mathematics

Article

# Restricted Distance-Type Gaussian Estimators Based on Density Power Divergence and Their Applications in Hypothesis Testing

**Ángel Felipe, María Jaenada, Pedro Miranda and Leandro Pardo ***

Department of Statistics and Operational Research, Complutense University of Madrid, 28040 Madrid, Spain
* Correspondence: lpardo@mat.ucm.es

**Abstract:** In this paper, we introduce the restricted minimum density power divergence Gaussian estimator (MDPDGE) and study its main asymptotic properties. In addition, we examine it robustness through its influence function analysis. Restricted estimators are required in many practical situations, such as testing composite null hypotheses, and we provide in this case constrained estimators to inherent restrictions of the underlying distribution. Furthermore, we derive robust Rao-type test statistics based on the MDPDGE for testing a simple null hypothesis, and we deduce explicit expressions for some main important distributions. Finally, we empirically evaluate the efficiency and robustness of the method through a simulation study.

**Keywords:** Gaussian estimator; minimum density power divergence Gaussian estimator; robustness; influence function; Rao-type tests; elliptical family of distributions

**MSC:** 62F30

check for updates

**Citation:** Felipe, Á.; Jaenada, M.; Miranda, P.; Pardo, L. Restricted Distance-Type Gaussian Estimators Based on Density Power Divergence and Their Applications in Hypothesis Testing. *Mathematics* **2023**, *11*, 1480. https://doi.org/10.3390/math11061480

Academic Editor: Zhonghua Li

Received: 15 February 2023
Revised: 10 March 2023
Accepted: 15 March 2023
Published: 17 March 2023



## 1. Introduction

Let $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n$ be independent and identically distributed observations from an $m$-dimensional random vector $\boldsymbol{Y}$ with probability density function $f_{\boldsymbol{\theta}}(\boldsymbol{y})$, $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$, where $\boldsymbol{\theta}$ is the vector of unknown parameters, $\boldsymbol{\Theta}$ is the corresponding parameter space and $d$ is the dimension of $\boldsymbol{\theta}, d \geq 1$. We denote,

$$E_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\mu}(\boldsymbol{\theta}) \text{ and } Cov_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\Sigma}(\boldsymbol{\theta}). \tag{1}$$

The log-likelihood function of the assumed model is given by

$$l(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log f_{\boldsymbol{\theta}}(\boldsymbol{y}_i)$$

for $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ observations of the $m$-dimensional random vectors $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$. Then, the maximum likelihood estimator (MLE) is computed as

$$\widehat{\boldsymbol{\theta}}_{\text{MLE}} = \max_{\boldsymbol{\theta} \in \Theta} l(\boldsymbol{\theta}). \tag{2}$$

In many real-life situations, the underlying density function, $f_\theta(\cdot)$, is unknown or its computation is quite difficult but contrariwise, the mean vector and variance–covariance matrices of the underlying distribution of the data as a function of $\boldsymbol{\theta}$, namely $\boldsymbol{\mu}(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}(\boldsymbol{\theta})$, are known.

In this case, Zhang [1] proposed a general procedure based on the Gaussian distribution for estimating the model parameter vector $\boldsymbol{\theta}$. In [1], it is assumed that the $m$-dimensional random vector $\boldsymbol{Y}$ comes from a multidimensional normal distribution with mean vector $\boldsymbol{\mu}(\boldsymbol{\theta})$ and variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})$. From a statistical point of view, this procedure can be justified on the basis of the maximum-entropy principle (see [2]), as the multidimensional normal distribution has maximum uncertainty in terms of Shannon entropy and is as well consistent with the given information, i.e., vector mean and variance–covariance matrix.

Then, an estimator of the model parameter $\boldsymbol{\theta}$ based on the Gaussian distribution can be obtained by maximizing the log-likelihood function as defined in (2) but using as $f_{\boldsymbol{\theta}}(\cdot)$ the probability density function of a normal distribution with known mean vector $\boldsymbol{\mu}(\boldsymbol{\theta})$ and variance–covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})$, corresponding to the true mean vector and variance–covariance matrix of the underlying distribution. That is, the Gaussian-based likelihood function of $\boldsymbol{\theta}$ is given by

$$l_G(\boldsymbol{\theta}) = -\frac{nm}{2}\log 2\pi - \frac{n}{2}\log|\boldsymbol{\Sigma}(\boldsymbol{\theta})| - \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta})) \tag{3}$$

for any $\boldsymbol{y}_1, ..., \boldsymbol{y}_n$ independent observations of the population $\boldsymbol{Y}$, and the Gaussian MLE of $\boldsymbol{\theta}$ is defined by

$$\widehat{\boldsymbol{\theta}}_G = \arg\max_{\boldsymbol{\theta}\in\Theta} l_G(\boldsymbol{\theta}).$$

The Gaussian estimator is an MLE and thus inherits all the good properties of the likelihood estimators. Consequently, it works well in terms of the asymptotic efficiency, but it has important robustness problems. That is, in the absence of contamination in data, the MLE consistently estimates the true value of the model parameter, but it may be quite heavily affected by outlying observations in the data. For this reason, Castilla and Zografos extended in [3] the concept of Gaussian estimator and defined a robust version of the estimator based on the density power divergence (DPD) introduced in Basu et al. in [4]. The DPD robustly quantifies the statistical difference between two distributions, and it has been widely used for developing robust inferential methods in many different statistical models. Given a set of observations, the robust minimum DPD estimator (MDPDE) is computed as the minimizer of the DPD between the assumed model distribution and the empirical distribution of the data. The MDPDE enjoys good asymptotic properties and produces robust estimators under general statistical models, as discussed later. The minimum density power divergence Gaussian estimator (MDPDGE) of the parameter $\boldsymbol{\theta}$ is defined for $\tau \geq 0$ as

$$\widehat{\boldsymbol{\theta}}_G^{\tau} = \arg\max_{\boldsymbol{\theta}\in\Theta\subset\mathbb{R}^d} H_n^{\tau}(\boldsymbol{\theta}), \tag{4}$$

where

$$\begin{aligned} H_n^{\tau}(\boldsymbol{\theta}) =& \frac{\tau+1}{\tau(2\pi)^{m\tau/2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{\tau/2}}\frac{1}{n}\left[\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\right. \\ & \left. - \frac{\tau}{(1+\tau)^{(m/2)+1}}\right] - \frac{1}{\tau} \\ =& a|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}\frac{1}{n}\left[\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))\right\} - b\right] - \frac{1}{\tau}, \end{aligned} \tag{5}$$

and

$$a = \frac{\tau+1}{\tau(2\pi)^{m\tau/2}} \text{ and } b = \frac{\tau}{(1+\tau)^{(m/2)+1}}. \tag{6}$$

The MDPDGE family is indexed by a tuning parameter $\tau$ controlling the trade-off between robustness and efficiency; the greater the value of $\tau$, the more robust the resulting estimator is, but the efficiency decreases. It has been shown in the literature that values of the tuning parameter above 1 do not provide sufficiently efficient estimators, and so, the tuning parameter would be chosen in the $[0,1]$ interval. Furthermore, at $\tau = 0$, the MDPDGE reduces to the Gaussian estimator of [1],

$$\widehat{\boldsymbol{\theta}}_G = \arg \max_{\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d} H_n^0(\boldsymbol{\theta})$$

with

$$H_n^0(\boldsymbol{\theta}) = \lim_{\tau \to 0} H_n^\tau(\boldsymbol{\theta}) = -\frac{n}{2} \log|\boldsymbol{\Sigma}(\boldsymbol{\theta})| - \frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} (\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta})). \tag{7}$$

Note that the above objective function does not perfectly match with the likelihood function of the model stated in (2), as it lacks the first term of the likelihood. However, this term does not depend on the parameter $\boldsymbol{\theta}$, and thus, both loss functions will lead to the same maximizer. Indeed, the loss in Equation (7) corresponds to the Kullback–Leiber divergence between the assumed normal distribution and the empirical distribution of the data, which justifies the MLE from the point of view of information theory (see [5–7]).

Furthermore, the MDPDGE is consistent and asymptotically normal, that is, given $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n$ independent and identically distributed vectors from the $m$-dimensional random vector $\boldsymbol{Y}$, the MDPDGE, $\widehat{\boldsymbol{\theta}}_G^\tau$, defined in (4) satisfies

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_G^\tau - \boldsymbol{\theta}\right) \underset{n \longrightarrow \infty}{\xrightarrow{\mathcal{L}}} \mathcal{N}(\boldsymbol{0}_d, \boldsymbol{J}_\tau(\boldsymbol{\theta})^{-1} \boldsymbol{K}_\tau(\boldsymbol{\theta}) \boldsymbol{J}_\tau(\boldsymbol{\theta})^{-1}), \tag{8}$$

being

$$\boldsymbol{J}_\tau(\boldsymbol{\theta}) = \left(J_\tau^{ij}(\boldsymbol{\theta})\right)_{i,j=1,\dots,d} \text{ and } \boldsymbol{K}_\tau(\boldsymbol{\theta}) = \left(K_\tau^{ij}(\boldsymbol{\theta})\right)_{i,j=1,\dots,d}.$$

and the elements $J_\tau^{ij}(\boldsymbol{\theta})$ and $K_\tau^{ij}(\boldsymbol{\theta})$ of the matrices $J_\tau(\boldsymbol{\theta})$ and $K_\tau(\boldsymbol{\theta})$ are given by

$$\begin{aligned} J_\tau^{ij}(\boldsymbol{\theta}) &= \left(\frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{1/2}}\right)^\tau \frac{1}{(1+\tau)^{(m/2)+2}} \\ &\times \left[(\tau+1) trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i} \left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j}\right)^T\right)\right. \\ &\left. + \Delta_\tau^i \Delta_\tau^j + \frac{1}{2} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j} \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)\right] \end{aligned} \tag{9}$$

and

$$\begin{aligned} K_\tau^{ij}(\boldsymbol{\theta}) &= \left(\frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{1/2}}\right)^{2\tau} \frac{1}{(1+2\tau)^{(m/2)+2}} \left[\Delta_{2\tau}^i \Delta_{2\tau}^j\right. \\ &+ (1+2\tau) trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i} \left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j}\right)^T\right) \\ &\left. + \frac{1}{2} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j} \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)\right] \\ &- \left(\frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{1/2}}\right)^{2\tau} \frac{1}{(1+\tau)^{m+2}} \Delta_\tau^i \Delta_\tau^j, \end{aligned} \tag{10}$$

with $\Delta_\tau^i = \frac{\tau}{2} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)$. By $trace(A)$, we mean the trace of matrix $A$.

The above asymptotic distribution follows from Theorem 2 in [8], where matrices $\boldsymbol{J}_\tau(\boldsymbol{\theta})$ and $\boldsymbol{K}_\tau(\boldsymbol{\theta})$ are defined for general statistical models.

The asymptotic distribution of $\widehat{\boldsymbol{\theta}_G}$ has been considered in many papers, see e.g., [9–13].

On the other hand, in some situations, we may have additional knowledge about the true parameter space. Then, these constraints should be included in the definition of the parameter space $\boldsymbol{\Theta}$. Here, we will consider restricted parameter spaces of the form

$$\Theta_0 = \{\boldsymbol{\theta} \in \Theta / \ \boldsymbol{g}(\boldsymbol{\theta}) = \mathbf{0}_r\}, \tag{11}$$

where $\boldsymbol{g} : \mathbb{R}^d \to \mathbb{R}^r$ is a vector-valued function mapping such that the $d \times r$ matrix

$$\boldsymbol{G}(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{g}^T(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \tag{12}$$

exists and is continuous in $\boldsymbol{\theta}$ and $\mathrm{rank}(\boldsymbol{G}(\boldsymbol{\theta})) = r$, and $\mathbf{0}_r$ denotes the null vector of dimension $r$. The notation $\Theta_0$ clues the use of the present restricted estimator for defining test statistics under composite null hypothesis.

