1/2]\Gamma_m[n_2 m/2]} h\left\{v\left[1 + (1 - \operatorname{tr}\mathbf{B})\operatorname{tr}\mathbf{F}\right]\right\}(1 - \operatorname{tr}\mathbf{B})^{(n_0+n_1)m/2-1}(dv) \wedge d\mathbf{F}) \wedge d\mathbf{B}). \tag{16}$$

This distribution shall be termed threematrix variate generalised gamma - beta type II - beta type I distribution. *And the corresponding marginal density function* $dF_{\mathbf{F},\mathbf{B}}(\mathbf{F},\mathbf{B})$ *is given by*

$$\frac{\Gamma_1[Nm/2]|\mathbf{F}|^{(n_1-m-1)/2}|\mathbf{B}|^{(n_2-m-1)/2}}{\Gamma_1[n_0 m/2]\Gamma_m[n_1/2]\Gamma_m[n_2 m/2]}\left[1 + (1 - \operatorname{tr}\mathbf{B})\operatorname{tr}\mathbf{F}\right]^{-Nm/2}(1 - \operatorname{tr}\mathbf{B})^{(n_0+n_1)m/2-1}(d\mathbf{F}) \wedge d\mathbf{B}), \tag{17}$$

a distribution that shall be called bimatrix variate beta type II - beta type I distribution.

Proof. Results (16) and (17) follow from (14) and (15), respectively, by substitutions $\mathbf{F} = \mathbf{T}'\mathbf{T}$ and $\mathbf{B} = \mathbf{R}'\mathbf{R}$ and the use of (Muirhead, 2005, Theorem 2.1.14, p. 66). Thus,

$$(d\mathbf{T}) \wedge d\mathbf{R}) = 2^{-2m}|\mathbf{F}|^{(n_1-m-1)/2}|\mathbf{B}|^{(n_2-m-1)/2}(d\mathbf{F}) \wedge d\mathbf{B}) \bigwedge_{i=1}^{2}\left(\mathbf{H}_{1_i}' d\mathbf{H}_{1_i}\right).$$

Then, the demonstration is complete after a repeated implementation of (1). □

Note that different families of distributions can be obtained by replicating the procedure of Theorem 8. For example, if the change of variable on $\mathbf{X}_2$ is not carried out the corresponding distribution obtained can be termed *threematrix* variate generalized gamma-elliptical-Pearson type II distribution, and so on. Moreover, if $n_2 = 1$ is set in Theorem 8, the corresponding distribution can be termed threematrix variate generalised gamma-Pearson type VII-vector Pearson type II distribution, a fact that shows the versatility of the method proposed in this paper. In this case, a simultaneous model can involve a random variable V, a random matrix $\mathbf{T}$ and a random vector $\mathbf{r} = V^{-1/2}\mathbf{x}_2$.

In addition, when the multimatrix variate distributions involve non singular matrix arguments, we can find their inverse distribution (or if their arguments are rectangular or singular matrices, we can propose their corresponding generalise inverse distributions). For the distribution defined in Corollary 1, there are different possible multimatrix variate inverse distributions according to the selected argument combination of 1, 2 or 3 matrices. The following case exhibits the inverse when the three arguments are considered.

Theorem 9. *Assume the hypotheses of the Corollary 9 and define $W = V^{-1}$, $\mathbf{A} = \mathbf{F}^{-1}$ and $\mathbf{U} = \mathbf{B}^{-1}$. Then the joint density $dF_{W,\mathbf{A},\mathbf{U}}(w,\mathbf{A},\mathbf{U})$ can be written as*

$$\frac{\pi^{Nm/2}v^{Nm/2-1}|\mathbf{A}|^{-n_1/2}|\mathbf{U}|^{-n_2/2}}{\Gamma_1[n_0m/2]\Gamma_m[n_1/2]\Gamma_m[n_2m/2]}h\left\{w^{-1}\left[1+\left(1-\operatorname{tr}\mathbf{U}^{-1}\right)\operatorname{tr}\mathbf{A}^{-1}\right]\right\}\left(1-\operatorname{tr}\mathbf{U}^{-1}\right)^{(n_0+n_1)m/2-1}(dw)\wedge d\mathbf{A})\wedge d\mathbf{U}). \quad (18)$$

