Generalized Multivariate Birnbaum-Saunders Distributions And Related Inferential Issues

Debasis Kundu[†] & N. Balakrishnan[‡] & Ahad Jamalizadeh[§]

Abstract

Birnbaum and Saunders introduced in 1969 a two-parameter lifetime distribution which has been used quite successfully to model a wide variety of univariate positively skewed data. Diaz-Garcia and Leiva-Sanchez [9] proposed a generalized Birnbaum-Saunders distribution by using an elliptically symmetric distribution in place of the normal distribution. Recently, Kundu et al. [17] introduced a bivariate Birnbaum-Saunders distribution, based on a transformation of a bivariate normal distribution, and discussed its properties and associated inferential issues. In this paper, we construct a generalized multivariate Birnbaum-Saunders distribution, by using the multivariate elliptically symmetric distribution as a base kernel for the transformation instead of the multivariate normal distribution. Different properties of this distribution are obtained in the general case. Special emphasis is placed on statistical inference for two particular cases: (i) multivariate normal kernel and (ii) multivariate-t kernels. We use the maximized log-likelihood values for selecting the best kernel function. Finally, a data analysis is presented for illustrative purposes.

Keywords: Birnbaum-Saunders Distribution; Generalized Birnbaum-Saunders Distribution; Maximum Likelihood Estimators; Fisher Information Matrix; Asymptotic Distribution; Monte Carlo Simulation; Multivariate Normal Distribution; Elliptically Symmetric Distribution; Akaike Information Criterion.

[†]Dept. of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India, e-mail: kundu@iitk.ac.in. Corresponding author.

[‡]Dept. of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada L8S 4K1.

[§]Dept. of Statistics, Faculty of Mathematics & Computer, Shahid Bahonar University of Kerman, Kerman, Iran 76169-14111.

1 Introduction

The univariate Birnbaum-Saunders (BS) distribution was originally introduced by Birnbaum and Saunders [5, 6] as a failure time distribution for fatigue failure of a unit caused under cyclic loading. The BS distribution can be defined in terms of a monotone transformation of the normal distribution. The cumulative distribution function (CDF) of a two-parameter BS random variable T is of the form

$$F_T(t;\alpha,\beta) = \Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t}\right)^{\frac{1}{2}}\right\}\right], \qquad t>0, \quad \alpha>0,\ \beta>0, \tag{1}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. The corresponding probability density function (PDF) is

$$f_T(t;\alpha,\beta) = \frac{1}{2\sqrt{2\pi}\alpha\beta}\left[\left(\frac{\beta}{t}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t}\right)^{\frac{3}{2}}\right] exp\left[-\frac{1}{2\alpha^2}\left(\frac{t}{\beta}+\frac{\beta}{t}-2\right)\right], \qquad t>0. \tag{2}$$

Hereafter, we will denote this distribution by BS(α,β). The parameters α and β are the shape and scale parameters, respectively. Moreover, β is the median of the BS distribution. For various developments on the BS distribution, one may refer to Xiao et al. [27], Vilca et al. [26], and the references cited therein.

Diaz-Garcia and Leiva-Sanchez [9] introduced a generalized Birnbaum-Saunders (GBS) distribution by replacing the standard normal distribution function $\Phi(\cdot)$ in (1) by an elliptically symmetric distribution. Recall that a random variable X follows an elliptically symmetric distribution if it has the probability density function (PDF)

$$f_{EC}(x;\mu,\sigma^2,h^{(1)}) = \frac{c}{\sigma}h^{(1)}\left[\frac{(x-\mu)^2}{\sigma^2}\right], \qquad x\in\mathbb{R}, \tag{3}$$

where $h^{(1)}(u) > 0$, with $u > 0$, is a real-valued function. It corresponds to the kernel of the PDF of X, and c is the normalizing constant. From now on, we will denote this by $EC(\mu, \sigma^2, h^{(1)})$. Thus, a random variable T is said to have the GBS distribution, with parameters α, β and kernel $h^{(1)}$, if its CDF is

$$F_T(t; \alpha, \beta, h^{(1)}) = P(T \leq t) = F_{EC}\left[\frac{1}{\alpha}\left\{\left(\frac{t}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t}\right)^{\frac{1}{2}}\right\}; h^{(1)}\right], \tag{4}$$

where $F_{EC}(\cdot; h^{(1)})$ denotes the CDF of $EC(0, 1, h^{(1)})$. These authors then discussed different properties of the GBS, and noted that the GBS can be more flexible than the BS distribution as it has a broader range for coefficients of skewness and kurtosis. It will, therefore, be more useful for data analytic purposes; see, for example, Leiva et al. [19].

Recently, Kundu et al. [17] introduced a five-parameter bivariate Birnbaum-Saunders distribution by using the same monotone transformation on two variables possessing jointly a bivariate normal distribution. The cumulative distribution function of the bivariate BS random vector $(T_1, T_2)^T$ is of the form

$$P(T_1 \leq t_1, T_2 \leq t_2) = \Phi_2\left[\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\beta_1}} - \sqrt{\frac{\beta_1}{t_1}}\right), \frac{1}{\alpha_2}\left(\sqrt{\frac{t_2}{\beta_2}} - \sqrt{\frac{\beta_2}{t_2}}\right); \rho\right] \tag{5}$$

for $t_1 > 0,\ t_2 > 0$, where $\alpha_1 > 0$, $\beta_1 > 0$, $\alpha_2 > 0$, $\beta_2 > 0$, $-1 < \rho < 1$, and $\Phi_2(u, v; \rho)$ is the joint cumulative distribution function of a standard bivariate normal vector $(Z_1, Z_2)^T$ with correlation coefficient ρ. When $(T_1, T_2)^T$ has the bivariate BS distribution as in (5), then evidently T_1 and T_2 have univariate BS distributions as marginals. These authors then discussed several properties of this bivariate BS distribution and associated inferential issues.

The main purpose of this paper is to introduce a generalized multivariate Birnbaum-Saunders (GMBS) distribution by using the multivariate elliptically symmetric distribution in place of the multivariate normal distribution. We then discuss different properties of the proposed GMBS distribution in the general case. For statistical inference, we consider two special cases in detail, namely, when the kernel function is (i) a multivariate normal distribution and (ii) a multivariate t-distribution with specified degrees of freedom. The problem of choosing a proper kernel is an important problem in data analysis. Here, we use the maximized log-likelihood values for choosing the best kernel function. Finally, we present a numerical example for the purpose of illustrating the model as well as the inferential methods developed here, in which it is shown that the GMBS distribution with multivariate t kernel fits better than the model based on the multivariate normal kernel.

The rest of this paper is organized as follows. In Section 2, we present all pertinent preliminary details. In Section 3, we introduce the GMBS distribution and discuss different properties. Maximum likelihood estimates of the model parameters and associated inference for two specific kernels of multivariate normal and multivariate t are described in Section 4. A numerical example is presented in Section 5 for illustrative purposes, and finally some concluding remarks are made in Section 6.

2 Preliminaries

2.1 Elliptically Symmetric Distribution

Definition 1: A p-dimensional random vector $\boldsymbol{X}$ is said to have an elliptically symmetric distribution with p-dimensional location vector $\boldsymbol{\mu}$, a $p \times p$ positive definite dispersion matrix $\boldsymbol{\Sigma}$, and the density generator $h^{(p)}(\cdot)$, if the PDF of $\boldsymbol{X}$ is of the form (see Fang et al. [10] and Anderson and Fang [2])

$$f_{EC_p}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, h^{(p)}) = c_p |\boldsymbol{\Sigma}|^{-\frac{1}{2}} h^{(p)}(w(\boldsymbol{x})), \quad \boldsymbol{x} \in \mathbb{R}^p, \tag{6}$$

where $w(\boldsymbol{x}) : \mathbb{R}^p \to \mathbb{R}_+$ with $w(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})$, $h^{(p)} : \mathbb{R}_+ \to \mathbb{R}_+$, $c_p > 0$, and

$$\int_{\mathbb{R}^p} f_{EC_p}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, h^{(p)}) \boldsymbol{dx} = 1.$$

In what follows, we shall denote $f_{EC_p}(\boldsymbol{x};\mathbf{0},\boldsymbol{\Sigma},h^{(p)})$ simply by $f_{EC_p}(\boldsymbol{x};\boldsymbol{\Sigma},h^{(p)})$ and the corresponding random vector by $\mathrm{EC}_p(\boldsymbol{\Sigma},h^{(p)})$. It should be mentioned here that a more general definition of an elliptically symmetric distribution in terms of the characteristic function also exists; see, for example, Cambanis et al. [7], but we shall restrict our attention to the one based on (6). It is well known that this family of distributions is closed under linear transformation, marginalization and conditioning. In particular, if $\boldsymbol{X} \sim \mathrm{EC}_p(\boldsymbol{\mu},\boldsymbol{\Sigma},h^{(p)})$, then $\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{X}-\boldsymbol{\mu}) \sim \mathrm{EC}_p(\mathbf{0},\boldsymbol{I}_p,h^{(p)})$, where $\boldsymbol{I}_p$ is the $p\times p$ identity matrix.

We now present different examples of the elliptically symmetric distribution, which are obtained with different choices of $h^{(p)}(\cdot)$.

Example 1: Multivariate (p-variate) Normal Distribution

$$h^{(p)}(x) = e^{-x/2} \quad \text{and} \quad c_p = (2\pi)^{-p/2}. \tag{7}$$

This is undoubtedly the most popular elliptically symmetric distribution.

Example 2: Symmetric Kotz type distribution

$$h^{(p)}(x) = x^{\beta-1}e^{-\lambda x^{\delta}} \quad \text{and} \quad c_p = \frac{\delta\Gamma(p/2)}{\pi^{p/2}\Gamma((2\beta+p-2)/2\delta)}\lambda^{(2\beta+p-2)/2\delta}, \tag{8}$$

where $\delta>0$, $\lambda>0$, $2\beta+p>2$. When $\beta=\delta=1$ and $\lambda=1/2$, it reduces to the p-variate normal distribution. When $\beta=1$ and $\lambda=1/2$, it is known as the multivariate power normal distribution.

Example 3: Multivariate t-distribution (with $\nu>0$ degrees of freedom)

$$h^{(p)}(x) = \left(1+\frac{x}{\nu}\right)^{-(\nu+p)/2} \quad \text{and} \quad c_p = \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)(\nu\pi)^{p/2}}. \tag{9}$$

Example 4: Symmetric multivariate Pearson Type VII distribution

$$h^{(p)}(x) = \left(1+\frac{x}{\theta}\right)^{-\xi} \quad \text{and} \quad c_p = \frac{\Gamma(\xi)}{\Gamma(\xi-p/2)(\theta\pi)^{p/2}}, \tag{10}$$

where $\xi > p/2$ and $\theta > 0$. When $\xi = (\nu + p)/2$ and $\theta = \nu$, it becomes the multivariate t-distribution. For $\theta = 1$ and $\xi = (p+1)/2$, it becomes the multivariate Cauchy distribution.