The most popular estimator of $\boldsymbol{\theta}$ satisfying the constraints in (11) is the restricted MLE (RMLE), which is naturally defined as the maximizer of the log-likelihood function of the model but is subject to the parameter space restrictions $\boldsymbol{g}(\boldsymbol{\theta}) = \mathbf{0}_r$ (see [14–17]). Unfortunately, the RMLE has the same robustness problems as the MLE, and so robust alternatives should be adopted in the presence of contamination in data. Several robust restricted estimators have been considered in the statistical literature to overcome the robustness drawback of the RMLE. For example, Pardo et al. introduced in [18] the restricted minimum Phi-divergence estimator and studied its properties. In [8], the restricted minimum density power divergence estimators (RMDPDE) are presented, and some applications on the testing hypothesis are studied. In [19], the theoretical robustness properties of the RMDPDE were studied. In [20,21], the restricted Rényi pseudodistance estimator is considered, and robust Rao-type tests are derived from it and developed. More recently, in [22], the RMDPD under normal distributions is studied, and independence tests under the normal assumption are developed, and in [23] the RMDPDE is applied in the context of independent but not identically distributed variables under heterocedastic linear regression models. Other interesting papers related to multivariate analysis are [24–27].

In this paper, we introduce and study the restricted minimum density power divergence Gaussian estimator (RMDPDGE). The aim of this study is to introduce an estimator dealing with situations where Gaussian estimators are useful; there are additional constraints on the parameter space, and the estimator should be robust in terms of contamination. To show the robustness, we compute the corresponding influence function (IF), showing that in general, it is bounded. As an application of these restricted estimators, we develop Rao-type test statistics. The idea in this case is to consider the additional constraints as the constraints defining the null hypothesis of the test. Finally, we show the robustness via a simulation study, in which the good behavior of these estimators is shown comparing it with the behavior of classical restricted estimators based on MLE. We compare the loss in efficiency when comparing these estimators with more specific estimators dealing with the *real* distribution, showing that the loss is affordable.

The rest of the paper is organized as follows: In Section 2, we introduce the RMDPDGE and we obtain its asymptotic distribution. Section 3 presents the influence function of the RMDPDGE and theoretically proves the robustness of the proposed estimators. Some statistical applications for testing are presented in Section 4, and an explicit expression of the Rao-type test statistics based on the RMDPDGE under exponential, Poisson and Lindley models are given. Section 5 empirically demonstrates the robustness of the method through a simulation study, and the advantages and disadvantages of the Gaussian assumption are discussed there. Section 6 presents some conclusions. The proofs of the main results stated in the paper are included in Appendix A.

## 2. Restricted Minimum Density Power Divergence Gaussian Estimators

In this section, we present the RMDPDGE under general equality non-linear constraints and we study its asymptotic distribution, showing the consistency of the estimator.

**Definition 1.** *Let $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n$ be independent and identically distributed observations from an m-dimensional random vector $\boldsymbol{Y}$ with $E_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\mu}(\boldsymbol{\theta})$ and $Cov_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\Sigma}(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$. The RMDPDGE, $\widetilde{\boldsymbol{\theta}}_G^{\tau}$, is defined by*

$$\widetilde{\boldsymbol{\theta}}_G^{\tau} = \arg\max_{\Theta_0} H_n^{\tau}(\boldsymbol{\theta}),$$

*where $H_n^{\tau}(\boldsymbol{\theta})$ is as given in Equation (5) and $\Theta_0 = \{\boldsymbol{\theta} \in \Theta / \ \boldsymbol{g}(\boldsymbol{\theta}_r) = \boldsymbol{0}_r\}$ is the restricted parameter space defined in (11).*

Before presenting the asymptotic distribution of the MDPDGE, we present some previous results whose proofs are included in Appendix A.

**Proposition 1.** *Let $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n$ be independent and identically distributed observations from an m-dimensional random vector $\boldsymbol{Y}$ with $E_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\mu}(\boldsymbol{\theta})$ and $Cov_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\Sigma}(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$. Then,*

$$\sqrt{n}\left(\frac{1}{\tau+1}\frac{\partial H_n^{\tau}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right) \underset{n\longrightarrow\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(\boldsymbol{0}_d, \boldsymbol{K}_{\boldsymbol{\tau}}(\boldsymbol{\theta})),$$

*where $\boldsymbol{K}_{\boldsymbol{\tau}}(\boldsymbol{\theta})$ was defined in (10).*

**Proof.** See Appendix A.2. □

**Proposition 2.** *Let $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n$ be independent and identically distributed observations from an m-dimensional random vector $\boldsymbol{Y}$ with $E_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\mu}(\boldsymbol{\theta})$ and $Cov_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\Sigma}(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$. Then,*

$$\frac{\partial^2 H_n^{\tau}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} -(\tau+1)\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta}),$$

*where $\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta})$ was defined in (9).*

**Proof.** See Appendix A.3. □

Next, we present the asymptotic distribution of $\widetilde{\boldsymbol{\theta}}_G^{\tau}$.

**Theorem 1.** *Let $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n$ be independent and identically distributed observations from an m-dimensional random vector $\boldsymbol{Y}$ with $E_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\mu}(\boldsymbol{\theta})$ and $Cov_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\Sigma}(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$. Suppose the true distribution of $\boldsymbol{Y}$ belongs to the model, and we consider $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$. Then, the RMDPDGE $\widetilde{\boldsymbol{\theta}}_G^{\tau}$ of $\boldsymbol{\theta}$ obtained under the constraints $\boldsymbol{g}(\boldsymbol{\theta}) = \boldsymbol{0}_r$ satisfies*

$$n^{1/2}(\widetilde{\boldsymbol{\theta}}_G^{\tau} - \boldsymbol{\theta}) \underset{n\longrightarrow\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(\boldsymbol{0}_d, \boldsymbol{M}_{\boldsymbol{\tau}}(\boldsymbol{\theta})), \tag{13}$$

*where*

$$\boldsymbol{M}_{\boldsymbol{\tau}}(\boldsymbol{\theta}) = \boldsymbol{P}_{\boldsymbol{\tau}}^*(\boldsymbol{\theta})\boldsymbol{K}_{\boldsymbol{\tau}}(\boldsymbol{\theta})\boldsymbol{P}_{\boldsymbol{\tau}}^*(\boldsymbol{\theta})^T,$$

$$\boldsymbol{P}_{\boldsymbol{\tau}}^*(\boldsymbol{\theta}) = \boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta})^{-1} - \boldsymbol{Q}_{\boldsymbol{\tau}}(\boldsymbol{\theta})\boldsymbol{G}(\boldsymbol{\theta})^T\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta})^{-1}, \tag{14}$$

$$\boldsymbol{Q}_{\boldsymbol{\tau}}(\boldsymbol{\theta}) = \boldsymbol{J}_{\boldsymbol{\tau}}^{-1}(\boldsymbol{\theta})\boldsymbol{G}(\boldsymbol{\theta})\left[\boldsymbol{G}(\boldsymbol{\theta})^T\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta})^{-1}\boldsymbol{G}(\boldsymbol{\theta})\right]^{-1}, \tag{15}$$

*and $\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta})$ and $\boldsymbol{K}_{\boldsymbol{\tau}}(\boldsymbol{\theta})$ were defined in (9) and (10), respectively.*

**Proof.** The estimating equations for the RMDPDGE are given by

$$\begin{cases} \frac{\partial}{\partial \boldsymbol{\theta}} H_n^{\tau}(\boldsymbol{\theta}) + \boldsymbol{G}(\boldsymbol{\theta})\boldsymbol{\lambda}_n = \boldsymbol{0}_d, \\ \boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_G^{\tau}) = \boldsymbol{0}_r, \end{cases} \tag{16}$$

where $\boldsymbol{\lambda}_n$ is a vector of Lagrangian multipliers. Now, we consider $\boldsymbol{\theta}_n = \boldsymbol{\theta} + \boldsymbol{m}n^{-1/2}$, where $||\boldsymbol{m}|| < k$, for $0 < k < \infty$. We have,

$$\frac{\partial}{\partial \boldsymbol{\theta}} H_n^\tau(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_n} = \frac{\partial}{\partial \boldsymbol{\theta}} H_n^\tau(\boldsymbol{\theta}) + \frac{\partial^2}{\partial \boldsymbol{\theta}^T \partial \boldsymbol{\theta}} H_n^\tau(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_*}(\boldsymbol{\theta}_n^* - \boldsymbol{\theta})$$

and

$$n^{1/2} \frac{\partial}{\partial \boldsymbol{\theta}} H_n^\tau(\boldsymbol{\theta})\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_n} = n^{1/2} \frac{\partial}{\partial \boldsymbol{\theta}} H_n^\tau(\boldsymbol{\theta}) + \frac{\partial^2}{\partial \boldsymbol{\theta}^T \partial \boldsymbol{\theta}} H_n^\tau(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_*} n^{1/2}(\boldsymbol{\theta}_n - \boldsymbol{\theta}), \tag{17}$$

where $\boldsymbol{\theta}^*$ belongs to the segment joining $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_n$. Since

$$\lim_{n \to \infty} \frac{\partial^2}{\partial \boldsymbol{\theta}^T \partial \boldsymbol{\theta}} H_n^\tau(\boldsymbol{\theta}) = -(\tau+1)\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta}),$$

we obtain

$$n^{1/2} \frac{\partial}{\partial \boldsymbol{\theta}} H_n^\tau(\boldsymbol{\theta})\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_n} = n^{1/2} \frac{\partial}{\partial \boldsymbol{\theta}} H_n^\tau(\boldsymbol{\theta}) - (\tau+1)n^{1/2}\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta})(\boldsymbol{\theta}_n - \boldsymbol{\theta}) + o_p(1). \tag{18}$$

Taking into account that $\boldsymbol{G}(\boldsymbol{\theta})$ is continuous in $\boldsymbol{\theta}$

$$n^{1/2}\boldsymbol{g}(\boldsymbol{\theta}_n) = \boldsymbol{G}(\boldsymbol{\theta})^T n^{1/2}(\boldsymbol{\theta}_n - \boldsymbol{\theta}) + o_p(1). \tag{19}$$

The RMDPDGE $\widetilde{\boldsymbol{\theta}}_G^\tau$ must satisfy the conditions in (16), and in view of (18), we have

$$n^{1/2} \frac{\partial}{\partial \boldsymbol{\theta}} H_n^\tau(\boldsymbol{\theta}) - (\tau+1)\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta})n^{1/2}(\widetilde{\boldsymbol{\theta}}_G^\tau - \boldsymbol{\theta}) + \boldsymbol{G}(\boldsymbol{\theta})n^{1/2}\boldsymbol{\lambda}_n + o_p(1) = \boldsymbol{0}_p. \tag{20}$$

From (19), it follows that

$$\boldsymbol{G}^T(\boldsymbol{\theta})n^{1/2}(\widetilde{\boldsymbol{\theta}}_G^\tau - \boldsymbol{\theta}) + o_p(1) = \boldsymbol{0}_r. \tag{21}$$

Now, we can express Equations (20) and (21) in matrix form as

$$\begin{pmatrix} (\tau+1)\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta}) & -\boldsymbol{G}(\boldsymbol{\theta}) \\ -\boldsymbol{G}^T(\boldsymbol{\theta}) & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} n^{1/2}(\widetilde{\boldsymbol{\theta}}_G^\tau - \boldsymbol{\theta}) \\ n^{1/2}\boldsymbol{\lambda}_n \end{pmatrix} = \begin{pmatrix} n^{1/2}\frac{\partial}{\partial \boldsymbol{\theta}} H_n^\tau(\boldsymbol{\theta}) \\ \boldsymbol{0} \end{pmatrix} + o_p(1).$$

Therefore,

$$\begin{pmatrix} n^{1/2}(\widetilde{\boldsymbol{\theta}}_G^\tau - \boldsymbol{\theta}) \\ n^{1/2}\boldsymbol{\lambda}_n \end{pmatrix} = \begin{pmatrix} (\tau+1)\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta}) & -\boldsymbol{G}(\boldsymbol{\theta}) \\ -\boldsymbol{G}^T(\boldsymbol{\theta}) & \boldsymbol{0} \end{pmatrix}^{-1} \begin{pmatrix} n^{1/2}\frac{\partial}{\partial \boldsymbol{\theta}} H_n^\tau(\boldsymbol{\theta}) \\ \boldsymbol{0}_r \end{pmatrix} + o_p(1).$$