This distribution shall be called threematrix variate inverse generalised gamma-inverse beta type II - inverse beta type I distribution. *And the marginal density function $dF_{\mathbf{A},\mathbf{U}}(\mathbf{A},\mathbf{U})$ is given by*

$$\frac{\Gamma_1[Nm/2]|\mathbf{A}|^{-n_1/2}|\mathbf{U}|^{-n_2/2}}{\Gamma_1[n_0m/2]\Gamma_m[n_1/2]\Gamma_m[n_2m/2]}\left[1+\left(1-\operatorname{tr}\mathbf{U}^{-1}\right)\operatorname{tr}\mathbf{A}^{-1}\right]^{-Nm/2}\left(1-\operatorname{tr}\mathbf{U}^{-1}\right)^{(n_0+n_1)m/2-1}(d\mathbf{A})\wedge d\mathbf{U}). \quad (19)$$

A distribution that shall be termed bimatrix variate inverse beta type II - inverse beta type I distribution.

Proof. The density functions (18) and (19) are achieved from (16) and (17) by defining $W = V^{-1}$, $\mathbf{A} = \mathbf{F}^{-1}$ and $\mathbf{U} = \mathbf{B}^{-1}$ and by applying (Muirhead, 2005, Theorem 2.1.8, p. 59). □

Observe that the parameter domain of the multimatrix variate distributions can be extended to the complex or real fields. However their geometrical and/or statistical explication perhaps can be lost. These distributions are valid if we replace $n_i/2$ by a_i, $n_0m/2$ by a_0 and $Nm/2$ by a. Where the a'^s are complex numbers with positive real part. From a practical point of view for parametric estimation, this domain extension allows the use of nonlinear optimisation rather integer nonlinear optimisation, among other possibilities.

Moreover, note that each particular distribution is indexed by a kernel $h(\cdot)$ which defines itself another family of distributions. This opens a wide set of possible distributions available for an expert according to a preferred knowledge of a plausible $h(\cdot)$.

Finally, each distribution can be re-parametrized in order to reach no isotropy or general settings of its density function. In analogy with the normal case, the expressions obtained here appear in their standard form. However, perform the substitution $\mathbf{T}_i = \mathbf{A}_i^{-1}(\mathbf{S}_i - \boldsymbol{\mu}_i)\mathbf{B}_i^{-1}/\sqrt{r_i}$ in (7); where $\mathbf{A}_i \in \Re^{n_i\times n_i}$, $\mathbf{B}_i \in \Re^{m\times m}$ are non-singular matrices such that $\boldsymbol{\Sigma}_i = \mathbf{A}_i'\mathbf{A}_i > \mathbf{0}$ and $\boldsymbol{\Theta}_i = \mathbf{B}_i\mathbf{B}_i' > \mathbf{0}$, $\boldsymbol{\mu}_i \in \Re^{n_i\times m}$ and $r_i > 0$, $i = 1,\ldots,k$. Then by (Muirhead, 2005, Theorem 2.1.5, p. 58), we obtain

$$(d\mathbf{T}_i) = r_i^{-n_im/2}|\boldsymbol{\Sigma}_i|^{-m/2}|\boldsymbol{\Theta}_i|^{-n_i/2}(d\mathbf{S}_i).$$