We need the following notation for further developments. The p-dimensional vectors $\boldsymbol{X}$ and $\boldsymbol{\mu}$ and the $p \times p$ matrix $\boldsymbol{\Sigma}$ are partitioned as follows:

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}, \tag{11}$$

where $\boldsymbol{X}_1$ and $\boldsymbol{\mu}_1$ are $q \times 1$ vectors and $\boldsymbol{\Sigma}_{11}$ is a $q \times q$ matrix, and all the remaining partitioned elements are defined so that the corresponding orders match. Then, we have the following lemma, proof of which can be found in Fang et al. [10], for example.

Lemma 1: If $\boldsymbol{X} \sim \mathrm{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, h^{(p)})$, then:

(a) $\boldsymbol{X}_1 \sim \mathrm{EC}_q(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}, h^{(q)})$ and $\boldsymbol{X}_2 \sim \mathrm{EC}_{p-q}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22}, h^{(p-q)})$;

(b) $\boldsymbol{X}_1 \mid (\boldsymbol{X}_2 = \boldsymbol{x}_2) \sim \mathrm{EC}_q(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}, h^{(q)}_{v(\boldsymbol{x}_2)})$,

where $v(\boldsymbol{x}_2) = (\boldsymbol{x}_2 - \boldsymbol{\mu}_2)^T\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)$, and $h^{(q)}$ and $h^{(q)}_a$ can be expressed in terms of $h^{(p)}$, as

$$h^{(q)}(u) = \frac{\pi^{(p-q)/2}}{\Gamma((p-q)/2)} \int_0^\infty x^{\frac{p-q}{2}-1} h^{(p)}(u+x)dx \quad \text{and} \quad h^{(q)}_a(u) = \frac{h^{(p)}(u+a)}{h^{(p-q)}(a)}.$$

3 Generalized Multivariate BS Distribution

3.1 Definition

Definition 2: Let $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^p$, where $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_p)^T$ and $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_p)^T$, with $\alpha_i > 0, \beta_i > 0$ for $i = 1, \cdots, p$. Let $\boldsymbol{\Gamma}$ be a $p \times p$ positive-definite correlation matrix. Then, the random vector $\boldsymbol{T} = (T_1, \cdots, T_p)^T$ is said to have a generalized multivariate Birnbaum-Saunders distribution with parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma})$ and the density generator $h^{(p)}$, denoted by

$\boldsymbol{T} \sim \text{GBS}_p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}, h^{(p)})$, if the CDF of $\boldsymbol{T}$, i.e., $P(\boldsymbol{T} \leq \boldsymbol{t}) = P(T_1 \leq t_1, \cdots, T_p \leq t_p)$, is given by

$$P(\boldsymbol{T} \leq \boldsymbol{t}) = F_{EC_p}\left[\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\beta_1}} - \sqrt{\frac{\beta_1}{t_1}}\right), \cdots, \frac{1}{\alpha_p}\left(\sqrt{\frac{t_p}{\beta_p}} - \sqrt{\frac{\beta_p}{t_p}}\right); \boldsymbol{\Gamma}, h^{(p)}\right] \tag{12}$$

for $\boldsymbol{t} > \boldsymbol{0}$, where $F_{EC_p}(\cdot; \boldsymbol{\Gamma}, h^{(p)})$ denotes the CDF of $EC_p(\boldsymbol{\Gamma}, h^{(p)})$. The corresponding joint PDF of $\boldsymbol{T} = (T_1, \cdots, T_p)^T$ is, for $\boldsymbol{t} > \boldsymbol{0}$,

$$\begin{aligned} f_{\boldsymbol{T}}(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}) &= f_{EC_p}\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\beta_1}} - \sqrt{\frac{\beta_1}{t_1}}\right), \cdots, \frac{1}{\alpha_p}\left(\sqrt{\frac{t_p}{\beta_p}} - \sqrt{\frac{\beta_p}{t_p}}\right); \boldsymbol{\Gamma}, h^{(p)}\right) \\ &\quad \times \prod_{i=1}^{p} \frac{1}{2\alpha_i\beta_i}\left\{\left(\frac{\beta_i}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta_i}{t_i}\right)^{\frac{3}{2}}\right\}, \end{aligned} \tag{13}$$

where $f_{EC_p}(\cdot)$ is as given in (6).

3.2 Marginal and Conditional Distributions

The following theorem provides the marginal and conditional distributions of $\text{GBS}_p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}, h^{(p)})$.

Theorem 1: Let $\boldsymbol{T} \sim \text{GBS}_p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}, h^{(p)})$, and further $\boldsymbol{T}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}$ be partitioned as follows:

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{T}_1 \\ \boldsymbol{T}_2 \end{pmatrix}, \quad \boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix}, \quad \boldsymbol{\Gamma} = \begin{pmatrix} \boldsymbol{\Gamma}_{11} & \boldsymbol{\Gamma}_{12} \\ \boldsymbol{\Gamma}_{21} & \boldsymbol{\Gamma}_{22} \end{pmatrix}, \tag{14}$$

where $\boldsymbol{T}_1, \boldsymbol{\alpha}_1, \boldsymbol{\beta}_1$ are all $q \times 1$ vectors and $\boldsymbol{\Gamma}_{11}$ is a $q \times q$ matrix, with the remaining elements all defined suitably so that the corresponding orders match. Then, we have:

(a) $\boldsymbol{T}_1 \sim \text{GBS}_q(\boldsymbol{\alpha}_1, \boldsymbol{\beta}_1, \boldsymbol{\Gamma}_{11}, h^{(q)})$ and $\boldsymbol{T}_2 \sim \text{GBS}_{p-q}(\boldsymbol{\alpha}_2, \boldsymbol{\beta}_2, \boldsymbol{\Gamma}_{22}, h^{(p-q)})$, where $h^{(q)}$ and $h^{(p-q)}$ can be obtained in terms of $h^{(p)}$;

(b) The conditional CDF of $\boldsymbol{T}_1$, given $\boldsymbol{T}_2 = \boldsymbol{t}_2$, is

$$P[\boldsymbol{T}_1 \leq \boldsymbol{t}_1 | \boldsymbol{T}_2 = \boldsymbol{t}_2] = F_{EC_q}(\boldsymbol{w}; \boldsymbol{\Gamma}_{11.2}, h^{(q)}_{a(\boldsymbol{v}_2)}),$$

where

$$\boldsymbol{w} = \boldsymbol{v}_1 - \boldsymbol{\Gamma}_{12}\boldsymbol{\Gamma}_{22}^{-1}\boldsymbol{v}_2, \qquad \boldsymbol{v} = (v_1 \cdots, v_p)^T, \qquad v_i = \frac{1}{\alpha_i}\left(\sqrt{\frac{t_i}{\beta_i}} - \sqrt{\frac{\beta_i}{t_i}}\right) \text{ for } i = 1, \cdots, p,$$

$$\boldsymbol{\Gamma}_{11.2} = \boldsymbol{\Gamma}_{11} - \boldsymbol{\Gamma}_{12}\boldsymbol{\Gamma}_{22}^{-1}\boldsymbol{\Gamma}_{21}, \quad \boldsymbol{v} = \begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \end{pmatrix}, \text{ and } a(\boldsymbol{v}_2) = \boldsymbol{v}_2^T\boldsymbol{\Gamma}_{22}^{-1}\boldsymbol{v}_2,$$

where $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are vectors of dimensions q and $(p-q)$, respectively;

(c) The conditional PDF of $\boldsymbol{T}_1$, given $\boldsymbol{T}_2 = \boldsymbol{t}_2$, is

$$f_{\boldsymbol{T}_1|(\boldsymbol{T}_2=\boldsymbol{t}_2)}(\boldsymbol{t}_1) = f_{EC_q}(\boldsymbol{w}; \boldsymbol{\Gamma}_{12}, h^{(q)}_{a(\boldsymbol{v}_2)}) \prod_{i=1}^{q} \frac{1}{2\alpha_i\beta_i}\left\{\left(\frac{\beta_i}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta_i}{t_i}\right)^{\frac{3}{2}}\right\}. \tag{15}$$

Proof: Part (a) follows readily from (12) by letting $t_{q+1} \to \infty, \cdots, t_p \to \infty$.

For the proof of Part (b), if $V_i = \frac{1}{\alpha_i}\left(\sqrt{\frac{T_i}{\beta_i}} - \sqrt{\frac{\beta_i}{T_i}}\right)$, then we have $\boldsymbol{V} = \begin{pmatrix} \boldsymbol{V}_1 \\ \boldsymbol{V}_2 \end{pmatrix} \sim EC_p\left(\boldsymbol{0}, \begin{bmatrix} \boldsymbol{\Gamma}_{11} & \boldsymbol{\Gamma}_{12} \\ \boldsymbol{\Gamma}_{21} & \boldsymbol{\Gamma}_{22} \end{bmatrix}, h^{(p)}\right)$, where $\boldsymbol{V}_1 = (V_1, \cdots, V_q)^T$ and $\boldsymbol{V}_2 = (V_{q+1}, \cdots, V_p)^T$ are $q \times 1$ and $(p-q) \times 1$ vectors, respectively, and $\boldsymbol{\Gamma}_{ij}$ (for $i, j = 1, 2$) are as defined before. So,

$$\Pr\left[\boldsymbol{T}_1 \leq \boldsymbol{t}_1 \,|\boldsymbol{T}_2 = \boldsymbol{t}_2\right] = \Pr\left[\boldsymbol{V}_1 \leq \boldsymbol{v}_1 \,|\boldsymbol{V}_2 = \boldsymbol{v}_2\right] = F_{EC_q}\left(\boldsymbol{w}; \boldsymbol{\Gamma}_{11.2}, h^{(q)}_{a(\boldsymbol{v}_2)}\right).$$

Part (c) follows immediately from Lemma 1. ∎

3.3 Distributions of Reciprocals

In this section, we establish some properties of T_i^{-1}. In addition to the notation used in Theorem 1, we further denote by $\frac{1}{\boldsymbol{a}} = \left(\frac{1}{a_1}, \cdots, \frac{1}{a_k}\right)^T$ for a vector $\boldsymbol{a} = (a_1, \cdots, a_k)^T \in \mathcal{R}^{+^k}$.