However,

$$\begin{pmatrix} (\tau+1)\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta}) & -\boldsymbol{G}(\boldsymbol{\theta}) \\ -\boldsymbol{G}^T(\boldsymbol{\theta}) & \boldsymbol{0} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{L}_\tau^*(\boldsymbol{\theta}) & \boldsymbol{Q}_\tau(\boldsymbol{\theta}) \\ \boldsymbol{Q}_\tau(\boldsymbol{\theta}_0)^T & \boldsymbol{R}_\tau(\boldsymbol{\theta}) \end{pmatrix},$$

where

$$\begin{aligned} \boldsymbol{L}_\tau^*(\boldsymbol{\theta}) &= \frac{1}{\tau+1}\Big(\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta})^{-1} - \boldsymbol{Q}_{\boldsymbol{\tau}}(\boldsymbol{\theta})\boldsymbol{G}(\boldsymbol{\theta})^T\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta})^{-1}\Big) \\ &= \frac{1}{\tau+1}\boldsymbol{P}_\tau^*(\boldsymbol{\theta}), \\ \boldsymbol{Q}_\tau(\boldsymbol{\theta}) &= \boldsymbol{J}_{\boldsymbol{\tau}}^{-1}(\boldsymbol{\theta})\boldsymbol{G}(\boldsymbol{\theta})\Big[\boldsymbol{G}(\boldsymbol{\theta})^T\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta})^{-1}\boldsymbol{G}(\boldsymbol{\theta})\Big]^{-1}, \\ \boldsymbol{R}_\tau(\boldsymbol{\theta}) &= \boldsymbol{G}(\boldsymbol{\theta})^T\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta})^{-1}\boldsymbol{G}(\boldsymbol{\theta}), \end{aligned}$$

and $\boldsymbol{P}_\tau^*(\boldsymbol{\theta}_0)$ and $\boldsymbol{Q}_\tau(\boldsymbol{\theta}_0)$ are as given in (14) and (15), respectively. Then,

$$n^{1/2}(\widetilde{\boldsymbol{\theta}}_G^\tau - \boldsymbol{\theta}) = (\tau+1)^{-1}\boldsymbol{P}_\tau^*(\boldsymbol{\theta})n^{1/2}\frac{\partial}{\partial\boldsymbol{\theta}}H_n^\tau(\boldsymbol{\theta}) + o_p(1), \tag{22}$$

and we know by Proposition 1 that

$$n^{1/2}(\tau+1)^{-1}\frac{\partial}{\partial\boldsymbol{\theta}}H_n^\tau(\boldsymbol{\theta}) \underset{n\longrightarrow\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(\mathbf{0}, \boldsymbol{K}_{\boldsymbol{\tau}}(\boldsymbol{\theta})). \tag{23}$$

Now, by (22) and (23), we have the desired result. □

**Remark 1.** *Notice that the result in (8) is a special case of the previous theorem when there are no restrictions on the parameter space, in the sense that* $\boldsymbol{G}$ *defined in (12) is the null matrix. In this case, matrix* $\boldsymbol{P}_{\boldsymbol{\tau}}^*(\boldsymbol{\theta})$ *given in (14) becomes* $\boldsymbol{P}_{\boldsymbol{\tau}}^*(\boldsymbol{\theta}) = \boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta})^{-1}$*. Therefore, the asymptotic variance–covariance matrix of the unrestricted estimator, i.e., the MDPDGE, may be reconstructed from the previous theorem.*

In order to compute the MDPDGE, we note that it is an optimum of a differentiable function $H_n^\tau$, so it must annul its first derivatives. We will use that

$$\frac{\partial|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}}{\partial\boldsymbol{\theta}} = -\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right),$$

and

$$\begin{aligned}\frac{\partial}{\partial\boldsymbol{\theta}}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta})) = & -2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta})) - (\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T \\ & \times\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta})).\end{aligned}$$

Now, taking derivatives in (5), for a certain fixed $\tau$, we have that

$$\begin{aligned}\frac{\partial}{\partial\boldsymbol{\theta}}H_n^\tau(\boldsymbol{\theta}) = & \frac{1}{n}\sum_{i=1}^n\Bigg\{-a\,\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right) \\ & \times\exp\left(-\frac{\tau}{2}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))\right) + ba\,\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2} \\ & \times trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right) + a\,\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2} \\ & \times\exp\left(-\frac{\tau}{2}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))\right) \\ & \times\Bigg[2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta})) \\ & +(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))\Bigg]\Bigg\} \\ = & \frac{1}{n}\sum_{i=1}^n\boldsymbol{\Psi}_{\boldsymbol{\tau}}(\boldsymbol{y}_i;\boldsymbol{\theta}),\end{aligned}$$

with

$$\begin{aligned}\mathbf{\Psi}_{\tau}(\boldsymbol{y}_i;\boldsymbol{\theta}) &= a\,\frac{\tau}{2}|\mathbf{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}\Bigg\{\Bigg[-trace\left(\mathbf{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\mathbf{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right) \\ &+\left(2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)^T\mathbf{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right. \\ &\left.+(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\mathbf{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\mathbf{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right)\Bigg] \\ &\times\exp\left(-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\mathbf{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right) \\ &+b\;trace\left(\mathbf{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\mathbf{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)\Bigg\}.\end{aligned} \tag{24}$$

Therefore, the estimating equations of the MDPDGE for a fixed parameter $\tau$ are given by

$$\sum_{i=1}^{n}\mathbf{\Psi}_{\tau}(\boldsymbol{y}_i;\boldsymbol{\theta}) = \mathbf{0}_d. \tag{25}$$

The previous estimating equations characterize the MDPDGE as an M-estimator and so its asymptotic distribution could have been also derived from the general theory of M-estimators. In particular, the MDPDGE, $\widehat{\boldsymbol{\theta}}_G^{\tau}$, satisfies for any $\tau \geq 0$

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_G^{\tau}-\boldsymbol{\theta}\right)\underset{n\longrightarrow\infty}{\overset{\mathcal{L}}{\longrightarrow}}\mathcal{N}\left(\mathbf{0}_d,\boldsymbol{S}^{-1}\boldsymbol{M}\boldsymbol{S}^{-1}\right), \tag{26}$$

with

$$\boldsymbol{S} = -E\left[\frac{\partial^2 H_n^{\tau}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right] \text{ and } \boldsymbol{M} = Cov\left[\sqrt{n}\frac{\partial}{\partial\boldsymbol{\theta}}H_n^{\tau}(\boldsymbol{\theta})\right]. \tag{27}$$

Based on Propositions 1 and 2, we can express the previous matrices as

$$\boldsymbol{S} = (\tau+1)\boldsymbol{J}_{\boldsymbol{\tau}}(\boldsymbol{\theta}) \text{ and } \boldsymbol{M} = (\tau+1)^2\boldsymbol{K}_{\boldsymbol{\tau}}(\boldsymbol{\theta}),$$

and we obtain the expressions established in (8). The asymptotic convergence in (26) offers an alternative proof of the asymptotic distribution of MDPDGE developed in [3] in terms of the transformed matrices $\boldsymbol{S}$ and $\boldsymbol{M}$ in Equation (27).

## 3. Influence Function for the RMDPDGE

To analyze the robustness of an estimator, Hampel et al. introduced in [28] the concept of an Influence Function (IF). Since then, the IF has been widely used in the statistical literature to measure robustness in different statistical contexts. Intuitively, the IF describes the effect of an infinitesimal contamination of the model on the estimation. Robust estimators should be less affected by contamination, and thus, IFs associated to locally robust (B-robust) estimators should be bounded.

The IF of an estimator, $\widetilde{\boldsymbol{\theta}}_G^{\tau}$, is defined in terms of its statistical functional $\widetilde{T}_{\tau}$ satisfying $\widetilde{T}_{\tau}(g) = \widetilde{\boldsymbol{\theta}}_G^{\tau}$, where $g$ is the true density function underlying the data. Given the density function $g$, we define its contaminated version at the point perturbation $\boldsymbol{y}_0$ as,

$$g = (1-\varepsilon)g + \varepsilon\Delta_{\boldsymbol{y}_0}, \tag{28}$$

where $\varepsilon$ is fraction of contamination and $\Delta_{\boldsymbol{y}_0}$ denotes the indicator function at $\boldsymbol{y}_0$. Then, the IF of $\widetilde{\theta}_G^{\tau}$ is defined as the derivative of the functional at $\varepsilon = 0$

$$IF(\boldsymbol{y}_0,\widetilde{T}_{\tau}) = \frac{\partial\widetilde{T}_{\tau}(g_{\varepsilon})}{\varepsilon}|_{\varepsilon=0}.$$

Hence, the above derivative quantifies the rate of change of the sample estimator when contamination occurs.

Let us now obtain the IF of RMDPDE. We consider the contaminated model

$$g_{\varepsilon}(\boldsymbol{y}) = (1-\varepsilon) f_{\boldsymbol{\theta}}(\boldsymbol{y}) + \varepsilon \Delta_{\boldsymbol{y}_0},$$

where $f_{\boldsymbol{\theta}}$ is the assumed probability density function of a normal population. The MDPDGE for the contaminated model is then given by $\widetilde{\boldsymbol{\theta}}^{\tau}_{G,\varepsilon} = \widetilde{T}_{\tau}(g_{\varepsilon})$.

By definition, $\widetilde{\boldsymbol{\theta}}^{\tau}_{G,\varepsilon}$ is the maximizer of $H^{\tau}_{n}(\boldsymbol{\theta})$ in (5), subject to the constraints $\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}^{\tau}_{G,\varepsilon}) = \mathbf{0}$. Using the characterization of the MDPDGE as an M-estimator, we have that the influence function of the MDPDGE is given by

$$IF(\boldsymbol{y}, \widetilde{T}_{\tau}, \boldsymbol{\theta}) = \boldsymbol{J}_{\tau}(\boldsymbol{\theta})^{-1} \boldsymbol{\Psi}_{\tau}(\boldsymbol{y}; \boldsymbol{\theta}), \tag{29}$$

where $\boldsymbol{J}_{\tau}(\boldsymbol{\theta})$ was defined in (9) and $\boldsymbol{\Psi}_{\tau}(\boldsymbol{y}; \boldsymbol{\theta})$ in (24). The influence function of the RMDPDGE will be obtained with the additional condition $\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}^{\tau}_{G,\varepsilon}) = \mathbf{0}$. Differentiating this last equation gives, at $\varepsilon = 0$,

$$\boldsymbol{G}(\boldsymbol{\theta})^{T} IF(\boldsymbol{y}, \widetilde{T}_{\tau}, \boldsymbol{\theta}) = \mathbf{0}. \tag{30}$$

Based on (29) and (30), we have

$$\begin{pmatrix} \boldsymbol{J}_{\tau}(\boldsymbol{\theta}) \\ \boldsymbol{G}(\boldsymbol{\theta})^{T} \end{pmatrix} IF(\boldsymbol{y}, \widetilde{T}_{\tau}, \boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{\Psi}_{\tau}(\boldsymbol{y}; \boldsymbol{\theta}) \\ \mathbf{0} \end{pmatrix}.$$

Therefore,

$$\left( \boldsymbol{J}_{\tau}(\boldsymbol{\theta})^{T}, \boldsymbol{G}(\boldsymbol{\theta}) \right) \begin{pmatrix} \boldsymbol{J}_{\tau}(\boldsymbol{\theta}) \\ \boldsymbol{G}(\boldsymbol{\theta})^{T} \end{pmatrix} IF(\boldsymbol{y}, \widetilde{T}_{\tau}, \boldsymbol{\theta}) = \boldsymbol{J}_{\tau}(\boldsymbol{\theta})^{T} \boldsymbol{\Psi}_{\tau}(\boldsymbol{y}; \boldsymbol{\theta}),$$

and the influence function of the RMDPDGE, $\widetilde{\boldsymbol{\theta}}^{\tau}_{G}$, is given by

$$\left( \boldsymbol{J}_{\tau}(\boldsymbol{\theta})^{T} \boldsymbol{J}_{\tau}(\boldsymbol{\theta})) + \boldsymbol{G}(\boldsymbol{\theta}) \boldsymbol{G}(\boldsymbol{\theta})^{T} \right)^{-1} \boldsymbol{J}_{\tau}(\boldsymbol{\theta})^{T} \boldsymbol{\Psi}_{\tau}(\boldsymbol{y}; \boldsymbol{\theta}). \tag{31}$$

We can observe that the influence function of $\widetilde{\boldsymbol{\theta}}^{\tau}_{G}$, obtained in (31), will be bounded if the influence function of the MDPDGE, $\widehat{\boldsymbol{\theta}}^{\tau}_{G}$, given in (29) is bounded. In general, it is not easy to see if it is bounded or not, but in particular situations, this can be solved. On the other hand, if there are no restrictions, $\boldsymbol{G}(\boldsymbol{\theta}) = \mathbf{0}$, and therefore, (31) coincides with (29).