And finally, the density $dF_{\mathbf{S}_1,\ldots,\mathbf{S}_k}(\mathbf{S}_1,\ldots,\mathbf{S}_k)$ can be written in the robust form

$$\frac{\Gamma_1[Nm/2]}{\pi^{(N-n_0)m/2}\Gamma_1[n_0m/2]\left(\displaystyle\prod_{i=1}^{k}r_i^{n_im/2}|\boldsymbol{\Sigma}_i|^{m/2}|\boldsymbol{\Theta}_i|^{n_i/2}\right)}\left[1+\operatorname{tr}\sum_{i=1}^{k}(\mathbf{S}_i-\boldsymbol{\mu}_i)'\boldsymbol{\Sigma}_i^{-1}(\mathbf{S}_i-\boldsymbol{\mu}_i)\boldsymbol{\Theta}_i^{-1}\right]^{-Nm/2}\bigwedge_{i=1}^{k}(d\mathbf{S}_i). \quad (20)$$

Observe that the particular values $k=1$, and $Nm/2=q$, $(N=n_0+n_1)$, turns (20) into the matrix variate Pearson type VII distribution, see (Gupta and Varga, 1993, Subsection 2.7.3.2, p. 54). And setting $q=(n_1m+r)/2$ in the preceding density, then the matrix variate T-distribution with r degrees of freedom is achieved. Furthermore, taking $r=1$ in the last density, then the matrix variate Cauchy distribution is obtained.

5 Application of molecular docking in SARS-CoV-2

Statistical applications supported by real data are usually difficult because the computation and underlying mathematical context force simplifications. Restrictions often related with inference based on independent sample information. It is extremely rare in nature to find time (space) probabilistic independent experiments. The most challenging problems in several areas are dependent in time and-or space. Some of them, such world pandemic situations are crucial and precise of urgent solutions. When a virus as SARS-CoV-2 arrives, a drug design is reacquired in a record time in order to avoid millions of deaths worldwide. A simplification of the docking problem involves two main steps: identify a target in the virus-protein, usually called pocket, and obtain a molecule (termed ligand) which must be place into an adequate site of the virus, for inhibiting its deathly activity. The target usually contains thousand of atoms with hundreds of possible unknown pockets were the small ligand can be placed. Discover the pocket and place the ligand is a colossal task for a human, however, the use of artificial intelligence can provide suitable sites to be considered by the biologist. Recently, a non conventional artificial intelligence method was provided in Ramirez et al. (2022) by a suitable recent discipline termed shape theory. The docking problem requires the use of certain invariance as translation, rotation and scaling, in order to avoid unnecessary computations. Instead of using the noisy Euclidean space latent in the classical artificial intelligence, the search is performed in quotient spaces with the demand symmetries. In this case, shape theory was applied in molecular docking for an automatic localization of ligand binding pockets in large proteins. The docking phenomena is modeled in Ramirez et al. (2022) by a physics chemistry punctuation function based on Lennard-Jones potential type $6-12$ and $6-10$. It includes the expected time dependence where the matrix representation at each state of the pocket-ligand goes into more stable and optimal low energy. Then the probabilistic dependent sample records n matrices in time ordered by a decreasing Lennard-Jones potential until the convergence is reached and the docking is stable. For this example, we have used the crystal structure of COVID-19 main protease an its inhibitor N3 due to Jin et al. (2020) and Liu et al. (2020). The protein consists of 2387 atoms and the ligand is constituted by 21 atoms. Once (Ramirez et al., 2022, Th. Sec. 3.) is applied, a plausible pocket is reached and the ligand is rotated rigidly into the cavity, until the energy is stabilized. In this case, $k=56$ locations of the ligand are dependently obtained inside a new cavity of 241 atoms. Figure 1 shows the dependent sample for this example, they constitutes 21×3 matrices $\mathbf{T}_i$, $i=1,\dots,56$.

Hence the sample matrices $\mathbf{F}_i=\mathbf{T}_i'\mathbf{T}_i$ shall be the core for estimation of the parameters in the likelihood, corresponding to the joint distribution of $\mathbf{F}_i$ in equation (11).