Theorem 2: Let $\boldsymbol{T} \sim \text{GBS}_p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}, h^{(p)})$. Then:

(a)

$$\begin{pmatrix} \boldsymbol{T}_1 \\ \frac{1}{\boldsymbol{T}_2} \end{pmatrix} \sim \text{GBS}_p\left(\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\beta}_1 \\ \frac{1}{\boldsymbol{\beta}_2} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Gamma}_{11} & -\boldsymbol{\Gamma}_{12} \\ -\boldsymbol{\Gamma}_{21} & \boldsymbol{\Gamma}_{22} \end{pmatrix}, h^{(p)}\right);$$

(b)

$$\begin{pmatrix} \frac{1}{\boldsymbol{T}_1} \\ \boldsymbol{T}_2 \end{pmatrix} \sim \text{GBS}_p\left(\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix}, \begin{pmatrix} \frac{1}{\boldsymbol{\beta}_1} \\ \boldsymbol{\beta}_2 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Gamma}_{11} & -\boldsymbol{\Gamma}_{12} \\ -\boldsymbol{\Gamma}_{21} & \boldsymbol{\Gamma}_{22} \end{pmatrix}, h^{(p)}\right);$$

(c)

$$\begin{pmatrix} \frac{1}{\boldsymbol{T}_1} \\ \frac{1}{\boldsymbol{T}_2} \end{pmatrix} \sim \text{GBS}_p\left(\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix}, \begin{pmatrix} \frac{1}{\boldsymbol{\beta}_1} \\ \frac{1}{\boldsymbol{\beta}_2} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Gamma}_{11} & \boldsymbol{\Gamma}_{12} \\ \boldsymbol{\Gamma}_{21} & \boldsymbol{\Gamma}_{22} \end{pmatrix}, h^{(p)}\right).$$

Proof: (a) Let us denote

$$\widetilde{\boldsymbol{\Gamma}} = \begin{pmatrix} \boldsymbol{\Gamma}_{11} & -\boldsymbol{\Gamma}_{12} \\ -\boldsymbol{\Gamma}_{21} & \boldsymbol{\Gamma}_{22} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Gamma}^{-1} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix}.$$

Then, we have from Rao [24] that

$$|\boldsymbol{\Gamma}| = |\widetilde{\boldsymbol{\Gamma}}| \quad \text{and} \quad \widetilde{\boldsymbol{\Gamma}}^{-1} = \begin{pmatrix} \boldsymbol{A}_{11} & -\boldsymbol{A}_{12} \\ -\boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix}. \tag{16}$$

Now, let us define the random variables $S_{q+1}, \cdots, S_p$ as $S_i = T_i^{-1}$ $(i = q+1, \cdots, p)$. Then, the joint PDF of $(T_1, \cdots, T_q, S_{q+1}, \cdots, S_p)$ can be obtained from (13) by performing the necessary transformation. Since, from (16), we have

$$f_{EC_p}(u_1, \cdots, u_q, -u_{q+1}, \cdots, -u_p; \boldsymbol{\Gamma}) = f_{EC_p}(u_1, \cdots, u_q, u_{q+1}, \cdots, u_p; \widetilde{\boldsymbol{\Gamma}}),$$

the result in Part (a) follows readily. The proofs of Parts (b) and (c) follow along the same lines. ∎

3.4 MTP$_2$ Property

Now, we will discuss the multivariate total positivity of order two (MTP$_2$) property, in the sense of Karlin and Rinott [14], of the joint PDF of the GMBS distribution in (13). We shall use the following standard notation here. For any two real numbers a and b, let $a \vee b = \min\{a, b\}$ and $a \wedge b = \max\{a, b\}$. For $\boldsymbol{x} = (x_1, \cdots, x_p)^T$ and $\boldsymbol{y} = (y_1, \cdots, y_p)^T$, let $\boldsymbol{x} \vee \boldsymbol{y} = (x_1 \vee y_1, \cdots, x_p \vee y_p)^T$ and $\boldsymbol{x} \wedge \boldsymbol{y} = (x_1 \wedge y_1, \cdots, x_p \wedge y_p)^T$. Then, a function $g : \mathbf{R}^p \to \mathbf{R}^+$

is said to be MTP$_2$, in the sense of Karlin and Rinott [14], if $g(\boldsymbol{x})g(\boldsymbol{y}) \le g(\boldsymbol{x} \wedge \boldsymbol{y})g(\boldsymbol{x} \vee \boldsymbol{y})$. We then have the following result for the MBS distribution.

THEOREM 3: Let $\boldsymbol{T}$ be a p variate Birnbaum-Saunders distribution with parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma})$. If $\alpha_1 = \cdots = \alpha_p$, $\beta_1 = \cdots, \beta_p$, and all the off-diagonal elements of $\boldsymbol{\Gamma}^{-1}$ are less than or equal to zero, then the PDF of $\boldsymbol{T}$ has MTP$_2$ property.

PROOF: Suppose $\alpha_1 = \cdots \alpha_p = \alpha$, $\beta_1 = \cdots = \beta_p = \beta$, and we take $\boldsymbol{t}_1 = (t_{11}, \cdots, t_{1p})^T$ and $\boldsymbol{t}_2 = (t_{21}, \cdots, t_{2p})^T$ to be any two p-dimensional vectors. Then, to prove that the PDF of $\boldsymbol{T}$ has MTP$_2$ property, it is sufficient to show that

$$\boldsymbol{x}_1^T \boldsymbol{\Gamma}^{-1} \boldsymbol{x}_1 + \boldsymbol{x}_2^T \boldsymbol{\Gamma}^{-1} \boldsymbol{x}_2 \ge (\boldsymbol{x}_1 \vee \boldsymbol{x}_2)^T \boldsymbol{\Gamma}^{-1} (\boldsymbol{x}_1 \vee \boldsymbol{x}_2) + (\boldsymbol{x}_1 \wedge \boldsymbol{x}_2)^T \boldsymbol{\Gamma}^{-1} (\boldsymbol{x}_1 \wedge \boldsymbol{x}_2), \qquad (17)$$

where $\boldsymbol{x}_i = (x_{i1}, \cdots, x_{ip})^T$, for $i = 1, 2$, and

$$x_{ij} = \frac{1}{\alpha}\left(\sqrt{\frac{t_{ij}}{\beta}} - \sqrt{\frac{\beta}{t_{ij}}}\right), \quad i = 1, 2, \quad j = 1, \cdots, p.$$

If the elements of $\boldsymbol{\Gamma}^{-1}$ are denoted by $((\gamma^{kj}))$, for $k, j = 1, \cdots, p$, then proving (17) is equivalent to showing

$$\sum_{\substack{k, j = 1 \\ k \ne j}}^{p} (x_{1k}x_{1j} + x_{2k}x_{2j})\gamma^{kj} \ge \sum_{\substack{k, j = 1 \\ k \ne j}}^{p} ((x_{1k} \wedge x_{2k})(x_{1j} \wedge x_{2j}) + (x_{1k} \vee x_{2k})(x_{1j} \vee x_{2j}))\gamma^{kj}. \quad (18)$$

For all $k, j = 1, \cdots, p$,

$$x_{1k}x_{1j} + x_{2k}x_{2j} \le (x_{1k} \wedge x_{2k})(x_{1j} \wedge x_{2j}) + (x_{1k} \vee x_{2k})(x_{1j} \vee x_{2j}),$$

which can be easily shown by taking any ordering of $x_{1k}, x_{1j}, x_{2k}, x_{2j}$. Now, the result follows since $\gamma^{kj} \le 0$. ∎

It may be mentioned that the same result may not be true for other forms of generalized multivariate Birnbaum-Saunders distributions; see, for example, Sampson [25] in this regard. Moreover, it is immediate that Theorem 3.3 of Kundu et al. [17] follows from Theorem 3 above.

3.5 Shannon Entropy for GBS$_p$ distribution

In this section, we present the Shannon entropy for the GBS$_p$ distribution. We first need the following results for the required developments; see Arellano-Valle et al. [3] for proofs.

Lemma 2: Let $\boldsymbol{X} \sim EC_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, h^{(p)})$. Then, the Shannon entropy of $\boldsymbol{X}$, denoted by $H_{\boldsymbol{X}}^{EC_p}$, is given by

$$H_{\boldsymbol{X}}^{EC_p} = \frac{1}{2}\ln|\boldsymbol{\Sigma}| + H_{\boldsymbol{X}_0}^{EC_p},$$

where $H_{\boldsymbol{X}_0}^{EC_p}$ is the Shannon entropy of $\boldsymbol{X}_0 \sim EC_p(\boldsymbol{0}_p, \boldsymbol{I}_p, h^{(p)})$ given by

$$H_{\boldsymbol{X}_0}^{EC_p} = -\int_0^{+\infty} \left[\ln\left\{h^{(p)}(s)\right\}\right] g(s)\,ds$$

with $g(s) = \dfrac{\pi^{p/2}}{\Gamma(p/2)} s^{s/2-1} h^{(p)}(s)\,,\ s > 0.$

By using Lemma 2, for the multivariate normal case, when $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we have

$$H_{\boldsymbol{X}}^{N_p} = \frac{1}{2}\ln|\boldsymbol{\Sigma}| + \frac{p}{2}\left\{1 + \ln(2\pi)\right\}. \tag{19}$$

Similarly, for the multivariate-t case, when $\boldsymbol{X} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, we have

$$H_{\boldsymbol{X}}^{t_p} = \frac{1}{2}\ln|\boldsymbol{\Sigma}| - \ln\left\{\frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)(\nu\pi)^{p/2}}\right\} + \frac{\nu+p}{2}\left\{\psi\left(\frac{\nu+p}{2}\right) - \psi\left(\frac{\nu}{2}\right)\right\}, \tag{20}$$

where $\psi(x) = \dfrac{d\{\ln\Gamma(x)\}}{dx} = \dfrac{\Gamma'(x)}{\Gamma(x)}$ is the digamma function. Now, we present our main result concerning the Shannon entropy for the GBS$_p$ distribution.

Theorem 4: Let $\boldsymbol{T} \sim GBS_p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}, h^{(p)})$. Then, the Shannon entropy of $\boldsymbol{T}$, denoted by $H_{\boldsymbol{T}}^{GBS_p}$, is given by

$$H_{\boldsymbol{T}}^{GBS_p} = H_{\mathbf{X}}^{EC_p} + p\ln(2) + \sum_{i=1}^{p}\ln(\alpha_i) + \sum_{i=1}^{p}\ln(\beta_i) + \frac{3}{2}\sum_{i=1}^{p} E\left[\ln(T_i)\right] - \sum_{i=1}^{p} E\left[\ln(T_i+1)\right], \tag{21}$$

where $H_{\boldsymbol{X}}^{EC_p}$ denotes the Shannon entropy of $\boldsymbol{X} \sim EC_p(\boldsymbol{0}, \boldsymbol{\Gamma}, h^{(p)})$ and $T_i \sim GBS\left(\alpha_i, 1, h^{(1)}\right)$, for $i = 1, 2, \cdots, p$.

Proof: It is easy to obtain the expression in (21) using the PDF of the generalized multivariate BS distribution as obtained in (13). ∎

Another equivalent form can be obtained as

$$\begin{aligned} H_{\boldsymbol{T}}^{GBS_p} &= H_{\boldsymbol{X}}^{EC_p} + p\ln(2) + \sum_{i=1}^{p}\ln(\alpha_i) + \sum_{i=1}^{p}\ln(\beta_i) + 3\sum_{i=1}^{p} E\left[\ln\left(\left\{V_i + \sqrt{1+V_i^2}\right\}\right)\right] \\ &\quad - \sum_{i=1}^{p} E\left[\ln\left(\left\{V_i + \sqrt{1+V_i^2}\right\}^2 + 1\right)\right], \end{aligned} \quad (22)$$

where $V_i \sim EC\left(0, \alpha_i^2, h^{(1)}\right)$, for $i = 1, 2, \cdots, p$. Upon substituting the expressions in (19) and (20) into (21) (or (22)), we can obtain the Shannon entropy for $\boldsymbol{T}$, in the case of multivariate normal and t kernels, respectively.

3.6 Generation

In this section, we consider some of the important special cases, and describe how the corresponding random vectors can be simulated.