In Section 4.1, we shall present the expression of $\boldsymbol{J}_{\tau}(\theta)$ and $\boldsymbol{\Psi}_{\tau}(y, \theta)$ for some models. Based on that explicit expressions, Figure 1 presents the influence function of the MDPDGE, $\widehat{\boldsymbol{\theta}}^{\tau}_{G}$, for the exponential model with $\theta = 4$ and $\tau = 0, 0.2$ and $0.8$ . At $\tau = 0$, the influence function of $\widehat{\boldsymbol{\theta}}^{\tau}_{G}$ is not bounded, whereas it is bounded at the positive values of the tuning parameter, $\tau = 0.2$ and $0.8$. This fact illustrates the robustness of the MDPDGE for $\tau > 0$ for the exponential model.

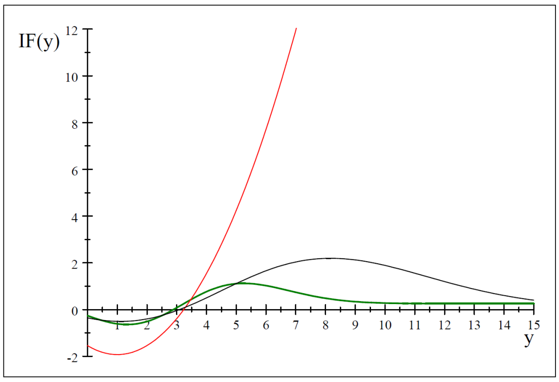


**Figure 1.** Influence function of the MDPDGE for the exponential model with $\tau = 0$ (red), $\tau = 0.2$ (black) and $\tau = 0.8$ (green).

## 4. Rao-Type Tests Based on RMDPDGE

Recently, many robust test statistics based on minimum distance estimators have been introduced in the statistical literature for testing under different statistical models. Among them, density power divergence and Rényi's pseudodistance-based test statistics have shown very competitive performance with respect to classical tests in many different problems. Distance-based test statistics are essentially of two types: Wald-type tests and Rao-type tests. Some applications of these tests can be seen at [8,20–23,29–37] and references therein. In this section, we introduce the Rao-type tests based on RMDPDGE, and we study their asymptotic properties, proving the consistency of the tests.

We analyze here a simple null hypothesis of the form

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \text{ vs. } H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0. \tag{32}$$

**Definition 2.** *Let $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n$ be independent and identically distributed observations from an m-dimensional random vector $\boldsymbol{Y}$ with $E_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\mu}(\boldsymbol{\theta})$ and $Cov_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\Sigma}(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$, and consider the testing problem defined in (32). The Rao-type test statistic based on RMDPDGE is defined by*

$$R_{\boldsymbol{\tau}}(\boldsymbol{\theta}_0) = \frac{1}{n}\boldsymbol{U}_n^{\tau}(\boldsymbol{\theta}_0)^T \boldsymbol{K}_{\boldsymbol{\tau}}(\boldsymbol{\theta}_0)^{-1}\boldsymbol{U}_n^{\tau}(\boldsymbol{\theta}_0), \tag{33}$$

*where*

$$\boldsymbol{U}_n^{\tau}(\boldsymbol{\theta}) = \left(\frac{1}{\tau+1}\sum_{i=1}^{n}\Psi_{\tau}^1(\boldsymbol{y}_i;\boldsymbol{\theta}), ..., \frac{1}{\tau+1}\sum_{i=1}^{n}\Psi_{\tau}^d(\boldsymbol{y}_i;\boldsymbol{\theta})\right)^T$$

*is the score function defining the estimating equations of the MDPDGE and*

$$\boldsymbol{\Psi}_{\boldsymbol{\tau}}(\boldsymbol{y}_i;\boldsymbol{\theta}) = \left(\Psi_{\tau}^1(\boldsymbol{y}_i;\boldsymbol{\theta}), ...., \Psi_{\tau}^d(\boldsymbol{y}_i;\boldsymbol{\theta})\right).$$

The next result establishes the asymptotic behavior of the proposed Rao-type test statistic.

**Theorem 2.** *Let $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n$ be independent and identically distributed observations from an m-dimensional random vector $\boldsymbol{Y}$ with $E_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\mu}(\boldsymbol{\theta})$ and $Cov_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\Sigma}(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$. Under the null hypothesis given in (32), it holds*

$$R_{\boldsymbol{\tau}}(\boldsymbol{\theta}_0) \underset{n\to\infty}{\overset{L}{\to}} \chi_d^2.$$

**Proof.** First, note that we can rewrite

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\tau+1}\boldsymbol{\Psi}_{\boldsymbol{\tau}}(\boldsymbol{y}_i;\boldsymbol{\theta}) = \sqrt{n}\frac{1}{\tau+1}\frac{\partial}{\partial\boldsymbol{\theta}}H_n^{\tau}(\boldsymbol{\theta}),$$

and hence, by Proposition 1, we can establish the asymptotic distribution of the $\tau$-score function $\boldsymbol{U}_n^{\tau}(\boldsymbol{\theta})$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\tau+1}\boldsymbol{\Psi}_{\boldsymbol{\tau}}(\boldsymbol{y}_i;\boldsymbol{\theta}) \underset{n\to\infty}{\overset{L}{\to}} \mathcal{N}\left(\boldsymbol{0}_p, \boldsymbol{K}_{\boldsymbol{\tau}}(\boldsymbol{\theta})\right).$$

Hence, as the $\tau$-score function is asymptotically normal,

$$\frac{1}{\sqrt{n}}\boldsymbol{U}_n^{\tau}(\boldsymbol{\theta}) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{\tau+1}\boldsymbol{\Psi}_{\boldsymbol{\tau}}(\boldsymbol{y}_i;\boldsymbol{\theta}) = \sqrt{n}\frac{1}{\tau+1}\frac{\partial}{\partial\boldsymbol{\theta}}H_n^{\tau}(\boldsymbol{\theta}) \underset{n\longrightarrow\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(\boldsymbol{0}, \boldsymbol{K}_{\boldsymbol{\tau}}(\boldsymbol{\theta})).$$

Then, applying a suitable transformation, the result follows. □

**Remark 2.** *The Rao-type statistic relies on the $\tau$-score function $\boldsymbol{U}_n^{\tau}(\boldsymbol{\theta})$ defining the estimating equations $\boldsymbol{U}_n^{\tau}(\boldsymbol{\theta}) = 0$. Therefore, if the simple null hypothesis holds, the $\tau$-score function vanishes and conversely, if the true parameter is far from the null hypothesis, large $\tau$-scores will be produced.*

*Based on Theorem 2, for large enough sample sizes, one can use the* $100(1-\alpha)$ *percentile of the chi-square distribution with d degrees of freedom* $\chi^2_{d,\alpha}$ *satisfying*

$$\Pr\left(\chi^2_d > \chi^2_{d,\alpha}\right) = \alpha,$$

*to define the reject region of the test with null hypothesis in (32) as*

$$RC = \{R_{\tau}(\boldsymbol{\theta}_0) > \chi^2_{d,\alpha}\}.$$

For illustrative purposes, we present here the application of the proposed method in elliptical distributions.

**Example 1.** *(Elliptical distributions). The m-dimensional random vector* $\boldsymbol{Y}$ *follows an elliptical distribution if its characteristic function has the form*

$$\varphi_{\boldsymbol{Y}}(\boldsymbol{t}) = \exp\left(i\boldsymbol{t}^2\boldsymbol{\mu}\right)\psi\left(\frac{1}{2}\boldsymbol{t}^2\boldsymbol{\Sigma}\boldsymbol{t}\right),$$

*where* $\boldsymbol{\mu}$ *is an m-dimensional vector,* $\boldsymbol{\Sigma}$ *is a positive definite matrix and* $\psi(t)$ *denotes the so-called characteristic generator function. The function* $\psi$ *may depend on the dimension of random vector* $\boldsymbol{Y}$*. In general, it does not hold that* $\boldsymbol{Y}$ *has a joint density function,* $f_{\boldsymbol{Y}}(\boldsymbol{y})$*, but if this density exists, it is given by*

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = c_m|\boldsymbol{\Sigma}|^{-\frac{1}{2}}g_m\left(\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right)$$

*for some density generator function* $g_m$ *which could depend on the dimension of the random vector. Moreover, if the density exists, the parameter* $c_m$ *is given explicitly by*

$$c_m = (2\pi)^{-\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)\left(\int x^{\frac{m}{2}-1}g_m(x)dx\right)^{-1}.$$

*The elliptical distribution family is in the following denoted by* $E_m(\boldsymbol{\mu},\boldsymbol{\Sigma},g_m)$*. For more details about the elliptical family* $E_m(\boldsymbol{\mu},\boldsymbol{\Sigma},g_m)$*, see [38–42] and references therein. In [40], for instance, it can be seen that the mean vector and variance–covariance matrix can be obtained as*

$$E[\boldsymbol{Y}] = \boldsymbol{\mu} \text{ and } Cov[\boldsymbol{Y}] = c_{\boldsymbol{Y}}\boldsymbol{\Sigma},$$

*where* $c_{\boldsymbol{Y}} = -2\psi'(0)$.

*For the elliptical model, the parameter to be estimated is* $\boldsymbol{\theta} = \left(\boldsymbol{\mu}^T, \boldsymbol{\Sigma}\right)$ *whose dimension is* $s = m + \frac{m(m+1)}{2}$*. In the following, we denote* $\boldsymbol{\mu}(\boldsymbol{\theta})$ *instead of* $\boldsymbol{\mu}$ *and* $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ *instead of* $\boldsymbol{\Sigma}$*, in order to be consistent with the paper notation.*

*Let us consider the testing problem*

$$H_0 : (\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta})) = (\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \text{ vs. } H_1 : (\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta})) \neq (\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0), \quad (34)$$

*where* $\boldsymbol{\mu}_0$ *and* $\boldsymbol{\Sigma}_0$ *are known. The Rao-type test statistic based on the MDPDGE for the elliptical model is given as*

$$R_{\tau}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) = \frac{1}{n}\boldsymbol{U}_n^{\tau}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)^T\boldsymbol{K}_{\tau}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)^{-1}\boldsymbol{U}_n^{\tau}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0),$$

*where*

$$\boldsymbol{U}_n^{\tau}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) = \sum_{i=1}^{n}\frac{1}{\tau+1}\boldsymbol{\Psi}_{\tau}(\boldsymbol{y}_i; \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0),$$

*with* $\boldsymbol{\Psi}_{\tau}(\boldsymbol{y}_i; \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ *and* $\boldsymbol{K}_{\tau}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ *as defined in (24) and (10), respectively, but replacing* $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ *by* $c_{\boldsymbol{Y}}\boldsymbol{\Sigma}$ *and* $\boldsymbol{\mu}(\boldsymbol{\theta})$ *by* $\boldsymbol{\mu}$*. Then, the null hypothesis in (34) should be rejected if*

$$R_{\boldsymbol{\tau}}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) > \chi^2_{m+\frac{m(m+1)}{2},\alpha},$$

*with* $\chi^2_{m+\frac{m(m+1)}{2},\alpha}$ *the* $(1-\alpha)$ *upper quantile of a chi-square with* $m + \frac{m(m+1)}{2}$ *degrees of freedom.*

We finally prove the consistency of the Rao-type test based on RMDPDGE. To simplify the statement of the next result, we first define the vector

$$\begin{aligned} \boldsymbol{Y}_\tau(\boldsymbol{\theta}) \;=\; & a\,\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}\Bigg\{-trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right) \\ & \exp\left(-\frac{\tau}{2}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right) + b\; trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right) \\ & +\;\exp\left(-\frac{\tau}{2}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right) \\ & \left[2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta})) \right.\\ & \left. +(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right]\Bigg\}, \end{aligned} \tag{35}$$

where $a$ and $b$ were defined in (6). We can observe that $\frac{\partial}{\partial\boldsymbol{\theta}}H_n(\boldsymbol{\theta})$ is the sample mean of a random sample of size $n$ from the $m$-dimensional population $\boldsymbol{Y}_\tau(\boldsymbol{\theta})$.