In this example, we consider a parameter space extension, from $n_i/2, i=0,\dots,56$ to corresponding positive reals $a_i, i=0,\dots,56$. Then the optimisation problem is reduced to find the maximum likelihood estimators of a_0, and $a=a_1=\cdots=a_{56}$ in the following dependent sample joint distribution:

$$\frac{\Gamma_1[(a_0+ka)m]}{\Gamma_1[a_0m]\Gamma_m^k[a]}\prod_{i=1}^{k}|\mathbf{F}_i|^{a-\frac{m+1}{2}}\left(1+\sum_{i=1}^{k}\operatorname{tr}\mathbf{F}_i\right)^{-(a_0+ka)m}.$$

For computations we have used the free licensed software R. The optimisation routine is performed in the package Optimx. The following estimation is obtained consistently by several methods and different random seeds for the initial parameters:

$$a_0=0.34397, a=0.19735$$

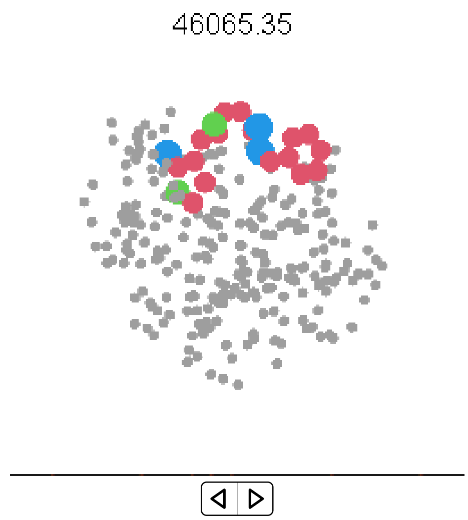

Figure 1: Dependent sample of 56 movements of the rigid ligand N3 with 21 atoms inside a new cavity of 241 atoms found on SARS-CoV-2 main protease. For illustration, the ligand is oversized respect the cavity (gray); their 21 atoms are colored as red (C), green (N), blue (O). The animation shows the progressive optimisation of the molecular docking according to a decreasing Lennard-Jones potential. First a positive energy indicates that the coupling is repulsive, however it is diminishing until negative values of strong affinity are reached and stabilized into an attractive complex ligand-pocket.

This estimation is also important because the invariance of the likelihood, under the complete family of spherical distribution, avoids the difficult a priori knowledge of the underlying model. Also, no fitting test is required. Note that the usual estimation based on an independent sample can not be applied here because the molecular docking is reached by the calibration of the Lennard-Jones potential which was set by the expert as a decreasing function of the dependent locus in the ligand.

In order to provide the consistency of the method in multiple scenarios, we compute an exhaustive optimization under a large set of models based on different samples. We have shown the estimation of the model for the 56 time-dependent molecular docking; we now consider the influence of each time by removing it from the experiment. In that case 56 models are estimated and the extremely very similar results can be seen in cluster 1 (black) of Figure 2 and Table 1.

If we define w as the number of extractions in the complete set of 56 dependent-time molecular docking, then after considering the provided estimations for $w = 0$ and $w = 1$, we now study all the plausible $\frac{56!}{(56-2)!2!} = 1540$ models for $w = 2$. Again the results behave as we expect, the maximal log likelihood decreases because the sample information is reduced, but 99% confidence intervals for $\hat{a}_0$ and $\hat{a}$ perform properly. According to Table 1, the resolution of the 1540 points is so narrow that cluster 2 (red) of Figure 2 seems to be a point. The cluster location sustains the preceding two estimations, as we expect. For $w = 3$ extraction, we have the large amount of $\frac{56!}{(56-3)!3!} = 27720$ possible combinations; Table 1 and reduced cluster 3 (green) of Figure 2 mantains the tendency. Finally, for $w = 4, 5, \ldots, 28$ the computations are colosal to be performed in the simple PC used for optimization, then we have studied all the ordered successive set of extractions from the following way. For $w = 4$, $(1:3,4), (1:3,5), \ldots, (1:3,56), (1:2,4,56), (1:2,5,56), \ldots, (1:2,55:56), (1,4,55:56), (1,5,55:56), \ldots, (1,54:56), (4,54:56), (5,54:56), \ldots, (53:56)$. For $w = 5$, $(1:4,5), (1:4,6), \ldots, (1:4,56), (1:3,5,56), (1:3,6,56), \ldots, (1:3,55,56), (1:2,5,55:56), (1:2,6,55:56), \ldots, (1:2,54:56), (1,5,54:56), (1,6,54:56), \ldots, (1,53:56), (5,53:56), (6,53:56), \ldots, (52:56)$, and so and on; the final set is given for $w = 28$, as $(1:27,28), (1:27,29), \ldots, (1:27,56), (1:26,28,56), (1:26,29,56), \ldots, (1:26,55:56), (1:25,28,55:56), (1:$