Case 1: Multivariate Birnbaum-Saunders Distribution

The random vector $\boldsymbol{T} = (T_1, \cdots, T_p)^T$ is said to have a p-variate Birnbaum-Saunders distribution if it has the joint PDF

$$\begin{aligned} f_{\boldsymbol{T}}(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}) &= \phi_p\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\beta_1}} - \sqrt{\frac{\beta_1}{t_1}}\right), \cdots, \frac{1}{\alpha_p}\left(\sqrt{\frac{t_p}{\beta_p}} - \sqrt{\frac{\beta_p}{t_p}}\right); \boldsymbol{\Gamma}\right) \\ &\quad \times \prod_{i=1}^{p} \frac{1}{2\alpha_i\beta_i}\left\{\left(\frac{\beta_i}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta_i}{t_i}\right)^{\frac{3}{2}}\right\} \end{aligned} \quad (23)$$

for $t_1 > 0, \cdots, t_p > 0$; here, for $\boldsymbol{u} = (u_1, \cdots, u_p)^T$,

$$\phi_p(u_1, \cdots, u_p; \boldsymbol{\Gamma}) = \frac{1}{(2\pi)^{\frac{p}{2}} |\boldsymbol{\Gamma}|^{\frac{1}{2}}} e^{-\frac{1}{2}\boldsymbol{u}^T \boldsymbol{\Gamma}^{-1} \boldsymbol{u}} \quad (24)$$

is the PDF of the standard normal vector with correlation matrix $\boldsymbol{\Gamma}$. Hereafter, the p-variate BS distribution, with joint PDF in (23), will be denoted by $\text{BS}_p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma})$. For $p = 2$, the bivariate Birnbaum-Saunders distribution has been discussed in detail by Kundu et al. [17]. The following algorithm can be adopted to generate $\boldsymbol{T} = (T_1, \cdots, T_p)^T$ from $\mathrm{BS}_p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma})$ in (23).

Algorithm 1

Step 1: Make a Cholesky decomposition of $\boldsymbol{\Gamma} = \boldsymbol{A}\boldsymbol{A}^T$ (say);

Step 2: Generate p independent standard normal random numbers, say, $U_1, \cdots, U_p$;

Step 3: Compute $\boldsymbol{Z} = (Z_1, \cdots, Z_p)^T = \boldsymbol{A}\ (U_1, \cdots, U_p)^T$;

Step 4: Perform the transformation

$$T_i = \beta_i \left[\frac{1}{2}\alpha_i Z_i + \sqrt{\left(\frac{1}{2}\alpha_i Z_i\right)^2 + 1} \right]^2 \quad \text{for} \quad i = 1, \cdots, p.$$

Then, $\boldsymbol{T} = (T_1, \cdots, T_p)^T$ has the required $\mathrm{BS}_p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma})$ density as in (23).

Case 2: GMBS Distribution Induced by a Multivariate-t Kernel

The random vector $\boldsymbol{T}$ is said to have a p-variate generalized multivariate BS distribution induced by the multivariate-t kernel with ν degrees of freedom (will be denoted by $\boldsymbol{T} \sim \mathrm{BS}t_p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}, \nu)$) if the joint PDF of $\boldsymbol{T}$ is

$$\begin{aligned} f_{\boldsymbol{T}}(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}, \nu) &= g_p\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\beta_1}} - \sqrt{\frac{\beta_1}{t_1}}\right), \cdots, \frac{1}{\alpha_p}\left(\sqrt{\frac{t_p}{\beta_p}} - \sqrt{\frac{\beta_p}{t_p}}\right); \boldsymbol{\Gamma}, \nu \right) \\ &\quad \times \prod_{i=1}^{p} \frac{1}{2\alpha_i\beta_i}\left\{ \left(\frac{\beta_i}{t_i}\right)^{1/2} + \left(\frac{\beta_i}{t_i}\right)^{3/2} \right\}, \end{aligned} \tag{25}$$

for $t_1 > 0, \cdots, t_p > 0$, where $g_p(\cdot\,; \boldsymbol{\Gamma}, \nu)$ is the PDF of $t_p(\boldsymbol{0}, \boldsymbol{\Gamma}, \nu)$, the p-variate Student's t distribution with location parameter $\boldsymbol{0}$, scale parameter $\boldsymbol{\Gamma}$, and ν degrees of freedom, given by

$$g_p(\boldsymbol{u}; \boldsymbol{\Gamma}, \nu) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)(\nu\pi)^{p/2}|\boldsymbol{\Gamma}|^{\frac{1}{2}}} \left(1 + \frac{\mathbf{u}^T\boldsymbol{\Gamma}^{-1}\mathbf{u}}{\nu}\right)^{-(\nu+p)/2} \tag{26}$$

for $\boldsymbol{u} = (u_1, u_2, \cdots, u_p)^T \in \mathbb{R}^p$.

Note that there are several ways to generate p variate Student's t distribution; see, for example, Kotz and Nadarajah [16]. Suppose $\boldsymbol{Z} = (Z_1, \cdots, Z_p)^T$ has been generated from a $t_p(\mathbf{0}, \boldsymbol{\Gamma}, \nu)$ distribution. Then, by using Step 4 of Algorithm 1, the required $\boldsymbol{T}$ having $\mathrm{BS}t_p(\alpha, \beta, \boldsymbol{\Gamma}, \nu)$ can be obtained.

Case 3: GMBS Distribution Induced by The Multivariate Power Normal Kernel

The random vector $\boldsymbol{T}$ is said to have a p-variate generalized multivariate BS distribution induced by the multivariate power normal kernel (will be denoted by $\boldsymbol{T} \sim \mathrm{BSPN}_p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}, \delta)$) if the joint PDF of $\boldsymbol{T}$ is

$$f_{\boldsymbol{T}}(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}, \delta) = h_p\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\beta_1}} - \sqrt{\frac{\beta_1}{t_1}}\right), \cdots, \frac{1}{\alpha_p}\left(\sqrt{\frac{t_p}{\beta_p}} - \sqrt{\frac{\beta_p}{t_p}}\right); \boldsymbol{\Gamma}, \delta\right) \times \prod_{i=1}^{p} \frac{1}{2\alpha_i\beta_i}\left\{\left(\frac{\beta_i}{t_i}\right)^{1/2} + \left(\frac{\beta_i}{t_i}\right)^{3/2}\right\}, \tag{27}$$

for $t_1 > 0, \cdots, t_p > 0$, where $h_p(\cdot\,; \boldsymbol{\Gamma}, \delta)$ is the PDF of the multivariate power normal distribution with location parameter $\mathbf{0}$, scale parameter $\boldsymbol{\Gamma}$, and shape parameter δ given by

$$h_p(\boldsymbol{u}; \boldsymbol{\Gamma}, \delta) = \frac{\Gamma\left(\frac{p}{2}\right)}{2^{\frac{p}{2\delta}} \Gamma\left(\frac{p}{2\delta}\right) \pi^{\frac{p}{2}} |\boldsymbol{\Gamma}|^{\frac{1}{2}}} e^{-\frac{1}{2}\left(\mathbf{u}^T \boldsymbol{\Gamma}^{-1} \mathbf{u}\right)^{\delta}} \tag{28}$$

for $\boldsymbol{u} = (u_1, u_2, \cdots, u_p)^T \in \mathbb{R}^p$.

Naik and Plungpongpun [22] proposed an efficient method for generating from a multivariate power normal distribution. Once $\boldsymbol{Z} = (Z_1, \cdots, Z_p)^T$ has been generated from a multivariate power normal distribution with location vector $\mathbf{0}$ and scale matrix $\boldsymbol{\Sigma}$, using their method, then by using Step 4 of Algorithm 1, the required $\boldsymbol{T}$ from $\mathrm{BSPN}_p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}, \delta)$ can be obtained.

4 Inference

In this section, we discuss the maximum likelihood estimates (MLEs) of the model parameters and the associated inference, based on the observed data $\{(t_{i1},\cdots,t_{ip})^T; i=1,\cdots,n\}$. We shall assume that the data are from a generalized Birnbaum-Saunders distribution with the kernel function being specified. We then consider two different kernel functions in detail, viz., (i) multivariate normal kernel and (ii) multivariate-t kernel with a specified degrees of freedom. It is worth mentioning that the multivariate-t kernel with one degree of freedom corresponds to the multivariate Cauchy kernel.

4.1 Multivariate Normal Kernel

4.1.1 Maximum Likelihood Estimation

The log-likelihood function, without the additive constant, is given by

$$\begin{aligned} l(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Gamma}|data) &= -\frac{n}{2}\ln|\boldsymbol{\Gamma}| - \frac{1}{2}\sum_{i=1}^{n}\boldsymbol{v}_i^T\Gamma^{-1}\boldsymbol{v}_i - n\sum_{j=1}^{p}\ln\alpha_j - n\sum_{j=1}^{p}\ln\beta_j \\ &\quad + \sum_{i=1}^{n}\sum_{j=1}^{p}\ln\left\{\left(\frac{\beta_{ij}}{t_{ij}}\right)^{\frac{1}{2}} + \left(\frac{\beta_{ij}}{t_{ij}}\right)^{\frac{3}{2}}\right\}, \end{aligned} \tag{29}$$

where

$$\boldsymbol{v}_i^T = \left[\frac{1}{\alpha_1}\left(\sqrt{\frac{t_{i1}}{\beta_1}} - \sqrt{\frac{\beta_1}{t_{i1}}}\right),\cdots,\frac{1}{\alpha_p}\left(\sqrt{\frac{t_{ip}}{\beta_p}} - \sqrt{\frac{\beta_p}{t_{ip}}}\right)\right]. \tag{30}$$

Then, the MLEs of the unknown parameters can be obtained by maximizing (29) with respect to the parameters $\boldsymbol{\alpha},\boldsymbol{\beta}$ and $\boldsymbol{\Gamma}$, which would require a $2p+\binom{p}{2}$ dimensional optimization process. For this reason, we adopt the following procedure in order to reduce the computational effort significantly. Observe that

$$\left[\left(\sqrt{\frac{T_1}{\beta_1}} - \sqrt{\frac{\beta_1}{T_1}}\right),\cdots,\left(\sqrt{\frac{T_p}{\beta_p}} - \sqrt{\frac{\beta_p}{T_p}}\right)\right]^T \sim N_p\left(\mathbf{0},\boldsymbol{D}\boldsymbol{\Gamma}\boldsymbol{D}^T\right), \tag{31}$$

where $\boldsymbol{D}$ is a diagonal matrix given by $\boldsymbol{D} = \text{diag}\{\alpha_1, \cdots, \alpha_p\}$. Therefore, for given $\boldsymbol{\beta}$, the MLEs of $\boldsymbol{\alpha}$ and $\boldsymbol{\Gamma}$ become

$$\widehat{\alpha}_j(\boldsymbol{\beta}) = \left(\frac{1}{n}\sum_{i=1}^{n}\left(\sqrt{\frac{t_{ij}}{\beta_j}} - \sqrt{\frac{\beta_j}{t_{ij}}}\right)^2\right)^{\frac{1}{2}} = \left(\frac{1}{\beta_j}\left\{\frac{1}{n}\sum_{i=1}^{n} t_{ij}\right\} + \beta_j\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{t_{ij}}\right\} - 2\right)^{\frac{1}{2}}, \quad j = 1, \cdots, p, \tag{32}$$

and

$$\widehat{\boldsymbol{\Gamma}}(\boldsymbol{\beta}) = \boldsymbol{P}(\boldsymbol{\beta})\boldsymbol{Q}(\boldsymbol{\beta})\boldsymbol{P}^T(\boldsymbol{\beta}); \tag{33}$$

here, $\boldsymbol{P}(\boldsymbol{\beta})$ is a diagonal matrix given by $\boldsymbol{P}(\boldsymbol{\beta}) = \text{diag}\{1/\widehat{\alpha}_1(\boldsymbol{\beta}), \cdots, 1/\widehat{\alpha}_p(\boldsymbol{\beta})\}$, and the elements $q_{jk}(\boldsymbol{\beta})$ of the matrix $\boldsymbol{Q}(\boldsymbol{\beta})$ are given by

$$q_{jk}(\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n}\left(\sqrt{\frac{t_{ij}}{\beta_j}} - \sqrt{\frac{\beta_j}{t_{ij}}}\right)\left(\sqrt{\frac{t_{ik}}{\beta_k}} - \sqrt{\frac{\beta_k}{t_{ik}}}\right) \quad \text{for } j, k = 1, \cdots, p. \tag{34}$$