**Theorem 3.** *Let* $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n$ *be independent and identically distributed observations from an m-dimensional random vector* $\boldsymbol{Y}$ *with* $E_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\mu}(\boldsymbol{\theta})$ *and* $Cov_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\Sigma}(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$. *Let* $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ *with* $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, *with* $\boldsymbol{\theta}_0$ *defined in (32), and let us assume that* $E_{\boldsymbol{\theta}}[\boldsymbol{Y}_\tau(\boldsymbol{\theta}_0)] \neq \boldsymbol{0}_d$. *Then,*

$$\lim_{n\to\infty} P_{\boldsymbol{\theta}}\left(R_{\boldsymbol{\tau}}(\boldsymbol{\theta}_0) > \chi^2_{d,\alpha}\right) = 1.$$

**Proof.** From the previous results, it holds that

$$\frac{1}{n}\boldsymbol{U}_n^\tau(\boldsymbol{\theta}_0) = \frac{1}{n}\sum_{i=1}^n \frac{1}{\tau+1}\boldsymbol{\Psi}_\tau(\boldsymbol{Y}_i;\boldsymbol{\theta}_0\,) = \frac{1}{\tau+1}\frac{\partial}{\partial\boldsymbol{\theta}}H_n^\tau(\boldsymbol{\theta}_0) \underset{n\to\infty}{\overset{P}{\to}} \frac{1}{\tau+1}E_{\boldsymbol{\theta}}[\boldsymbol{Y}_\tau(\boldsymbol{\theta}_0)],$$

where $\boldsymbol{Y}_\tau(\boldsymbol{\theta}_0)$ is as defined in (35). Therefore,

$$\begin{aligned} P_\theta\left(R_{\boldsymbol{\tau}}(\boldsymbol{\theta}_0) > \chi^2_{d,\alpha}\right) &= P_\theta\left(\tfrac{1}{n}R_{\boldsymbol{\tau}}(\boldsymbol{\theta}_0) > \tfrac{1}{n}\chi^2_{d,\alpha}\right) \\ &\underset{n\to\infty}{\longrightarrow} \mathrm{I}\left(\frac{1}{(\tau+1)^2}E_{\boldsymbol{\theta}}[\boldsymbol{Y}_\tau(\boldsymbol{\theta}_0)]\boldsymbol{K}_\tau^{-1}(\boldsymbol{\theta})E_{\boldsymbol{\theta}}^T[\boldsymbol{Y}_\tau(\boldsymbol{\theta}_0)] > 0\right) = 1, \end{aligned}$$

where $\mathrm{I}(\cdot)$ is the indicator function. □

A natural question that arises here is how the asymptotic power of different test statistics considered for testing the hypothesis in (34) could be compared. Lehmann [43] stated that contiguous alternative hypotheses are of great interest for such purposes, as their associated power functions do not converge to 1. In this regard, we next derive the asymptotic distribution of $R_{\boldsymbol{\tau}}(\boldsymbol{\theta}_0)$ under local Pitman-type alternative hypotheses of the form

$$H_{1,n} : \boldsymbol{\theta} = \boldsymbol{\theta}_n := \boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{l},$$

where $\boldsymbol{l}$ is a $d$-dimensional normal vector. The next result determines the asymptotic power of the Rao-type test based on RMDPDGE under a contiguous alternative hypothesis.

**Theorem 4.** *Let $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n$ be independent and identically distributed observations from an m-dimensional random vector $\boldsymbol{Y}$ with $E_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\mu}(\boldsymbol{\theta})$ and $Cov_{\boldsymbol{\theta}}[\boldsymbol{Y}] = \boldsymbol{\Sigma}(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$. Under the contiguous alternative hypothesis of the form*

$$H_{1,n} : \boldsymbol{\theta}_n = \boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{l},$$

*the asymptotic distribution of the Rao-type test based on RMDPDGE, $R_{\boldsymbol{\tau}}(\boldsymbol{\theta}_0)$, is a non-central chi-square distribution with d degrees of freedom and non-centrality parameter given by*

$$\delta_\tau(\boldsymbol{\theta}_0, \boldsymbol{l}) = \boldsymbol{l}^T \boldsymbol{J}_\tau(\boldsymbol{\theta}_0)\boldsymbol{K}_\tau^{-1}(\boldsymbol{\theta}_0)\boldsymbol{J}_\tau(\boldsymbol{\theta}_0)\boldsymbol{l}.$$

**Proof.** Consider the Taylor series expansion

$$\frac{1}{\sqrt{n}}\boldsymbol{U}_n^\tau(\boldsymbol{\theta}_n) = \frac{1}{\sqrt{n}}\boldsymbol{U}_n^\tau(\boldsymbol{\theta}_0) + \frac{1}{n}\frac{\partial \boldsymbol{U}_n^\tau(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_n^*}\boldsymbol{l},$$

where $\boldsymbol{\theta}_n^*$ belongs to the line segment joining $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_0 + \frac{1}{\sqrt{n}}\boldsymbol{l}$. Now, by Proposition 2

$$\frac{1}{n}\frac{\partial \boldsymbol{U}_n^\tau(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} = \frac{1}{\tau+1}\frac{\partial^2 H_n^\tau(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\, \partial \boldsymbol{\theta}^T} \xrightarrow[n\longrightarrow\infty]{\mathcal{P}} -\boldsymbol{J}_\tau(\boldsymbol{\theta}).$$

Therefore,

$$\frac{1}{\sqrt{n}}\boldsymbol{U}_n^\tau(\boldsymbol{\theta})\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0+n^{-1/2}\boldsymbol{d}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(-\boldsymbol{J}_\tau(\boldsymbol{\theta}_0)\boldsymbol{l},\ \boldsymbol{K}_\tau(\boldsymbol{\theta}_0)),$$

and

$$R_{\boldsymbol{\tau}}(\boldsymbol{\theta}_0) \xrightarrow[n\to\infty]{\mathcal{L}} \chi_p^2(\delta_\tau(\boldsymbol{\theta}_0, \boldsymbol{l})),$$

with $\delta_\tau(\boldsymbol{\theta}_0, \boldsymbol{d})$ given by

$$\delta_\tau(\boldsymbol{\theta}_0, \boldsymbol{l}) = \boldsymbol{l}^T \boldsymbol{J}_\tau(\boldsymbol{\theta}_0)\boldsymbol{K}_\tau^{-1}(\boldsymbol{\theta}_0)\boldsymbol{J}_\tau(\boldsymbol{\theta}_0)\boldsymbol{l}.$$

Hence, the result holds. □

**Remark 3.** *The previous result can be used for defining an approximation to the power function under any alternative hypothesis, $\boldsymbol{\theta} \in \Theta \setminus \Theta_0$, given as*

$$\boldsymbol{\theta} = \boldsymbol{\theta} - \boldsymbol{\theta}_0 + \boldsymbol{\theta}_0 = \sqrt{n}\frac{1}{\sqrt{n}}(\boldsymbol{\theta} - \boldsymbol{\theta}_0) + \boldsymbol{\theta}_0 = \boldsymbol{\theta}_0 + n^{-1/2}\boldsymbol{l},$$

*with $\boldsymbol{l} = \sqrt{n}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$.*

**Remark 4.** *In this section, we have dealt with a Rao-type test for a simple null hypothesis. This family can be extended to a composite null hypothesis. If we are interested in testing $H_0 : \boldsymbol{\theta} \in \boldsymbol{\Theta}_0 = \{\boldsymbol{\theta} \in \Theta /\ \boldsymbol{g}(\boldsymbol{\theta}) = \boldsymbol{0}_r\}$, we can consider the family of Rao-type tests given by*

$$R_{\boldsymbol{\tau}}\left(\widetilde{\boldsymbol{\theta}}_G^\tau\right) = \frac{1}{n}\boldsymbol{U}_n^\tau(\widetilde{\boldsymbol{\theta}}_G^\tau)^T \boldsymbol{Q}_{\boldsymbol{\tau}}(\widetilde{\boldsymbol{\theta}}_G^\tau)\left[\boldsymbol{Q}_{\boldsymbol{\tau}}(\widetilde{\boldsymbol{\theta}}_G^\tau)\boldsymbol{K}_\tau\left(\widetilde{\boldsymbol{\theta}}_G^\tau\right)\boldsymbol{Q}_{\boldsymbol{\tau}}(\widetilde{\boldsymbol{\theta}}_G^\tau)\right]^{-1}\boldsymbol{Q}_{\boldsymbol{\tau}}(\widetilde{\boldsymbol{\theta}}_G^\tau)^T \boldsymbol{U}_n^\tau(\widetilde{\boldsymbol{\theta}}_G^\tau). \tag{36}$$

*However, the extension of the presented results for the family of robust test statistics defined in (36) is not trivial, and it will be established in future research.*

*In particular, the simple null hypothesis in (32) can be written as a composite null hypothesis with $\boldsymbol{g}(\boldsymbol{\theta}) = \boldsymbol{\theta} - \boldsymbol{\theta}_0$. In this case, $\boldsymbol{G}(\boldsymbol{\theta})$ reduces to the identity matrix of dimension p and the restricted estimator $\widetilde{\boldsymbol{\theta}}_G^\tau$ coincides with $\boldsymbol{\theta}_0$. In this case, the Rao-type test statistic $R_{\boldsymbol{\tau}}\left(\widetilde{\boldsymbol{\theta}}_G^\tau\right)$ coincides with the proposed $R_{\boldsymbol{\tau}}(\boldsymbol{\theta}_0)$ given in (33). Rao-type test statistics based on RMDPDE have been developed in [22].*

### 4.1. Rao-Type Tests Based on MDPDGE for Univariate Distributions

Let $Y_1, ..., Y_n$ be a random sample from the population $Y$, with

$$E[Y] = \mu(\theta) \text{ and } Var[Y] = \sigma^2(\theta).$$

Based on (25), the estimating equation is given by

$$\sum_{i=1}^{n} \Psi_\tau(y_i, \theta) = 0,$$

with

$$\begin{aligned} \Psi_\tau(y_i,\theta) &= \frac{(\tau+1)\left(\sigma^2(\theta)\right)^{-\tau/2}}{2(2\pi)^{\tau/2}}\left\{\left[-\frac{\partial \log \sigma^2(\theta)}{\partial\theta} + \frac{\partial \log \sigma^2(\theta)}{\partial\theta}\right.\right. \\ &\left.\left(\frac{y_i-\mu(\theta)}{\sigma^2(\theta)}\right)^2 + 2\frac{\partial\mu(\theta)}{\partial\theta}(y_i-\mu(\theta))\frac{1}{\sigma^2(\theta)}\right] \\ &\left.\times\exp\left(-\frac{\tau}{2\sigma^2(\theta)}(y_i-\mu(\theta))^2\right) + \frac{\tau}{(1+\tau)^{3/2}}\frac{\partial\log\sigma^2(\theta)}{\partial\theta}\right\}. \end{aligned} \tag{37}$$

Moreover, the expressions of $J_\tau(\theta)$ and $K_\tau(\theta)$ are, respectively, given by

$$\begin{aligned} J_\tau(\theta) &= \frac{1}{\left(2\pi\sigma(\theta)^2\right)^{\frac{\tau}{2}}}\frac{1}{(1+\tau)^{5/2}}\left[(\tau+1)\sigma^{-2}(\theta)\left(\frac{\partial\mu(\theta)}{\partial\theta}\right)^2 + \frac{\tau^2}{4}\left(\frac{\partial\log\sigma^2(\theta)}{\partial\theta}\right)^2\right. \\ &\left.+\frac{1}{2}\left(\frac{\partial\log\sigma^2(\theta)}{\partial\theta}\right)^2\right] \end{aligned}$$

and

$$\begin{aligned} K_\tau(\theta) &= \left(\frac{1}{(2\pi)^{1/2}\sigma(\theta)}\right)^{2\tau}\left\{\frac{1}{(1+2\tau)^{5/2}}\left[\tau^2\left(\frac{\partial\log\sigma^2(\theta)}{\partial\theta}\right)^2\right.\right. \\ &\left.+(1+2\tau)\sigma^{-2}(\theta)\left(\frac{\partial\mu(\theta)}{\partial\theta}\right)^2 + \frac{1}{2}\left(\frac{\partial\log\sigma^2(\theta)}{\partial\theta}\right)^2\right] \\ &\left.-\frac{\tau^2}{4(1+\tau)^3}\left(\frac{\partial\log\sigma^2(\theta)}{\partial\theta}\right)^2\right\}. \end{aligned} \tag{38}$$

Therefore, if we are interesting in testing

$$H_0 : \theta = \theta_0 \text{ vs. } H_1 : \theta \neq \theta_0,$$

the Rao-type tests based on RMDPDGE are given by

$$R_{\boldsymbol{\tau}}(\theta_0) = \frac{1}{n}U_n^{\boldsymbol{\tau}}(\theta_0)^2 K_{\boldsymbol{\tau}}(\theta_0)^{-1},$$

where

$$U_n^{\boldsymbol{\tau}}(\theta_0) = \frac{1}{\tau+1}\sum_{i=1}^{n}\Psi_\tau(y_i,\theta)$$

and $\Psi_\tau(y_i, \theta)$ and $K_\tau(\theta)$ are given in (37) and (38). The null hypothesis is rejected if

$$R_{\boldsymbol{\tau}}(\theta_0) > \chi^2_{1,\alpha},$$

where $\chi^2_{1,\alpha}$ is the upper $1-\alpha$ quantile of a chi-square distribution with 1 degree of freedom.