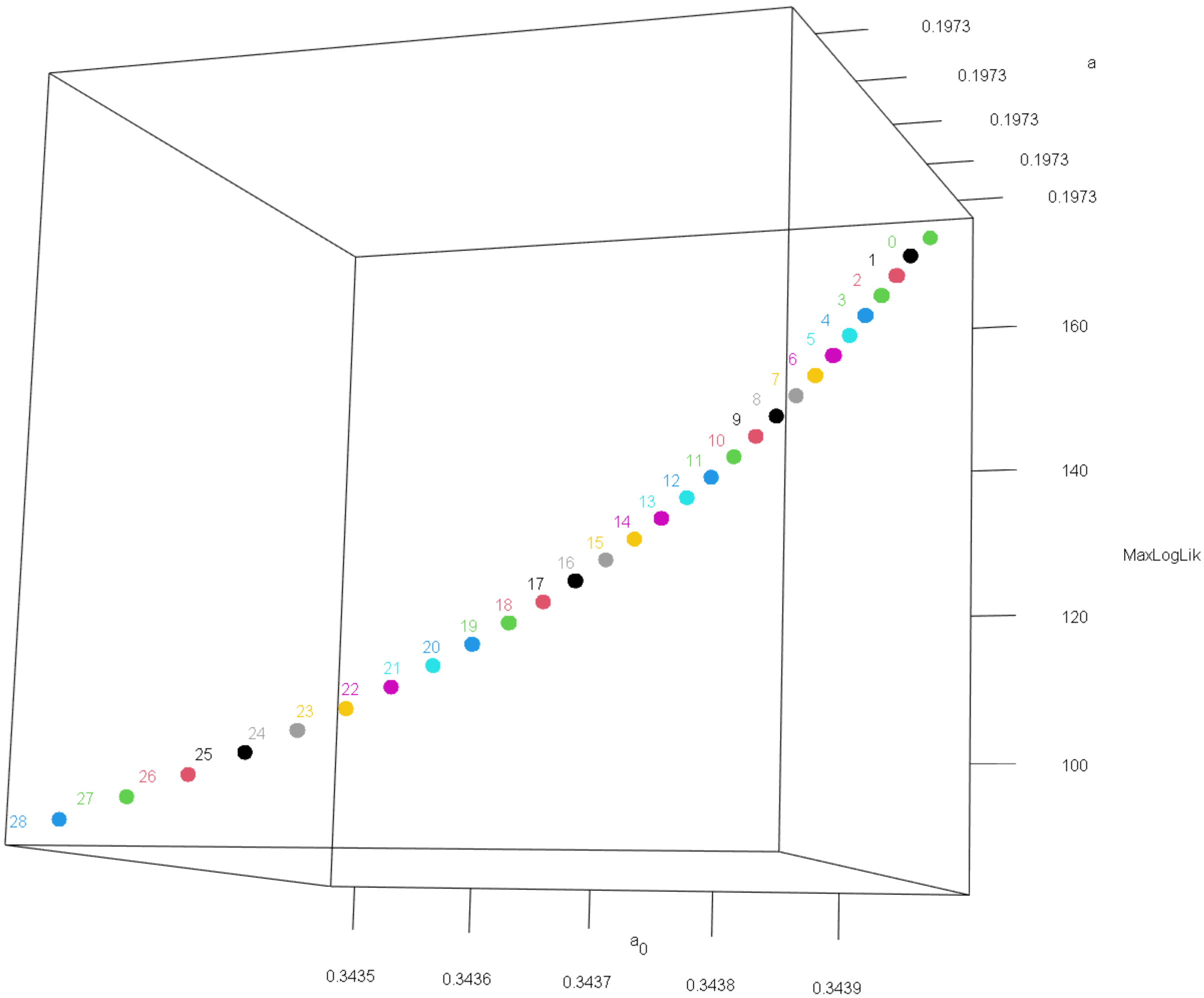

Figure 2: Evolution of maximal log likelihood information under several models with controlled substraction of dependent-time molecular docking. The models and estimations are described in Table 1. The full estimation for the complete set of 56 dependent-time is given in point 0 (black), with maximal log likelihood. The 56 possible models for a $w = 1$ subtraction are summarized in cluster 2 (red) which is over exceeded in size because the nitid resolution of the estimations. Clusters 2 (red) and 3 (green) represent the corresponding complete 1540 and 27720 models for extractions $w = 2$ and $w = 3$; again for displaying the estimation evolution the size is over exceeded, and they are confined in a very small space. The remaining extractions from $w = 4$ to $w = 28$ follows a particular but representative subset sequence given the prominent number of combination. They resulting expected monotone decreasing maximal log likelihood is provided and just verifies the gradual lost of meaningful sample information but the estimation of a_0 and a rounds the original estimation for the complete 56 dependent-time experiment.

Table 1: Exhaustive estimation for molecular docking under subtraction of w dependent-times samples. For each experiment from $w = 0$ to $w = 28$ extractions, the number M of studied models are shown with the corresponding 99% confidence intervals for estimations of "a_0", "a" and maximum log likelihood (MaxLogLik).