Thus, we obtain the p-dimensional profile log-likelihood function $l(\widehat{\boldsymbol{\alpha}}(\boldsymbol{\beta}), \boldsymbol{\beta}, \widehat{\boldsymbol{\Gamma}}(\boldsymbol{\beta})|data)$. The MLE of $\boldsymbol{\beta}$ can then be obtained by maximizing the p-dimensional profile log-likelihood function, and once we get the MLE of $\boldsymbol{\beta}$, say $\widehat{\boldsymbol{\beta}}$, the MLEs of $\boldsymbol{\alpha}$ and $\boldsymbol{\Gamma}$ can be obtained readily by substituting $\widehat{\boldsymbol{\beta}}$ in place of $\boldsymbol{\beta}$ in (32) and (33), respectively.

However, for computing the MLEs of the unknown parameters, we need to maximize the profile log-likelihood function of $\boldsymbol{\beta}$ and we may use the Newton-Raphson iterative process for this purpose. Finding a proper p-dimensional initial guess value of $\boldsymbol{\beta}$ becomes quite important in this case. Modified moment estimators, similar to those proposed by Ng et al. [23], can be used effectively for this purpose, and they are as follows:

$$\beta_j^{(0)} = \left(\frac{1}{n}\sum_{i=1}^{n} t_{ij} \bigg/ \frac{1}{n}\sum_{i=1}^{n}\frac{1}{t_{ij}}\right)^{\frac{1}{2}}, \qquad j = 1, \cdots, p. \tag{35}$$

Note that if $\boldsymbol{\beta}$ is known, then the MLEs of $\boldsymbol{\alpha}$ and $\boldsymbol{\Gamma}$ can be obtained explicitly.

Now, we discuss the asymptotic properties of the MLEs when all the parameters are unknown, and also when some parameters are known. First, we shall consider the case when all the parameters are unknown.

THEOREM 5: If $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma})$ is the parameter vector and $\widehat{\boldsymbol{\theta}}$ denotes the corresponding MLE, then

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N_m(\mathbf{0}, \boldsymbol{J}^{-1}), \tag{36}$$

with $m = p + \binom{p}{2}$ being the dimension of the vector $\boldsymbol{\theta}$. Here, $\xrightarrow{d}$ denotes convergence in distribution while $N_m(\mathbf{0}, \boldsymbol{J}^{-1})$ denotes the m-variate normal distribution with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{J}^{-1}$, with $\boldsymbol{J}$ being the Fisher information matrix. Expressions of all the elements of the Fisher information matrix $\boldsymbol{J}$ are presented in the Appendix.

PROOF: Since the multivariate BS distribution satisfies all the regularity conditions for the MLEs to be consistent and asymptotically normally distributed, the result follows from the standard asymptotic properties of MLEs. ∎

If $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are known, then the MLE of $\boldsymbol{\Gamma}$ is $\widehat{\boldsymbol{\Gamma}} = \boldsymbol{D}^{-1}\boldsymbol{Q}(\boldsymbol{\beta})\boldsymbol{D}^{-1}$, where the elements of the matrix $\boldsymbol{Q}(\boldsymbol{\beta})$ are as in (34) and the matrix $\boldsymbol{D}$ is as defined earlier. From (31), it immediately follows in this case that $\widehat{\boldsymbol{\Gamma}}$ has a Wishart distribution with parameters p and $\boldsymbol{\Gamma}$. Furthermore, if only $\boldsymbol{\beta}$ is known, it is clear that $\widehat{\alpha}_j^2(\boldsymbol{\beta})$, defined in (32), is distributed as χ_1^2 for $j = 1, \cdots, p$.

4.1.2 MODIFIED MOMENT ESTIMATORS

Since the MLEs do not have explicit form and need to be obtained by solving p non-linear equations, we propose the following modified moment estimators for the unknown parameters, along the lines of Kundu et al. [17]. The modified moment estimators can be obtained by equating the moments and inverse moments with the corresponding sample quantities. If we denote

$$s_j = \frac{1}{n}\sum_{i=1}^{n} t_{ij} \quad \text{and} \quad r_j = \left[\frac{1}{n}\sum_{i=1}^{n} t_{ij}^{-1}\right]^{-1}, \quad j = 1, \cdots, p,$$

then the modified moment estimators of α_j, β_j and ρ_{jk}, for $j, k = 1, \cdots p$, are

$$\widetilde{\alpha}_j = \left\{2\left[\left(\frac{s_j}{r_j}\right)^{1/2} - 1\right]\right\}^{1/2}, \quad \widetilde{\beta}_j = (s_j r_j)^{1/2}, \tag{37}$$

and

$$\widetilde{\rho}_{jk} = \frac{\sum_{i=1}^{n}\left(\sqrt{\frac{t_{ij}}{\widetilde{\beta}_j}} - \sqrt{\frac{\widetilde{\beta}_j}{t_{ij}}}\right)\left(\sqrt{\frac{t_{ik}}{\widetilde{\beta}_k}} - \sqrt{\frac{\widetilde{\beta}_k}{t_{ik}}}\right)}{\sqrt{\sum_{i=1}^{n}\left(\sqrt{\frac{t_{ij}}{\widetilde{\beta}_j}} - \sqrt{\frac{\widetilde{\beta}_j}{t_{ij}}}\right)^2}\sqrt{\sum_{i=1}^{n}\left(\sqrt{\frac{t_{ik}}{\widetilde{\beta}_k}} - \sqrt{\frac{\widetilde{\beta}_k}{t_{ik}}}\right)^2}}. \tag{38}$$

4.2 Multivariate-t Kernel

For a given ν, the log-likelihood function, without the additive constant, is given by

$$\begin{aligned} l(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}|data, \nu) &= -\frac{n}{2}\ln|\boldsymbol{\Gamma}| - \frac{\nu+p}{2}\sum_{i=1}^{n}\ln\left(1 + \frac{\boldsymbol{v}_i^T\Gamma^{-1}\boldsymbol{v}_i}{\nu}\right) - n\sum_{j=1}^{p}\ln\alpha_j - n\sum_{j=1}^{p}\ln\beta_j \\ &+ \sum_{i=1}^{n}\sum_{j=1}^{p}\ln\left\{\left(\frac{\beta_{ij}}{t_{ij}}\right)^{\frac{1}{2}} + \left(\frac{\beta_{ij}}{t_{ij}}\right)^{\frac{3}{2}}\right\}, \end{aligned} \tag{39}$$

where $\boldsymbol{v}_i$, for $i = 1, \cdots, n$, are as defined earlier in (30). Here also, the MLEs of the unknown parameters can be obtained by maximizing (39), but it involves a $2p + \binom{p}{2}$ optimization process. Hence, as done in the case of multivariate normal kernel, we adopt the following procedure to reduce the computational burden. In this case, we have

$$\left[\left(\sqrt{\frac{T_1}{\beta_1}} - \sqrt{\frac{\beta_1}{T_1}}\right), \cdots, \left(\sqrt{\frac{T_p}{\beta_p}} - \sqrt{\frac{\beta_p}{T_p}}\right)\right]^T \sim t_p\left(\mathbf{0}, \boldsymbol{D}\boldsymbol{\Gamma}\boldsymbol{D}^T, \nu\right), \tag{40}$$

where the matrix $\boldsymbol{D}$ is the same diagonal matrix as defined earlier. If we define $\boldsymbol{R} = \boldsymbol{D}\boldsymbol{\Gamma}\boldsymbol{D}^T$, then the MLE of $\boldsymbol{R}$ can be obtained as the solution of the equation

$$\boldsymbol{R} = \frac{1}{n}\sum_{i=1}^{n} w_i \boldsymbol{u}_i \boldsymbol{u}_i^T, \tag{41}$$

where $w_i = (\nu + p)/(\nu + s_i)$, $s_i = \boldsymbol{u}_i^T \boldsymbol{R}^{-1} \boldsymbol{u}_i$, and

$$\boldsymbol{u}_i^T = \left[\left(\sqrt{\frac{t_{i1}}{\beta_1}} - \sqrt{\frac{\beta_1}{t_{i1}}}\right), \cdots, \left(\sqrt{\frac{t_{ip}}{\beta_p}} - \sqrt{\frac{\beta_p}{t_{ip}}}\right)\right]; \tag{42}$$

see, for example, Nadarajah and Kotz [21]. The following simple iterative process

$$\boldsymbol{R}^{(m+1)} = \frac{1}{n}\sum_{i=1}^{n} w_i^{(m)} \boldsymbol{u}_i \boldsymbol{u}_i^T, \tag{43}$$

where

$$w_i^{(m)} = (\nu + p)/\{\nu + \boldsymbol{u}_i^T (\boldsymbol{R}^{(m)})^{-1} \boldsymbol{u}_i\}$$

can be used to find the solution of (41); see Nadarajah and Kotz [21]. Then, if $\widehat{\boldsymbol{R}}(\boldsymbol{\beta}) = ((r_{jk}(\boldsymbol{\beta})))$ is the solution of (41), the MLEs of $\boldsymbol{\alpha}$ and $\boldsymbol{\Gamma}$ are given by

$$\widehat{\alpha}_j(\boldsymbol{\beta}) = \sqrt{r_{jj}}, \quad j = 1, \cdots, p, \tag{44}$$

and

$$\widehat{\boldsymbol{\Gamma}}(\boldsymbol{\beta}) = \boldsymbol{P}(\boldsymbol{\beta}) \widehat{\boldsymbol{R}}(\boldsymbol{\beta}) \boldsymbol{P}^T(\boldsymbol{\beta}); \tag{45}$$

here, the diagonal matrix $\boldsymbol{P}(\boldsymbol{\beta}) = \text{diag}\{1/\widehat{\alpha}_1(\boldsymbol{\beta}), \cdots, 1/\widehat{\alpha}_p(\boldsymbol{\beta})\}$ is same as defined earlier. Therefore, the MLE of $\boldsymbol{\beta}$ can be obtained by maximizing the profile log-likelihood function of $\boldsymbol{\beta}$. Once we get the MLE of $\boldsymbol{\beta}$, the MLEs of $\boldsymbol{\alpha}$ and $\boldsymbol{\Gamma}$ can be obtained, as given above, in Eqs. (44) and (45) respectively.