We finally derive explicit expressions of the Rao-type test statistics under Poisson, exponential and Lindley models.

4.1.1. Poisson Model

Let us assume that the random variable $Y$ is Poisson with parameter $\theta$. In this case, it is well known that $E[Y] = Var[Y] = \theta$, and so the RMDPDGE, for $\tau > 0$, is given by

$$\widehat{\theta}_G^{\tau} = \arg\max_{\theta}\left\{ \frac{\tau+1}{\tau(2\pi\theta)^{\frac{\tau}{2}}}\left(\frac{1}{n}\sum_{i=1}^{n}\exp\left(-\frac{\tau}{2\theta}(y_i-\theta)^2\right) - \frac{\tau}{(1+\tau)^{3/2}}\right) - \frac{1}{\tau}\right\}.$$

At $\tau = 0$, the RMDPDGE reduces to the Gaussian MLE,

$$\widehat{\theta}_G = \arg\max_{\theta}\left\{ -\frac{1}{2}\log 2\pi - \frac{1}{2}\log\theta - \frac{1}{n}\sum_{i=1}^{n}\frac{1}{2\theta}(y_i-\theta)^2\right\}.$$

On the other hand, the score function $\Psi_\tau(\cdot)$ is given by

$$\Psi_\tau(y_i,\theta) = \frac{\tau+1}{2(2\pi\theta)^{\frac{\tau}{2}}\theta^2}\left\{\left(-\theta^2-\theta+y_i^2\right)\exp\left(-\frac{\tau}{2\theta}(y_i-\theta)^2\right) + \frac{\tau\theta}{(1+\tau)^{\frac{3}{2}}}\right\},$$

and naturally, at $\tau = 0$, we obtain the score function of the Gaussian MLE presented in [1]

$$\Psi_0(y_i,\theta) = \frac{1}{2\theta^2}\left(-2\theta^2 + y_i^2\right).$$

On the other hand, the matrix $K_\tau(\theta)$ under the Poisson model has the explicit expression

$$K_\tau(\theta) = \left(\frac{1}{2\pi}\right)^{\tau}\frac{1}{2\theta^{2+\tau}}\left\{\frac{1}{(1+2\tau)^{5/2}}\left(\left(2\tau^2+2\theta+4\theta\tau+1\right) - \frac{\tau^2}{2(1+\tau)^3}\right)\right\},$$

and hence, the Rao-type tests based on RMDPDGE, $R_{\boldsymbol{\tau}}(\theta_0)$, for testing a simple null hypothesis is given, for $\tau > 0$, by

$$\begin{aligned} R_{\boldsymbol{\tau}}(\theta_0) &= \frac{1}{n}\frac{1}{\left(2(2\pi\theta)^{\frac{\tau}{2}}\theta^2\right)^2}\left(\sum_{i=1}^{n}\left(\left(-2\theta_0^2+y_i^2\right)\exp\left(-\frac{\tau}{2\theta_0}(y_i-\theta_0)^2\right) + \frac{\tau\theta}{(1+\tau)^{\frac{3}{2}}}\right)\right)^2 \\ &\quad \times(2\pi)^{\tau}(2\theta^{2+\tau})(1+2\tau)^{5/2}\left\{\left(\left(2\tau^2+2\theta+4\theta\tau+1\right) - \frac{\tau^2}{2(1+\tau)^3}\right)\right\}^{-1}. \end{aligned}$$

Again, for $\tau = 0$, we obtain the expression of the classical Rao test based on the Gaussian MLE,

$$R_0(\theta_0) = \frac{1}{4n}\left(\sum_{i=1}^{n}\left(\frac{-2\theta_0^2+y_i^2}{\theta_0^2}\right)\right)^2\frac{2\theta_0^2}{2\theta_0+1}.$$

The null hypothesis is rejected if

$$R_{\boldsymbol{\tau}}(\theta_0) > \chi^2_{1,\alpha}.$$

4.1.2. Exponential Model

Let us assume now that the random variable $Y$ comes from an exponential distribution with probability density function

$$f_\theta(x) = \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right),\ x > 0. \tag{39}$$

In this case, the true mean and variance are given by $E[Y] = \theta$ and $Var[Y] = \theta^2$. The RMDPDGE under the exponential model, for $\tau > 0$, is given by

$$\widehat{\theta}_G^\tau = \arg\max_\theta \left\{ \frac{\tau+1}{\tau} \left( \frac{1}{\theta\sqrt{2\pi}} \right)^\tau \left( \frac{1}{n} \sum_{i=1}^n \exp\left( -\frac{\tau}{2} \left( \frac{y_i - \theta}{\theta} \right)^2 \right) - \frac{\tau}{(1+\tau)^{3/2}} \right) - \frac{1}{\tau} \right\},$$

and for $\tau = 0$, we have the Gaussian MLE

$$\widehat{\theta}_G = \arg\max_\theta \left\{ -\frac{1}{2}\log 2\pi - \log\theta - \frac{1}{n}\sum_{i=1}^n \frac{1}{2}\left( \frac{y_i-\theta}{\theta} \right)^2 \right\}.$$

On the other hand, the score function is

$$\Psi_\tau(y_i, \theta) = \frac{(\tau+1)}{\theta^{\tau+3}\left(\sqrt{2\pi}\right)^\tau} \left\{ \left( y_i^2 - y_i\theta - \theta^2 \right) \exp\left( -\frac{\tau}{2}\left( \frac{y_i-\theta}{\theta} \right)^2 \right) + \frac{\tau\theta^2}{(1+\tau)^{\frac{3}{2}}} \right\},$$

and for $\tau = 0$, we recover the score function of the Gaussian MLE,

$$\Psi_0(y_i, \theta) = \frac{1}{\theta^3}\left( y_i^2 - y_i\theta + \theta^2 \right).$$

The value $K_\tau(\theta)$ has the expression

$$K_\tau(\theta) = \frac{1}{(2\pi)^\tau \theta^{2(\tau+1)}} \left\{ \frac{1}{(1+2\tau)^{5/2}}\left( 4\tau^2 + 2\tau + 3 \right) - \frac{\tau^2}{(1+\tau)^3} \right\},$$

and at $\tau = 0$

$$K_0(\theta) = \frac{2}{\theta^2}.$$

Correspondingly, the Rao-type tests based on RMDPDGE for testing

$$H_0 : \theta = \theta_0 \text{ vs. } H_1 : \theta \neq \theta_0,$$

is given, for $\tau > 0$, by

$$\begin{aligned} R_{\boldsymbol{\tau}}(\theta_0) \quad = \quad & \frac{1}{n}\frac{1}{\theta_0^{2\tau+6}(2\pi)^\tau} \left( \sum_{i=1}^n \left\{ \left( y_i^2 - y_i\theta_0 - \theta_0^2 \right) \exp\left( -\frac{\tau}{2}\left( \frac{y_i-\theta_0}{\theta_0} \right)^2 \right) \right. \right. \\ & \left. \left. + \frac{\tau\theta_0^2}{(1+\tau)^{\frac{3}{2}}} \right\} \right)^2 \times (2\pi)^\tau \theta^{2(\tau+1)} \left\{ \frac{1}{(1+2\tau)^{5/2}}\left( 4\tau^2 + 2\tau + 3 \right) - \frac{\tau^2}{(1+\tau)^3} \right\}^{-1}, \end{aligned} \tag{40}$$

and by

$$R_{\mathbf{0}}(\theta_0) = \frac{1}{2n}\sum_{i=1}^n \frac{\left( y_i^2 - y_i\theta_0 - \theta_0^2 \right)^2}{\theta_0^4} \tag{41}$$

for $\tau = 0$.

#### 4.1.3. Lindley Model

Let us finally assume that the random variable $Y$ comes from a Lindley distribution [44] with probability density function

$$f_\theta(x) = \frac{\theta^2}{\theta+1}(1+x)\exp(-\theta x),\ x > 0, \theta > 0.$$

In this case,

$$E[Y] = \frac{\theta+2}{\theta(\theta+1)} \text{ and } V[Y] = \frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2}.$$

The RMDPGE under the Lindlley model, for $\tau > 0$, is given by

$$\begin{aligned}\widehat{\theta}_G &= \arg\max_{\theta}\left\{\frac{\tau+1}{\tau}\left(\frac{1}{2\pi}\right)^{\tau}\left(\frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2}\right)^{-\frac{\tau}{2}}\right. \\ &\left.\left(\frac{1}{n}\sum_{i=1}^{n}\exp\left(-\frac{\tau}{2}\left(y_i-\frac{\theta+2}{\theta(\theta+1)}\right)^2\left(\frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2}\right)^{-1}\right)-\frac{\tau}{(1+\tau)^{\frac{3}{2}}}\right)-\frac{1}{\tau}\right\},\end{aligned}$$

and for $\tau = 0$, we have

$$\widehat{\theta}_G = \arg\max_{\theta}\left\{\frac{1}{2}\ln\frac{1}{2\pi}-\frac{1}{2}\ln\frac{(\theta^2+4\theta+2)}{\theta^2(\theta+1)^2}-\frac{1}{n}\sum_{i=1}^{n}\frac{(\theta-y_i\theta-y_i\theta^2+2)^2}{2(\theta^2+4\theta+2)}\right\}.$$

On the other hand, the score funtion, $\Psi_\tau(y_i,\theta)$ is given by

$$\begin{aligned}&\frac{(\tau+1)}{2(2\pi)^{\frac{\tau}{2}}}\left(\frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2}\right)^{-\frac{\tau}{2}}\left\{\left[\frac{2\theta^3+12\theta^2+12\theta+4}{\theta^4+5\theta^3+6\theta^2+2\theta}\right.\right. \\ &-\frac{2\theta^3+12\theta^2+12\theta+4}{\theta^4+5\theta^3+6\theta^2+2\theta}\frac{(\theta-\theta y_i-\theta^2y_i+2)^2}{\theta^2+4\theta+2}-2\left.\left(y_i-\frac{(\theta+2)}{\theta(\theta+1)}\right)\right] \\ &\left.\exp\left(-\frac{\tau}{2}\frac{(\theta-\theta y_i-\theta^2y_i+2)^2}{\theta^2+4\theta+2}\right)-\frac{\tau}{(1+\tau)^{\frac{3}{2}}}\frac{2\theta^3+12\theta^2+12\theta+4}{\theta^4+5\theta^3+6\theta^2+2\theta}\right\}\end{aligned} \tag{42}$$

and $K_\tau(\theta)$ has the expression

$$\begin{aligned}&\left(\frac{\theta^2(\theta+1)^2}{2\pi(\theta^2+4\theta+2)}\right)^{\tau}\left\{\frac{1}{(1+2\tau)^{5/2}}\left[\tau^2\left(\frac{2\theta^3+12\theta^2+12\theta+4}{\theta^4+5\theta^3+6\theta^2+2\theta}\right)^2\right.\right. \\ &\left.+(1+2\tau)\frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2}+\frac{1}{2}\left(\frac{2\theta^3+12\theta^2+12\theta+4}{\theta^4+5\theta^3+6\theta^2+2\theta}\right)^2\right] \\ &\left.-\frac{\tau^2}{4(1+\tau)^3}\left(\frac{2\theta^3+12\theta^2+12\theta+4}{\theta^4+5\theta^3+6\theta^2+2\theta}\right)^2\right\}\end{aligned} \tag{43}$$

and for $\tau = 0$,

$$K_0(\theta) = \frac{(2\theta^3+12\theta^2+12\theta+4)^2}{2(\theta^4+5\theta^3+6\theta^2+2\theta)^2}+\frac{(\theta+2)}{\theta(\theta+1)}.$$

Therefore, the Rao-type test based on RMDPDGE for testing

$$H_0: \theta = \theta_0 \text{ vs. } H_1: \theta \neq \theta_0$$

is given by

$$R_{\boldsymbol{\tau}}(\theta_0) = \frac{1}{n}U_n^{\boldsymbol{\tau}}(\theta_0)^2K_{\boldsymbol{\tau}}(\theta_0)^{-1} > \chi^2_{1,\alpha},$$

where

$$U_n^{\boldsymbol{\tau}}(\theta_0) = \frac{1}{\tau+1}\sum_{i=1}^{n}\Psi_\tau(y_i,\theta),$$

with $\Psi_\tau(y_i,\theta)$ is defined in (42) and $K_\tau(\theta)$ in (43).