w	M	a_0	a	MaxLogLik
0	1	0.34396995340027808	0.19734727197157109	172.475
1	56	[0.3439581, 0.3439584]	[0.19734433, 0.19734435]	[169.3689293, 169.3689294]
2	1540	[0.343949, 0.343950]	[0.19734254, 0.19734257]	[166.2630596, 166.2630598]
3	27720	[0.343940, 0.343941]	[0.19734068, 0.19734070]	[163.1573540, 163.1573550]
4	206	[0.3439306, 0.3439312]	[0.19733875, 0.19733878]	[160.0518201, 160.0518202]
5	250	[0.3439209, 0.3439214]	[0.19733674, 0.19733676]	[156.9464627, 156.9464628]
6	291	[0.3439107, 0.3439115]	[0.19733466, 0.19733469]	[153.8412893, 153.8412895]
7	329	[0.3439001, 0.3439007]	[0.19733249, 0.19733251]	[150.7363072, 150.7363074]
8	364	[0.3438892, 0.3438894]	[0.19733023, 0.19733025]	[147.6315241, 147.6315243]
9	396	[0.3438777, 0.3438779]	[0.19732788, 0.19732790]	[144.5269480, 144.5269483]
10	425	[0.3438657, 0.3438662]	[0.19732543, 0.19732545]	[141.4225879, 141.4225882]
11	451	[0.3438532, 0.3438534]	[0.19732288, 0.19732289]	[138.3184529, 138.3184532]
12	474	[0.3438401, 0.3438403]	[0.19732021, 0.19732023]	[135.2145529, 135.2145531]
13	494	[0.3438264, 0.3438267]	[0.19731742, 0.19731743]	[132.1108983, 132.1108986]
14	511	[0.3438120, 0.3438125]	[0.19731450, 0.19731452]	[129.0075005, 129.0075008]
15	525	[0.3437971, 0.3437975]	[0.19731144, 0.19731146]	[125.9043716, 125.9043719]
16	536	[0.3437815, 0.3437817]	[0.19730824, 0.19730825]	[122.8015243, 122.8015246]
17	544	[0.3437648, 0.3437655]	[0.19730485, 0.19730488]	[119.6989727, 119.6989729]
18	549	[0.3437477, 0.3437479]	[0.19730132, 0.19730134]	[116.5967315, 116.5967318]
19	551	[0.3437292, 0.3437299]	[0.19729759, 0.19729764]	[113.4948170, 113.4948172]
20	550	[0.3437101, 0.3437107]	[0.19729367, 0.19729368]	[110.3932463, 110.3932466]
21	546	[0.3436900, 0.3436902]	[0.19728953, 0.19728954]	[107.2920384, 107.2920387]
22	539	[0.3436686, 0.3436688]	[0.19728514, 0.19728515]	[104.1912136, 104.1912139]
23	529	[0.3436459, 0.3436461]	[0.19728050, 0.19728051]	[101.0907940, 101.0907943]
24	516	[0.3436216, 0.3436223]	[0.19727558, 0.19727562]	[97.99080370, 97.99080400]
25	500	[0.3435963, 0.3435965]	[0.19727036, 0.19727037]	[94.89126900, 94.89126930]
26	481	[0.3435691, 0.3435693]	[0.19726479, 0.19726480]	[91.79221870, 91.79221900]
27	459	[0.3435401, 0.3435402]	[0.19725886, 0.19725887]	[88.69368450, 88.69368480]
28	434	[0.3435091, 0.3435093]	[0.19725252, 0.19725253]	[85.59570110, 85.59570150]