5 Illustrative Example

5.1 Multivariate Normal Kernel

In this section, we analyze a multivariate data by using the proposed generalized multivariate BS distribution with multivariate normal kernel, for the purpose of illustration. These data, taken from Johnson and Wichern [12] (page 34), represent the mineral contents of four major bones of 25 new born babies. Here, T_1, T_2, T_3 and T_4 represent dominant radius, radius, dominant ulna and ulna, respectively. The data are not presented here, but the summary statistics of the sample mean, variance and skewness of the individual T_i's and their reciprocals are all presented in Table 1.

Table 1: The sample mean, variance and coefficient of skewness of T_i and T_i^{-1} for $i = 1, \cdots, 4$.

Variables → Statistics ↓	T_1	T_1^{-1}	T_2	T_2^{-1}	T_3	T_3^{-1}	T_4	T_4^{-1}
Mean	0.844	1.211	0.818	1.245	0.704	1.452	0.694	1.474
Variance	0.012	0.041	0.011	0.034	0.011	0.050	0.010	0.052
Skewness	-0.793	2.468	-0.543	1.679	-0.022	0.381	-0.133	0.755

It is clear from Table 1 that all T_i and T_i^{-1} (for $i = 1, \cdots, 4$) are quite skewed. To get an idea about the hazard functions of T_i's and T_i^{-1}'s, we have plotted in Figures 1 and 2 the marginal scaled TTT transforms of T_i's and T_i^{-1}'s, respectively, as suggested by Aarset [1]; see Kundu et al. (2008) and Azevedo et al. (2012) for a detailed analysis of hazard functions of univariate BS distributions based on normal and t kernels.

Since all of them are concave first and then convex, the plots do seem to suggest that the hazard functions are all unimodal. For checking whether the Birnbaum-Saunders distribution can be used for fitting the marginal distributions, the modified moment estimates of α_i and β_i (for $i = 1, \cdots, 4$), as proposed by Ng et al. [23], were computed. Using these values, the Kolmogorov-Smirnov (KS) distances between the empirical distribution function and the fitted distribution function and the corresponding p-values (determined by Monte Carlo simulations) were computed, and these results are presented in Table 2. The obtained results suggest that the Birnbaum-Saunders distribution fits all the marginals very well.

All these suggest that we could fit 4-variate Birnbaum-Saunders distribution to the considered data. Using modified moment estimates as initial guess, we found the MLEs of β_1, β_2, β_3 and β_4 to be 0.8547, 0.7907, 0.7363 and 0.8161, respectively. Finally, the corresponding maximum likelihood estimates of α_1, α_2, α_3 and α_4 were obtained to be 0.1491, 0.1393, 0.1625 and 0.2304, respectively, and the corresponding maximized log-likelihood value to be 182.561896. The 95% non-parametric bootstrap confidence intervals of β_1, β_2, β_3 and

Table 2: The modified moment estimates of α_i and β_i, the KS distance between the empirical distribution function and the fitted distribution function, and the corresponding p values.

	α	β	KS distance	p
T_1	0.1473	0.8347	0.161	0.537
T_2	0.1372	0.8107	0.145	0.671
T_3	0.1525	0.6963	0.109	0.929
T_4	0.1503	0.6861	0.094	0.979

β_4 were then obtained as (0.8069, 0.9025), (0.7475, 0.8339), (0.6950, 0.7776) and (0.7760, 0.8562), respectively. Similarly, the 95% non-parametric bootstrap confidence intervals of α_1, α_2, α_3 and α_4 were obtained as (0.1085, 0.1897), (0.1015, 0.1771), (0.1204, 0.2046) and (0.1890, 0.2718), respectively. The maximum likelihood estimate of $\mathbf{\Gamma}$ is obtained as

$$\widehat{\mathbf{\Gamma}} = \begin{bmatrix} 1.000 & 0.767 & 0.715 & 0.515 \\ 0.767 & 1.000 & 0.612 & 0.381 \\ 0.715 & 0.612 & 1.000 & 0.693 \\ 0.515 & 0.381 & 0.693 & 1.000 \end{bmatrix}. \tag{46}$$

Now, suppose we are interested in testing the hypotheses

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 (= \beta, \text{ say}) \quad vs. \quad H_1 : \text{they are not all equal.}$$

In this case, we find the maximum likelihood estimate of the common β as 0.7689, and the constrained maximum likelihood estimates of α_1, α_2, α_3 and α_4 to be 0.1689, 0.1471, 0.1823 and 0.1890, respectively. The corresponding constrained maximum likelihood estimate of $\mathbf{\Gamma}$ is obtained as

$$\widetilde{\mathbf{\Gamma}} = \begin{bmatrix} 1.000 & 0.841 & 0.253 & 0.131 \\ 0.841 & 1.000 & 0.372 & 0.378 \\ 0.252 & 0.372 & 1.000 & 0.800 \\ 0.131 & 0.378 & 0.800 & 1.000 \end{bmatrix}, \tag{47}$$

with the corresponding maximized log-likelihood value as 124.81564. So, by using the likelihood ratio test, we conclude that there is no evidence to support H_0 since we have $p < 10^{-8}$ in this case.

5.2 Multivariate-t Kernel

In this section, we re-analyze the same data considered in the preceding subsection by using the generalized 4-variate Birnbaum-Saunders distribution with mutivariate-t kernel. We varied the degrees of freedom ν from 1 to 20 for profile analysis with respect to ν. We computed the MLEs of all the parameters and the corresponding maximized log-likelihood values for different choices of ν. We observed that the maximized log-likelihood values increase first and then decrease. The maximized log-likelihood values, computed as a function of the degrees of freedom ν, through this discrete profile search, are presented in Table 3, and these values have also been plotted in Figure 3.

Table 3: The maximized log-likelihood value vs. degrees of freedom $\nu = 1(1)20$.

ν	Maximized log-likelihood	ν	Maximized log-likelihood	ν	Maximized log-likelihood	ν	Maximized log-likelihood
1	181.510422	2	187.544418	3	189.382248	4	189.978531
5	190.111679	6	190.048141	7	189.984222	8	189.889908
9	189.769516	10	189.638428	11	189.505157	12	189.374313
13	189.266373	14	189.175812	15	189.088974	16	189.006287
17	188.927994	18	188.854111	19	188.784485	20	188.718903

We observe that the maximum occurs at $\nu = 5$, with the associated log-likelihood value (without the additive constant) being 190.112. It is important to mention here that the selection of the best t-kernel function through the maximized log-likelihood value is equivalent to selecting by the Akaike Information Criterion since the number of model parameters remains the same when ν varies. Furthermore, this maximized log-likelihood value of 190.112 for the multivariate-t kernel with $\nu = 5$ degrees of freedom is significantly larger than the corresponding value of 182.562 for the multivariate normal kernel, which does provide evidence to the fact that the multivariate-t kernel provides a much better fit for these data.

Now, we provide detailed results for the case $\nu = 5$. In this case, the MLEs of β_1, β_2,

β_3 and β_4 are found to be 0.8686, 0.8469, 0.7349, 0.7156, respectively. The corresponding 95% confidence intervals, obtained by the use of non-parametric bootstrap method, are (0.7975, 0.9423), (0.7991, 0.9147), (0.6614, 0.7967), and (0.6357, 0.7855), respectively. The MLEs of α_1, α_2, α_3 and α_4 are 0.1057, 0.1357, 0.1413 and 0.1372 and the associated 95% non-parametric bootstrap confidence intervals are (0.0745, 0.1387), (0.0975, 0.1689), (0.1115, 0.1712), and (0.1015, 0.1655), respectively. The maximum likelihood estimate of $\mathbf{\Gamma}$ is obtained as

$$\widehat{\mathbf{\Gamma}} = \begin{bmatrix} 1.000 & 0.812 & 0.699 & 0.493 \\ 0.812 & 1.000 & 0.598 & 0.412 \\ 0.699 & 0.598 & 1.000 & 0.711 \\ 0.493 & 0.412 & 0.711 & 1.000 \end{bmatrix}. \tag{48}$$

6 Concluding Remarks

In this paper, we have introduced a p-variate generalized Birnbaum-Saunders distribution, and derived many of its properties. The maximum likelihood estimates of the model parameters for two special cases have been discussed in detail. It has been observed that in both these cases, the maximum likelihood estimates can be obtained numerically by an optimization process in conjunction with the profile likelihood method. Explicit expressions for the elements of the Fisher information matrix in the case of multivariate normal kernel have been provided. For illustrative purposes, one data set has been analyzed in which case it has been shown that the generalized Birnbaum-Saunders distribution with the multivariate-t kernel provides a better fit than the one with the multivariate normal kernel.

The estimation of the parameters of the generalized multivariate Birnbaum-Saunders distribution, for an arbitrary kernel function, as well as model selection and model discrimination within this general family still remain as challenging open problems. Work is currently under progress on these problems and we hope to report these findings in a future paper.

Acknowledgements

The authors thank the referees and the associate editor for their constructive comments and suggestions on an earlier version of this manuscript which led to this improved revised version.

Appendix A: Fisher Information Matrix

For deriving the elements of the expected Fisher information matrix, the following expressions and results are useful. Let $\boldsymbol{T} \sim \mathrm{BS}_p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma})$ and $\boldsymbol{\Gamma} = ((\gamma_{ik}))$. Then:

(a)

$$E\left\{\left(\sqrt{\frac{T_i}{\beta_i}} - \sqrt{\frac{\beta_i}{T_i}}\right)\left(\sqrt{\frac{T_k}{\beta_k}} - \sqrt{\frac{\beta_k}{T_k}}\right)\right\} = \alpha_i \alpha_k \gamma_{ik}, \quad i \neq k = 1, \cdots, p; \tag{49}$$

(b)

$$E\left(\sqrt{\frac{T_k}{\beta_k}} - \sqrt{\frac{\beta_k}{T_k}}\right)^2 = \alpha_k^2, \quad k = 1, \cdots, p. \tag{50}$$

If we denote

$$E\left(\sqrt{\frac{T_i T_k}{\beta_i \beta_k}}\right) = \psi_1(\alpha_i, \alpha_k, \gamma_{ik}), \quad i \neq k = 1 \cdots, p, \tag{51}$$

then we immediately have, for $i \neq k = 1, \cdots, p$,

$$E\left(\sqrt{\frac{\beta_i \beta_k}{T_i T_k}}\right) = \psi_1(\alpha_i, \alpha_k, \gamma_{ik}) \quad \text{and} \quad E\left(\sqrt{\frac{\beta_i T_k}{T_i \beta_k}}\right) = \psi_1(\alpha_i, \alpha_k, -\gamma_{ik}).$$