## 5. Simulation Study

We analyze here the performance of the Rao-type tests based on the MDPDGE, $R_{\boldsymbol{\tau}}(\theta_0)$, in terms of robustness and efficiency. We compare the proposed general method assuming Gaussian distribution with Rao-type test statistics based on the true parametric distribution underlying the data.

We consider the exponential model with density function $f_{\theta_0}(x)$ given in (39). For the exponential model, the Rao-type test statistics based on MDPDGE is for $\tau > 0$ given in (40) and for $\tau = 0$ given in (41). To evaluate the robustness of the tests, we generate samples from an exponential mixture,

$$f_{\theta_0}^{\varepsilon}(x) = (1-\varepsilon) f_{\theta_0}(x) + \varepsilon f_{2\theta_0}(x),$$

where $\theta_0$ denotes the parameter of the exponential distribution and $\varepsilon$ is the contamination proportion. The uncontaminated model is thus obtained by setting $\varepsilon = 0$.

For comparison purposes, we have also considered the robust Rao-type tests based on the restricted MDPDE, which was introduced and studied [30]. The efficiency loss caused by the Gaussian assumption should be advertised by the poorer performance of the Rao-type tests based on the restricted MDPDGE with respect to their analogous based on the restricted MDPDE. For the exponential model, the family Rao-type test statistics based on the restricted MDPDE is given, for $\beta > 0$, by

$$S_n^{\beta}(\theta_0) = \left(\frac{4\beta^2+1}{(2\beta+1)^3} - \frac{\beta^2}{(\beta+1)^4}\right)^{-1} \frac{1}{n}\left(\frac{1}{\theta_0}\sum_{i=1}^{n}(y_i-\theta_0)\exp\left(-\frac{\beta y_i}{\theta_0}\right) + \frac{n\beta}{(\beta+1)^2}\right)^2.$$

For $\beta = 0$, the above test reduces to the classical Rao test given by

$$S_n(\theta_0) = S_{\beta=0,n}(\theta_0) = \left(\sqrt{n}\frac{\bar{X}_n - \theta_0}{\theta_0}\right)^2.$$

We consider the testing problem

$$H_0 : \theta = \theta_0 \text{ vs. } H_1 : \theta \neq \theta_0,$$

and we empirically examine the level and power of both Rao-type test statistics, the usual test based on the parametric model and the Gaussian-based test by setting the true value of the parameter $\theta_0 = 2$ and $\theta_0 = 1$, respectively. Different sample sizes were considered, namely $n = 10, 20, 30, 40, 50, 60, 70, 80, 90, 100$ and 200, but simulation results were quite similar and so, for brevity, we only report here results for $n = 20$ and $n = 40$.

The empirical level of the test is computed

$$\widehat{\alpha}_n(\varepsilon) = \frac{\text{Number of times}\left\{R_n^{\tau}(\theta_0) \text{ (or } S_n^{\beta}(\theta_0)) > \chi^2_{1,0.05} = 3.84146\right\}}{\text{Number of simulated samples}}.$$

We set $\varepsilon = 0\%, 5\%, 10\%$ and $20\%$ of contamination proportions and perform the Monte-Carlo study over $R$ = 10,000 replications. The tuning parameters $\tau$ and $\beta$ are fixed from a grid of values, namely $\{0, 0.1, ..., 0.7\}$.

Simulation results are presented in Tables 1 and 2 for $n = 20$ and $n = 40$, respectively. The empirical powers are denoted by $\widehat{\pi}_n(\varepsilon)$. The robustness advantage in terms of level of both Rao-type tests considered, $R_{\boldsymbol{\tau}}(\theta_0)$ and $S_n^{\beta}(\theta_0)$ with positive values of the turning parameter with respect to the test statistics with $\tau = 0$ and $\beta = 0$ is clearly shown, as their simulated levels are closer to the nominal value in the presence of contamination.

**Table 1.** Simulated levels for different contamination proportions and different tuning parameters $\tau, \beta = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7$ for the Rao-type tests $R_{\boldsymbol{\tau}}(\theta_0)$ and $S_n^{\beta}(\theta_0)$ for $n = 20$.

| $\tau$ | $\widehat{\alpha}_n(0)$ | $\widehat{\alpha}_n(0.05)$ | $\widehat{\alpha}_n(0.10)$ | $\widehat{\alpha}_n(0.20)$ | $\widehat{\pi}_n(0)$ | $\widehat{\pi}_n(0.1)$ | $\widehat{\pi}_n(0.15)$ | $\widehat{\pi}_n(0.20)$ |
|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.2601 | 0.3093 | 0.3453 | 0.4661 | 0.9278 | 0.6791 | 0.6887 | 0.5088 |
| 0.1 | 0.1895 | 0.1748 | 0.1561 | 0.1989 | 0.9544 | 0.7213 | 0.7301 | 0.0595 |
| 0.2 | 0.2120 | 0.1776 | 0.1417 | 0.1174 | 0.9747 | 0.8398 | 0.8430 | 0.6395 |
| 0.3 | 0.2532 | 0.2113 | 0.1660 | 0.1275 | 0.9826 | 0.8963 | 0.8961 | 0.7301 |
| 0.4 | 0.2963 | 0.2447 | 0.1986 | 0.1471 | 0.9863 | 0.9228 | 0.9257 | 0.7893 |
| 0.5 | 0.3243 | 0.2773 | 0.2307 | 0.1695 | 0.9875 | 0.9363 | 0.9386 | 0.8254 |
| 0.6 | 0.3512 | 0.3055 | 0.2599 | 0.1899 | 0.9885 | 0.9441 | 0.9437 | 0.8434 |
| 0.7 | 0.3751 | 0.3258 | 0.2762 | 0.2060 | 0.9884 | 0.9466 | 0.9469 | 0.8541 |
| $\beta$ | | | | | | | | |
| 0.0 | 0.0453 | 0.0682 | 0.1048 | 0.1909 | 0.7200 | 0.4365 | 0.4384 | 0.2323 |
| 0.1 | 0.0476 | 0.0602 | 0.0780 | 0.1417 | 0.7799 | 0.5223 | 0.5267 | 0.3029 |
| 0.2 | 0.0498 | 0.0552 | 0.0667 | 0.1103 | 0.7922 | 0.5751 | 0.5780 | 0.3558 |
| 0.3 | 0.0494 | 0.0517 | 0.0584 | 0.0897 | 0.7882 | 0.5997 | 0.6024 | 0.3878 |
| 0.4 | 0.0489 | 0.0505 | 0.0535 | 0.0773 | 0.7779 | 0.6067 | 0.6058 | 0.4106 |
| 0.5 | 0.0494 | 0.0498 | 0.0504 | 0.0692 | 0.7634 | 0.6048 | 0.6037 | 0.4221 |
| 0.6 | 0.0491 | 0.0504 | 0.0497 | 0.0647 | 0.7492 | 0.6008 | 0.5986 | 0.4265 |
| 0.7 | 0.0502 | 0.0495 | 0.0494 | 0.0613 | 0.7348 | 0.5932 | 0.5919 | 0.4259 |

**Table 2.** Simulated levels for different contamination proportions and different tuning parameters $\tau, \beta = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7$ for the Rao-type tests $R_{\boldsymbol{\tau}}(\theta_0)$ and $S_n^{\beta}(\theta_0)$ for $n = 40$.

| $\tau$ | $\widehat{\alpha}_n(0)$ | $\widehat{\alpha}_n(0.05)$ | $\widehat{\alpha}_n(0.10)$ | $\widehat{\alpha}_n(0.20)$ | $\widehat{\pi}_n(0)$ | $\widehat{\pi}_n(0.1)$ | $\widehat{\pi}_n(0.15)$ | $\widehat{\pi}_n(0.20)$ |
|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.3014 | 0.3588 | 0.4407 | 0.5919 | 0.9948 | 0.8064 | 0.7591 | 0.5957 |
| 0.1 | 0.2393 | 0.1934 | 0.1757 | 0.2032 | 0.9991 | 0.9540 | 0.9229 | 0.7712 |
| 0.2 | 0.7712 | 0.2559 | 0.1970 | 0.1317 | 0.9995 | 0.9916 | 0.9846 | 0.9204 |
| 0.3 | 0.4257 | 0.3485 | 0.2782 | 0.1753 | 0.9997 | 0.9997 | 0.9953 | 0.9694 |
| 0.4 | 0.5021 | 0.4294 | 0.3572 | 0.2388 | 0.9999 | 0.9989 | 0.9978 | 0.9851 |
| 0.5 | 0.5642 | 0.4920 | 0.4253 | 0.2993 | 0.9999 | 0.9992 | 0.9986 | 0.9908 |
| 0.6 | 0.6084 | 0.5415 | 0.4742 | 0.3491 | 1.0000 | 0.9992 | 0.9994 | 0.9935 |
| 0.7 | 0.6416 | 0.5755 | 0.5081 | 0.3831 | 1.0000 | 0.9994 | 0.9994 | 0.9948 |
| $\beta$ | | | | | | | | |
| 0.0 | 0.0467 | 0.0758 | 0.1309 | 0.2728 | 0.9838 | 0.8093 | 0.7483 | 0.4905 |
| 0.1 | 0.0469 | 0.0623 | 0.0959 | 0.1987 | 0.9870 | 0.8770 | 0.8317 | 0.6072 |
| 0.2 | 0.0464 | 0.0554 | 0.0800 | 0.1526 | 0.9862 | 0.9010 | 0.8687 | 0.6778 |
| 0.3 | 0.0481 | 0.0529 | 0.0704 | 0.1220 | 0.9846 | 0.9084 | 0.8804 | 0.7169 |
| 0.4 | 0.0483 | 0.0518 | 0.0649 | 0.1036 | 0.9808 | 0.9059 | 0.8809 | 0.7316 |
| 0.5 | 0.0500 | 0.0519 | 0.0618 | 0.0929 | 0.9756 | 0.9008 | 0.8742 | 0.7338 |
| 0.6 | 0.0500 | 0.0501 | 0.0577 | 0.0858 | 0.9689 | 0.8914 | 0.8662 | 0.7317 |
| 0.7 | 0.0504 | 0.0519 | 0.0562 | 0.0801 | 0.9634 | 0.8813 | 0.8562 | 0.7258 |

Regarding the power of the tests for uncontaminated scenarios, there are values at least as good as the corresponding to $\tau = 0$ and $\beta = 0$, and for contaminated data, the power corresponding to $\tau > 0$ and $\beta > 0$ is higher.

The loss of efficiency caused by the Guassian assumption can be measured by the discrepancy of the estimated levels and powers between the family of Rao-type tests based on the restricted MDPDGE and the restricted MDPDE. As expected, empirical levels of the test statistics based on the restricted MDPDGE are higher than the corresponding levels of the test based on the restricted MDPDE. However, the test statistic based on the parametric model, $S_n^{\beta}(\theta_0)$, is quite conservative and so the corresponding powers are higher than those of the proposed tests, $R_{\boldsymbol{\tau}}(\theta_0)$. Based on the presented results, it seems that the proposed Rao-type tests, $R_{\boldsymbol{\tau}}(\theta_0)$, performs reasonably well and offers an appealing alternative for situations where the probability density function of the true model is unknown or it is very complicated to work with it.

## 6. Conclusions

In this paper, we have considered the situation in which the parametric distribution of the variables is unknown and the only available information is given in terms of the mean vector and the variance–covariance matrix. Hence, we only know that the mean vector and variance–covariance matrix depend on the unknown values of a parameter vector $\boldsymbol{\theta}$. To deal with this problem, Zhang [1] proposed a procedure to estimate $\boldsymbol{\theta}$ assuming that the underlying distribution was Gaussian. This assumption is justified in terms of the maximum entropy of the unknown distribution. However, the estimator developed in [1] is not robust. This procedure was extended using DPD leading to robust estimations of $\boldsymbol{\theta}$. In this paper, we have dealt with the case in which additional constraints must be imposed to the estimated parameters, thus leading to a new family of estimators that we have named RMDPGE. For these estimators, we have derived their asymptotic distribution and we have studied their robustness properties in terms of the corresponding IF. As an application, we have developed robust Rao-type test statistics under the null hypothesis, where the null hypothesis is indeed the restricted version of the estimator. Finally, we have tested the performance of these test statistics via a simulation study. From the results of this study, we empirically showed that the Rao-type tests considered in this paper seem to have a good performance in terms of efficiency and are more robust than the corresponding approach in [1].