$25, 29, 55 : 56), \ldots, (1 : 25, 54 : 56), \ldots, (1, 28, 31 : 56), (1, 29, 31 : 56), (1, 30 : 56), (28, 30 : 56), (29 : 56)$. The corresponding confidence intervals and narrow clusters follow the expected evolution given in Figure 2 and Table 1.

Finally, in the preceding example we have shown that the joint distribution functions of Sections 3 and 4 are easily computable and applied to real data. They provide an expected performance under multiple sub-samples in the rigorous dependent-time molecular docking setting. This is an important opportunity for several similar phenomena with multimatrix dependent samples.

6 Conclusions

This work introduces the class of *multimatrix variate distributions* by embedding previously isolated bimatrix-variate models within the general framework of elliptically contoured families. Most of the proposed distributions are invariant under the class of spherical models, an important property that eliminates the need for prior specification of the underlying joint distribution. The resulting models are computationally tractable, and several new properties and combinations involving three matrices are derived. These results address the problem of matrix-dependent samples by providing flexible constructions of joint distributions with desirable marginal distributions under independence. The versatility of the proposed framework allows scalars, vectors, and matrices to be incorporated simultaneously within a single joint distribution. Moreover, the computational feasibility of these models opens the way for further applications, including the evaluation of multiple probabilities on cones. Finally, the paper presents an application to molecular docking, specifically analyzing a known inhibitor within a newly identified cavity of the SARS-CoV-2 main protease.

Acknowledgements

The authors wish to thank the Editor and the anonymous reviewers for their constructive comments on the preliminary version of this paper.

References

Bekker, A., Roux, J. J. J., Ehlers, E. and Arashi, M. (2001). Bimatrix variate beta type IV distribution: relation to Wilks's statistics and bimatrix variate Kummer-beta type IV distribution. *Communication in Statistics-Theory and Methods*, **40**, 4165–4178.

Chen, J. J. and Novick, M. R. (1984). Bayesian analysis for binomial models with generalized beta prior distributions. *Journal of Educational Statistics*, **9**, 163–175.