An explicit expression for $\psi_1(\cdot)$ has been given by Kundu et al. [17]. Moreover,

$$\frac{\partial}{\partial \gamma_{ik}}\left(\boldsymbol{\Gamma}^{-1}\right) = -\boldsymbol{\Gamma}^{-1}\left(\frac{\partial}{\partial \gamma_{ik}}\boldsymbol{\Gamma}\right)\boldsymbol{\Gamma}^{-1} = \boldsymbol{B}^{ik} = ((b^{ik}_{j_1,j_2})) \ \text{(say)}, \quad i, k, j_1, j_2 = 1, \cdots, p,$$

$$\frac{\partial^2}{\partial \gamma_{ik}^2}\left(\boldsymbol{\Gamma}^{-1}\right) = 2\boldsymbol{\Gamma}^{-1}\left(\frac{\partial}{\partial \gamma_{ik}}\boldsymbol{\Gamma}\right)\boldsymbol{\Gamma}^{-1}\left(\frac{\partial}{\partial \gamma_{ik}}\boldsymbol{\Gamma}\right)\boldsymbol{\Gamma}^{-1} = 2\mathbf{A}^{ik} = 2((a^{ik}_{j_1,j_2})) \ \text{(say)},$$

for $i, k, j_1, j_2 = 1, \cdots, p$. Furthermore,

$$\begin{aligned} \frac{\partial^2}{\partial \gamma_{ik} \partial \gamma_{st}}\left(\boldsymbol{\Gamma}^{-1}\right) &= -\boldsymbol{\Gamma}^{-1}\left(\frac{\partial}{\partial \gamma_{ik}}\boldsymbol{\Gamma}\right)\boldsymbol{\Gamma}^{-1}\left(\frac{\partial}{\partial \gamma_{st}}\boldsymbol{\Gamma}\right)\boldsymbol{\Gamma}^{-1} - \boldsymbol{\Gamma}^{-1}\left(\frac{\partial}{\partial \gamma_{st}}\boldsymbol{\Gamma}\right)\boldsymbol{\Gamma}^{-1}\left(\frac{\partial}{\partial \gamma_{ik}}\boldsymbol{\Gamma}\right)\boldsymbol{\Gamma}^{-1} \\ &= -\boldsymbol{C}^{ikst} = ((c^{ikst}_{j_1,j_2})), \quad \text{(say)}, \end{aligned}$$

for $(i,k) \neq (s,t)$ or $(i,k) \neq (t,s)$, $i,k,s,t,j_1,j_2 = 1,\cdots,p$. Let us denote $\mathbf{\Gamma}^{-1} = ((\gamma^{ik}))$, and

$$f(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{\Gamma}) = -\frac{1}{2}\ln|\mathbf{\Gamma}| - \frac{1}{2}\boldsymbol{V}^T\Gamma^{-1}\boldsymbol{V} - \sum_{j=1}^{p}\ln\alpha_j - \sum_{j=1}^{p}\ln\beta_j + \sum_{j=1}^{p}\ln\left\{\left(\frac{\beta_j}{T_j}\right)^{\frac{1}{2}} + \left(\frac{\beta_j}{T_j}\right)^{\frac{3}{2}}\right\}, \tag{52}$$

where

$$\boldsymbol{V}^T = \left[\frac{1}{\alpha_1}\left(\sqrt{\frac{T_1}{\beta_1}} - \sqrt{\frac{\beta_1}{T_1}}\right),\cdots,\frac{1}{\alpha_p}\left(\sqrt{\frac{T_p}{\beta_p}} - \sqrt{\frac{\beta_p}{T_p}}\right)\right].$$

Then,

$$-E\left(\frac{\partial^2 f}{\partial\alpha_i^2}\right) = \frac{1}{\alpha_i^2}\left[3\gamma^{ii} + 2\sum_{k=1,k\neq i}^{p}\gamma^{ik}\gamma_{ik} - 1\right], \quad -E\left(\frac{\partial^2 f}{\partial\alpha_i\partial\alpha_k}\right) = -\frac{1}{\alpha_i\alpha_k}\gamma_{ik}\gamma^{ik},$$

$$\begin{aligned} -E\left(\frac{\partial^2 f}{\partial\beta_i^2}\right) &= \frac{1}{\beta_i^2}\left[-\frac{1}{2} + J(\alpha_i) + \frac{1}{\alpha_i^2}\left(1 + \frac{\alpha_i^2}{2}\right)\right] \\ &\quad + \frac{1}{\beta_i^2}\left[\frac{\gamma^{ii}}{4}(5+\alpha_i^2) + \sum_{k=1,k\neq i}^{p}\frac{1}{2\alpha_i\alpha_k}\gamma^{ik}\left(\psi_1(\alpha_i,\alpha_k,\gamma_{ik}) - \psi_1(\alpha_i,\alpha_k,-\gamma_{ik})\right)\right], \end{aligned}$$

where

$$J(\alpha) = \int_{-\infty}^{\infty}(1+g(\alpha u))^{-2}d\Phi(u) \quad \text{and} \quad g(u) = 1 + \frac{1}{2}u^2 + u\left(1+\frac{u^2}{4}\right)^{\frac{1}{2}},$$

$$-E\left(\frac{\partial^2 f}{\partial\beta_i\partial\beta_k}\right) = \frac{\gamma^{ik}}{2\alpha_i\alpha_k\beta_i\beta_k}\left[\psi_1(\alpha_i,\alpha_k,\gamma_{ik}) + \gamma_1(\alpha_i,\alpha_k,-\psi_{ik})\right],$$

$$-E\left(\frac{\partial^2 f}{\partial\gamma_{ik}^2}\right) = \sum_{j_1=1}^{p}\sum_{j_2=1}^{p}a_{j_1,j_2}^{ik}\gamma_{j_1,j_2} + \frac{1}{2}c_{ik},$$

where

$$c_{ik} = \begin{cases} \frac{|\Gamma|-\gamma_{ii}^2}{|\Gamma|^2} & \text{if} \quad i=k, \\ \frac{2|\Gamma|-4\gamma_{ik}^2}{|\Gamma|^2} & \text{if} \quad i\neq k. \end{cases}$$

For $(i,k)\neq(s,t)$ or $(i,k)\neq(t,s)$, and for $i\neq k$, $s\neq t$, we have

$$-E\left(\frac{\partial^2 f}{\partial\gamma_{ik}\partial\gamma_{st}}\right) = -\frac{1}{2}\sum_{j_1=1}^{p}\sum_{j_2=1}^{p}\gamma_{j_1,j_2}c_{j_1,j_2}^{i,k,s,t} - d(i,k,s,t),$$

where $d(i,k,s,t)$ is given by

$$d(i,k,s,t) = \frac{1}{|\mathbf{\Gamma}|^2}\begin{cases} 2\gamma_{ik}\gamma_{st} & \text{if} \quad i\neq k, s\neq t \\ \gamma_{ik}\gamma_{ss} & \text{if} \quad i\neq k, s=t \\ \frac{1}{2}\gamma_{ii}\gamma_{ss} & \text{if} \quad i=k, s=t \end{cases},$$

$$-E\left(\frac{\partial^2 f}{\partial\gamma_{jk}\partial\alpha_i}\right) = -\frac{2}{\alpha_i}\sum_{m=1}^{p} b_{im}^{jk}\gamma_{im} \quad \text{and} \quad -E\left(\frac{\partial^2 f}{\partial\gamma_{jk}\partial\beta_i}\right) = 0.$$

References

[1] Aarset, M.V. (1987). "How to identify a bathtub shaped hazard rate?", *IEEE Transactions on Reliability*, vol. 36, 106–108.

[2] Anderson, T.W. and Fang, K.T. (1989). *Statistical Inference in Elliptical Contoured and Related Distributions*, Allerton Press, New York.

[3] Arellano-Valle, R.B., Contreras-Reyes, J.E. and Genton, M.G. (2012), "Shannon entropy and mutual information for multivariate skew-elliptical distributions", *Scandinavian Journal of Statistics* (to appear), DOI:10.1111/j.1467-9469.2011.00774.x.

[4] Azevedo, C., Leiva, V., Athayde, E. and Balakrishnan, N. (2012), "Shape and change point analyses of the Birnbaum-Saunders-t hazard rate and associated estimation", *Computational Statistics and Data Analysis*, vol. 56, 3887 - 3897.

[5] Birnbaum, Z.W. and Saunders, S.C. (1969a). "A new family of life distributions", *Journal of Applied Probability*, vol. 6, 319–327.

[6] Birnbaum, Z.W. and Saunders, S.C. (1969b). "Estimation for a family of life distributions with applications to fatigue", *Journal of Applied Probability*, vol. 6, 328–347.

[7] Cambanis, S., Huang, S. and Simons, G. (1981). "On the theory of elliptically contoured distribution", *Journal of Multivariate Analysis*, vol. 11, 365–385.

[8] Chhikara, R.S. and Folks, J.L. (1989). *The Inverse Gaussian Distribution*, Marcel Dekker, New York.

[9] Diaz-Garcia, J.A. and Leiva-Sanchez, V. (2005). "A new family of life distributions based on the elliptically contoured distributions", *Journal of Statistical Planning and Inference*, vol. 128, 445–457.

[10] Fang, K.T., Kotz, S. and Ng, K.W. (1990). *Symmetric Multivariate and Related Distributions*, Chapman & Hall, London, England.

[11] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions* – Vol. 2, Second edition, John Wiley & Sons, New York.

[12] Johnson, R.A. and Wichern, D.W. (1999). *Applied Multivariate Analysis*, Fourth Edition, Prentice-Hall, New Jersey.

[13] Jorgensen, B., Seshadri, V. and Whitmore, G.A. (1991). "On the mixture of the inverse Gaussian distribution with its complementary reciprocal", *Scandinavian Journal of Statistics*, vol. 18, 77–89.

[14] Karlin, S. and Rinott, Y. (1980). "Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions", *Journal of Multivariate Analysis*, vol. 10, 467–498.

[15] Kotz, S., Balakrishnan, N. and Johnson, N.L. (2000). *Continuous Multivariate Distributions* – Vol. 1, Second edition, John Wiley & Sons, New York.

[16] Kotz, S. and Nadarajah, S. (2004). *Multivariate t Distribution and Their Applications*, Cambridge University Press, Cambridge, England.

[17] Kundu, D., Balakrishnan, N. and Jamalizadeh, A. (2010). "Bivariate Birnbaum-Saunders distribution and associated inference", *Journal of Multivariate Analysis*, vol. 101, 113–125.

[18] Kundu, D., Kannan, N. and Balakrishnan, N. (2008). "On the hazard function of Birnbaum-Saunders distribution and associated inference", *Computational Statistics & Data Analysis*, vol. 52, 2692–2702.

[19] Leiva, V., Riquelme, M., Balakrishnan, N. and Sanhueza, A. (2008). "Lifetime analysis based on the generalized Birnbaum-Saunders distribution", *Computational Statistics & Data Analysis*, vol. 52, 2079–2097.

[20] Lemonte, A. J., Simas, A. B. and Cribari-Neto, F. (2008). "Bootstrap-based improved estimators for the two-parameter Birnbaum-Saunders distribution", *Journal of Statistical Computation and Simulation*, vol. 78, 37–49.

[21] Nadarajah, S. and Kotz, S. (2008). "Estimation methods for the multivariate t distribution", *Acta Applied Mathematics*, vol. 102, 99–118.

[22] Naik, D. and Plungpongpun, K. (2006). "A Kotz-type distribution for multivariate statistical inference", In *Advances in Distribution Theory, Order Statistics, and Inference* (Eds., N. Balakrishnan, E. Castillo and J.M. Sarabia), pp 111–124, Birkhäuser, Boston.