Díaz-García, J. A. and Gutiérrez-Jáimez, R. (2006). The distribution of the residual from general elliptical multivariate linear model. *Journal of Multivariate Analysis*, **97**, 1829–1841.

Díaz-García, J. A. and Gutiérrez-Jáimez, R. (2010a). Bimatrix variate generalised beta distributions. *South African Statistical Journal*, **44**, 193–208.

Díaz-García, J. A. and Gutiérrez-Jáimez, R. (2010b). Complex bimatrix variate generalised beta distributions. *Linear Algebra and its Applications*, **432**, 571–582.

Díaz-García, J. A. and Gutiérrez-Jáimez, R. (2011). Noncentral bimatrix variate generalised beta distributions. *Metrika*, **73**, 317–333.

Díaz-García, J. A. and Caro-Lopera, F. J. (2022). Multimatricvariate distribution under elliptical models. *Journal of Statistical Planning and Inference*, **216**, 109-117.

Díaz-García, J. A., Caro-Lopera, F. J. and Pérez Ramírez, F. O. (2022). Multivector variate distributions: an application in Finance. *Sankhyā*, **84**, 534–555.

Dickey, J. M. (1967). Matricvariate generalizations of the multivariate t- distribution and the inverted multivariate t-distribution. *Annals of Mathematical Statistics*, **38**, 511–518.

Ehlers, R. (2011). *Bimatrix Variate Distributions of Wishart Ratios with Application.* Doctoral dissertation, Faculty of Natural & Agricultural Sciences University of Pretoria, Pretoria. http://hdl.handle.net/2263/31284.

Fang, K. T. and Zhang, Y. T. (1990). *Generalized Multivariate Analysis.* Science Press, Springer-Verlag, Beijing.

Fang, K. T., Zhang, Y. T. and Ng, K. W. (1990). *Symmetric Multivariate and realted distributions.* Springer-Science+Business Media, B. V., New Delhi.

Goodall, K. I. and Mardia, K. V. (1993). Multivariate aspects of shape theory. *Annals of Statistics*, **21**, 848–866.

Gupta, A. K. and Varga, T. (1993). *Elliptically Contoured Models in Statistics.* Kluwer Academic Publishers, Dordrecht.

Jin, Z., Du, X., Xu, Y. et al. (2020). Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature*, **582**, 289–293.

Libby, D. L. and Novick, M. R. (1982). Multivariate Generalized beta distributions with applications to utility assessment. *Journal of Educational Statistics*, **7**, 271–294.

Liu, X., Zhang, B., Jin, Z., Yang, H. and Rao, Z. (2020). The crystal structure of COVID-19 main protease in complex with an inhibitor N3. *Protein DataBank*, **10**, no. 10.2210.

Muirhead, R. J. (2005). *Aspects of Multivariate Statistical Theory.* John Wiley & Sons, New York.

Nadarajah, S. (2007). A bivariate gamma model for drought. *Water Resources Research*, **43**, W08501.

Nadarajah, S. (2013). A bivariate distribution with gamma and beta marginals with application to drought data. *Journal of Applied Statistics*, **36**, 277–301.

Olkin, I. and Liu, R. (2003). A bivariate beta distribution. *Statistics and Probability Letters*, **62**, 407–412.

Olkin, I. and Rubin, H. (1964). Multivariate beta distributions and independence properties of Wishart distribution. *Annals of Mathematical Statistics*, **35**, 261–269. Correction 1966, 37(1), 297.

Ramírez-Velásquez, I., Bedoya-Calle, Á. H., Vélez, E., and Caro-Lopera, F. J. (2022). Shape theory applied to molecular docking and automatic localization of ligand binding pockets in large proteins. *ACS Omega*, **7**, 45991–46002.

Sarabia, J. M., Prieto, F. and Jordá, V. (2014). Bivariate beta-generated distributions with application to well-being data. *Journal of Statistical Distributions and Applications*, **1**, 15.