[23] Ng, H.K.T., Kundu, D. and Balakrishnan, N. (2003). "Modified moment estimation for the two-parameter Birnbaum-Saunders distribution", *Computational Statistics & Data Analysis*, vol. 43, 283–298.

[24] Rao, C.R. (1973). *Linear Statistical Inference and Its Applications*, John Wiley & Sons, New York.

[25] Sampson, A. (1983). "Positive dependence properties of elliptically symmetric distributions", *Journal of Multivariate Analysis*, vol. 13, 375-381.

[26] Vilca, F., Santana, L., Leiva, V. and Balakrishnan, N. (2011). "Estimation of extreme percentiles in Birnbaum-Saunders distribution", *Computational Statistics & Data Analysis*, vol. 55, 1665–1678.

[27] Xiao, Q., Liu, Z., Balakrishnan, N. and Lu, X. (2010). "Estimation of the Birnbaum-Saunders distribution with current status data", *Computational Statistics & Data Analysis*, vol. 54, 326–332.

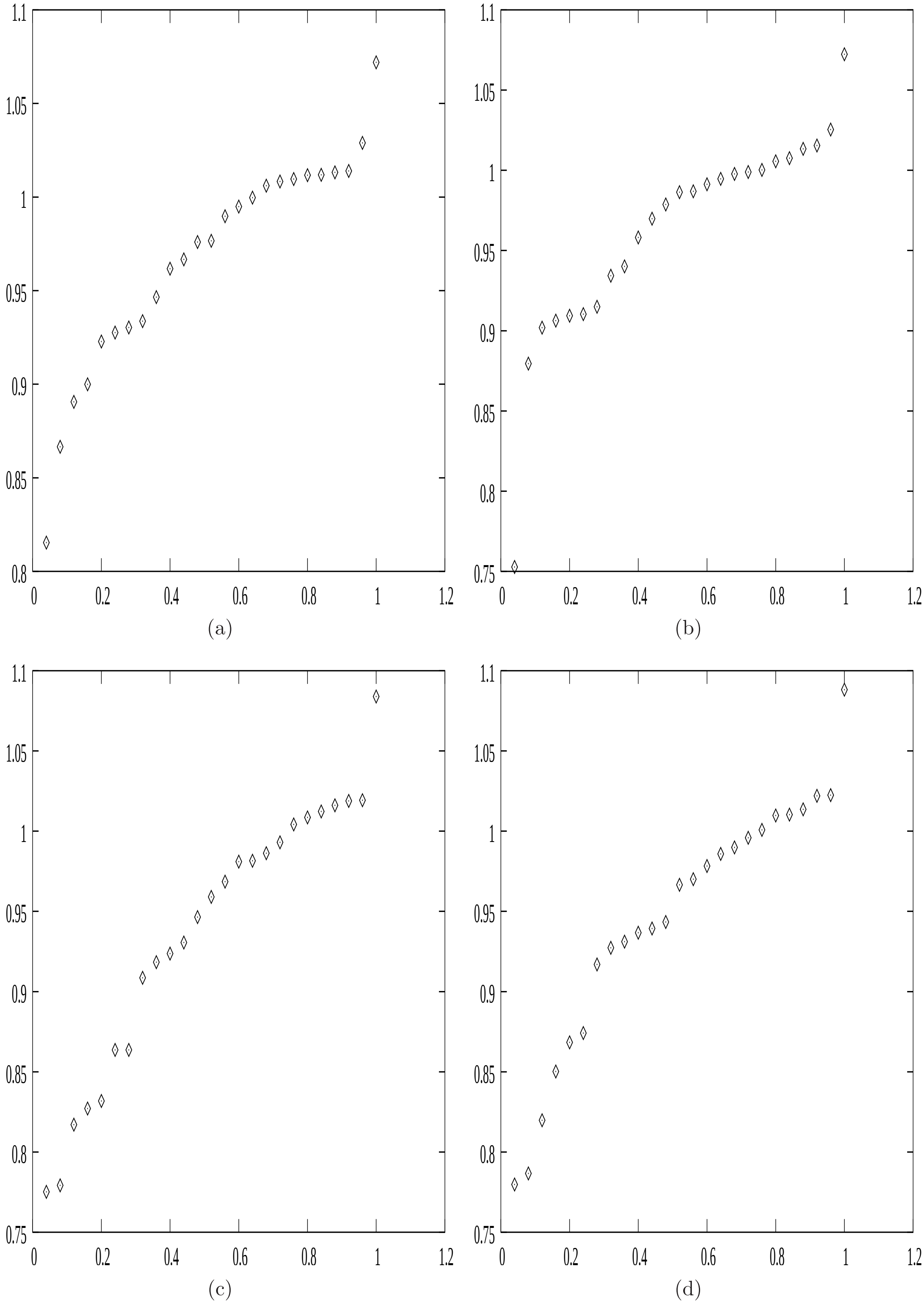

Figure 1: The marginal scaled TTT transform of (a) T_1, (b) T_2, (c) T_3, and (d) T_4.

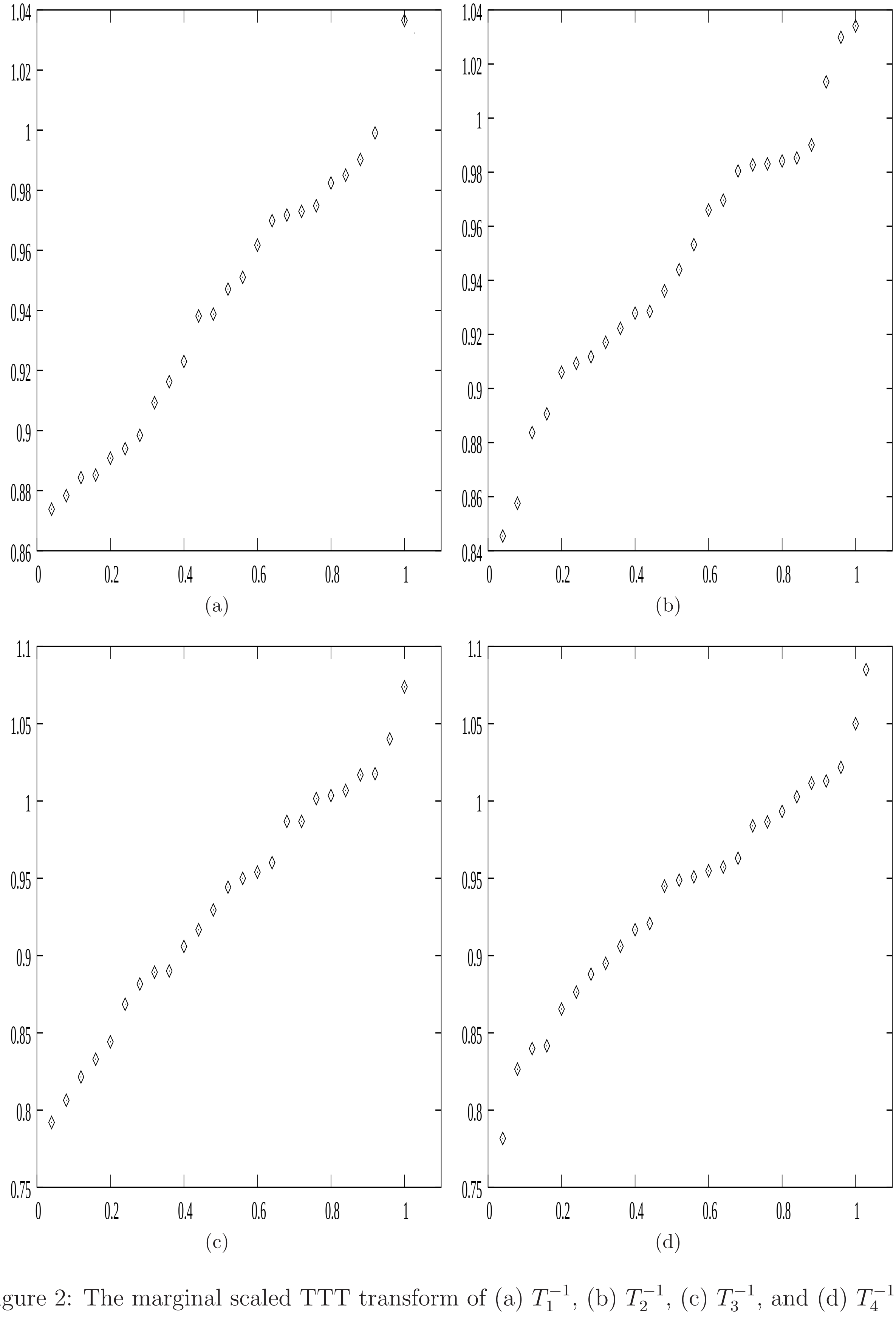

Figure 2: The marginal scaled TTT transform of (a) T_1^{-1}, (b) T_2^{-1}, (c) T_3^{-1}, and (d) T_4^{-1}.

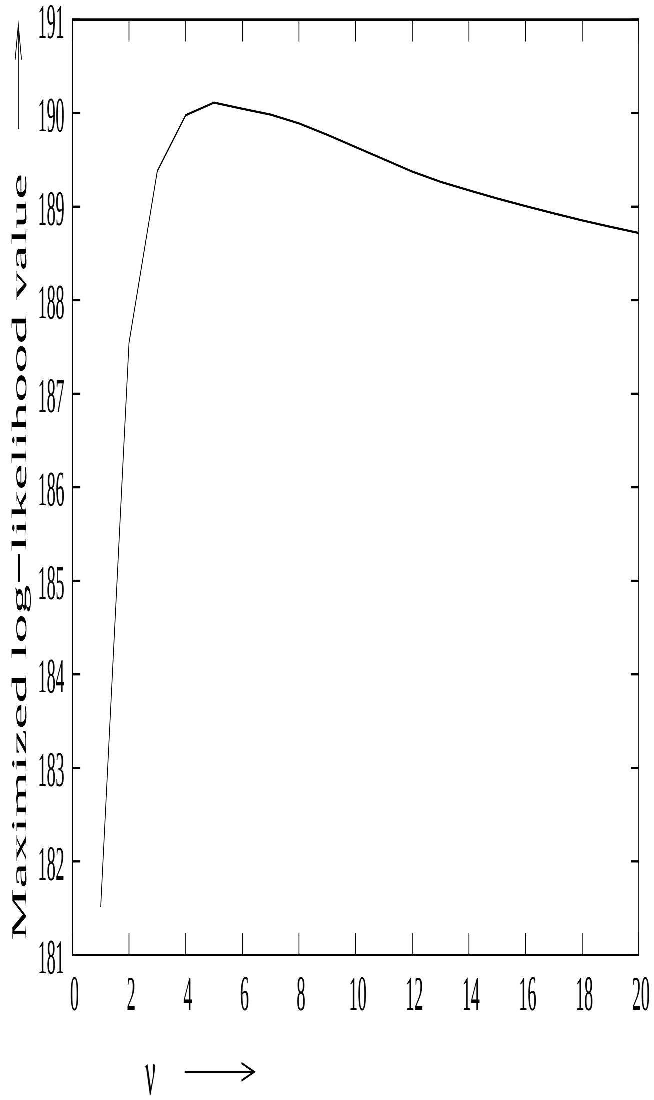

Figure 3: The maximized log-likelihood value as a profile function of $\nu = 1(1)20$.