There are several problems to be treated in future research. The most natural seems to develop Rao-type test statistics for composite null hypothesis and study the results obtained in terms of efficiency and robustness.

**Author Contributions:** Conceptualization, Á.F., M.J., P.M. and L.P.; methodology, Á.F., M.J., P.M. and L.P.; software, Á.F., M.J., P.M. and L.P.; validation, Á.F., M.J., P.M. and L.P.; formal analysis, Á.F., M.J., P.M. and L.P.; investigation, Á.F., M.J., P.M. and L.P.; resources, Á.F., M.J., P.M. and L.P.; data curation, Á.F., M.J., P.M. and L.P.; writing—original draft preparation, Á.F., M.J., P.M. and L.P.; writing—review and editing, Á.F., M.J., P.M. and L.P.; visualization, Á.F., M.J., P.M. and L.P.; supervision, Á.F., M.J., P.M. and L.P.; project administration, Á.F., M.J., P.M. and L.P.; funding acquisition, Á.F., M.J., P.M. and L.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Spanish Grant PID2021-124933NB-I00.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** Jaenada, M., Miranda P. and Pardo L. are members of the Interdisciplinary Mathematics Institute.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| MLE | Maximum likelihood estimator |
| MDPDE | Minimum density power divergence estimator |
| MDPDGE | Minimum density power divergence Gaussian estimator |
| RMDPDE | Restricted minimum density power divergence estimator |
| RMDPDGE | Restricted minimum density power divergence Gaussian estimator |

## Appendix A. Derivatives Calculation and Proofs of the Main Results

### *Appendix A.1. Previous Results*

In different parts of Appendix, the following results are applied.

**Lemma A1.** *The following results can be shown:*

1. $\dfrac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i} = |\boldsymbol{\Sigma}(\boldsymbol{\theta})| \, trace\left( \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \dfrac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i} \right).$
2. $\frac{\partial trace(\boldsymbol{\Sigma}(\boldsymbol{\theta}))}{\partial \theta_i} = trace\left( \dfrac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i} \right).$
3. $\dfrac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}^{-1} = -\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \dfrac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i} \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}.$

**Lemma A2.** *Let $\boldsymbol{Y}$ be a normal population with vector mean $\boldsymbol{\mu}$ and variance–covariance $\boldsymbol{\Sigma}$. Then, we have*

$$E\left[ (\boldsymbol{Y} - \boldsymbol{\mu})^T \boldsymbol{A} (\boldsymbol{Y} - \boldsymbol{\mu}) \right] = Trace(\boldsymbol{A\Sigma}),$$

$$E\left[ (\boldsymbol{Y} - \boldsymbol{\mu})^T \boldsymbol{A} (\boldsymbol{Y} - \boldsymbol{\mu}) (\boldsymbol{Y} - \boldsymbol{\mu})^T \boldsymbol{B} (\boldsymbol{Y} - \boldsymbol{\mu}) \right] = Trace\left( \boldsymbol{A\Sigma} \left( \boldsymbol{B} + \boldsymbol{B}^T \right) \boldsymbol{\Sigma} \right) + Trace(\boldsymbol{A\Sigma}) Trace(\boldsymbol{B\Sigma}),$$

$$E\left[ (\boldsymbol{Y} - \boldsymbol{\mu})^T \boldsymbol{A} (\boldsymbol{Y} - \boldsymbol{\mu}) (\boldsymbol{Y} - \boldsymbol{\mu}) \right] = \boldsymbol{0}.$$

For more details about these results, see for instance [45].

*Appendix A.2*

**Proof of Proposition 1.** The expression of $H_n^\tau(\boldsymbol{\theta})$ introduced in (5) is given by

$$H_n^\tau(\boldsymbol{\theta}) = a|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2} \left( \frac{1}{n} \sum_{i=1}^{n} \exp\left\{ -\frac{\tau}{2} (\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} (\boldsymbol{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta})) \right\} - b \right) - \frac{1}{\tau}.$$

Consider the $d$-dimensional random vector $\boldsymbol{Y}_\tau(\boldsymbol{\theta})$ defined in (35). Applying the Central Limit Theorem, we have

$$\sqrt{n} \frac{\partial}{\partial \boldsymbol{\theta}} H_n^\tau(\boldsymbol{\theta}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{\Psi}_{\boldsymbol{\tau}}(\boldsymbol{y}_i; \boldsymbol{\theta}) \underset{n \longrightarrow \infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(\boldsymbol{0}_m, \boldsymbol{S}_{\boldsymbol{\tau}}(\boldsymbol{\theta}_0)),$$

with

$$\boldsymbol{S}_{\boldsymbol{\tau}}(\boldsymbol{\theta}_0) = Cov[\boldsymbol{Y}_\tau(\boldsymbol{\theta})] = E\left[ \boldsymbol{Y}_\tau(\boldsymbol{\theta})^T \boldsymbol{Y}_\tau(\boldsymbol{\theta}) \right],$$

because

$$E[\boldsymbol{Y}_\tau(\boldsymbol{\theta})] = \boldsymbol{0}_d.$$

To see that $E[\boldsymbol{Y}_\tau(\boldsymbol{\theta})] = \mathbf{0}_d$, consider

$$
\begin{aligned}
E[\boldsymbol{Y}_\tau(\boldsymbol{\theta})] &= a\,\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}E\Bigg[-trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)\\
&\quad \exp\Big\{-\frac{\tau}{2}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\Big\} + b\;trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)\\
&\quad +\;\exp\Big\{-\frac{\tau}{2}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\Big\}\\
&\quad \Bigg[-2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\\
&\quad +(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\Bigg]\Bigg]\\
&= a\,\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}\Bigg\{-trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)\frac{1}{(\tau+1)^{m/2}}\\
&\quad +\frac{\tau}{(\tau+1)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)\\
&\quad +\frac{1}{(\tau+1)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)\Bigg\}\\
&= \mathbf{0}_d.
\end{aligned}
$$

We can observe that $\boldsymbol{Y}_\tau(\boldsymbol{\theta})$ is a $d$-dimensional vector whose $j$-th component is

$$
\begin{aligned}
Y_\tau^j(\boldsymbol{\theta}) &= a\,\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}\Bigg\{-trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\\
&\quad \exp\Big\{-\frac{\tau}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\Big\} + b\;trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\\
&\quad +\;\exp\Big\{-\frac{\tau}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\Big\}\\
&\quad \Bigg[-2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\\
&\quad +(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\Bigg]\Bigg\},\; j=1,\ldots,d.
\end{aligned}
$$

Therefore, the element $(i,j)$ of the matrix $\boldsymbol{S}_{\boldsymbol{\tau}}(\boldsymbol{\theta}_0)$ is given by

$$
E\left[Y_\tau^i(\boldsymbol{\theta})Y_\tau^j(\boldsymbol{\theta})\right].
$$

We are going to obtain $Y_\tau^i(\boldsymbol{\theta})Y_\tau^j(\boldsymbol{\theta})$. First,

$$\begin{aligned} Y_\tau^i(\boldsymbol{\theta})Y_\tau^j(\boldsymbol{\theta}) \quad = \quad & \Bigg\{a\ \frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}\Bigg[-trace\Bigg(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\Bigg) \\ & \exp\Big\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\Big\} + b\ trace\Bigg(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\Bigg) \\ & +\ \exp\Big\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\Big\} \\ & \Bigg[2\Bigg(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\Bigg)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta})) \\ & +(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\Bigg(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\Bigg)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\Bigg]\Bigg]\Bigg\} \\ & \Bigg\{a\ \frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}\Bigg[-trace\Bigg(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\Bigg) \\ & \exp\Big\{-\frac{\tau}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\Big\} + b\ trace\Bigg(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\Bigg) \\ & +\ \exp\Big\{-\frac{\tau}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\Big\} \\ & \Bigg[2\Bigg(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\Bigg)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta})) \\ & +(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\Bigg(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\Bigg)(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\Bigg]\Bigg]\Bigg\}. \end{aligned}$$

Therefore, $Y_\tau^i(\boldsymbol{\theta})Y_\tau^j(\boldsymbol{\theta})$ is given by

$$
\begin{aligned}
& a^2\left(\frac{\tau}{2}\right)^2|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau}\Bigg\{trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\\
& \times\exp\left\{-\frac{2\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\\
& -trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)b\\
& \times\exp\left\{-\frac{2\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\\
& -trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)\exp\left\{-\frac{2\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\\
& \times\left[2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))+(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right]\\
& -b\ trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\\
& +b^2trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\\
& +b\ trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\\
& \times\left[2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))+(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right]\\
& -trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)\exp\left\{-\frac{2\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\\
& \times\left[2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))+(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right]\\
& +b\ trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)\exp\left\{-\frac{2\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\\
& \times\left[2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))+(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right]\\
& +\exp\left\{-\frac{2\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\\
& \times\left[2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))+(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right]\\
& \times\left[2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))+(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right]\Bigg\}.
\end{aligned}
$$

Consequently, we can write $Y_\tau^i(\boldsymbol{\theta})Y_\tau^j(\boldsymbol{\theta})$ by

$$Y_\tau^i(\boldsymbol{\theta})Y_\tau^j(\boldsymbol{\theta})=a^2\left(\frac{\tau}{2}\right)^2|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau}\{C_1+C_2+C_3+C_4+C_5+C_6+C_7+C_8+C_9\},$$

and thus,

$$E\left[Y_i(\boldsymbol{\theta})Y_j(\boldsymbol{\theta})\right] = a^2\left(\frac{\tau}{2}\right)^2|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau} \times E[C_1 + C_2 + C_3 + C_4 + C_5 + C_6 + C_7 + C_8 + C_9]. \quad \text{(A1)}$$

Now, we are going to calculate the different expectations appearing in Equation A2. We have

$$C_1 = trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\exp\left\{-\frac{2\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}.$$

Therefore,

$$\begin{aligned}
E[C_1] &= trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\\
&\quad\times\int\frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\\
&\quad\times\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{2\tau}\right)^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}d\boldsymbol{y}\\
&= trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\\
&\times\left(\frac{1}{2\tau+1}\right)^{\frac{m}{2}}\int\frac{1}{(2\pi)^{m/2}\left|\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{2\tau+1}\right|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{2\tau+1}\right)^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}d\boldsymbol{y}\\
&= trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\left(\frac{1}{2\tau+1}\right)^{\frac{m}{2}}.
\end{aligned}$$

The expression of $C_2$ is given by

$$\begin{aligned}
C_2 &= -trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)b\\
&\quad\exp\left\{-\frac{2\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\},
\end{aligned}$$

and thus,

$$\begin{aligned}
E[C_2] &= -\frac{\tau}{(1+\tau)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\\
&\quad\times\int\frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\\
&\quad\times\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\tau}\right)^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}d\boldsymbol{y}\\
&= -\frac{\tau}{(1+\tau)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\\
&\quad\times\frac{1}{(1+\tau)^{\frac{m}{2}}}\int\frac{1}{(2\pi)^{m/2}\left|\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\tau+1}\right|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\tau+1}\right)^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}d\boldsymbol{y}\\
&= -\frac{\tau}{(1+\tau)^{m+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right).
\end{aligned}$$

The expression of $C_3$ is given by

$$\begin{aligned} C_3 &= -trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)\exp\left\{-\frac{2\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\} \\ &\times\left[2\left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))+(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right]. \end{aligned}$$

Then,

$$\begin{aligned} E[C_3] &= -trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)\frac{1}{(1+2\tau)^{m/2}} \\ &\times\int\frac{1}{(2\pi)^{m/2}\left|\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{2\tau+1}\right|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\} \\ &\times(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))d\boldsymbol{y} \\ &= -trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\right)\frac{1}{(1+2\tau)^{\frac{m}{2}+1}}. \end{aligned}$$

Related to $C_4$, we have

$$C_4 = -b\ trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\right)\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\},$$

and

$$\begin{aligned} E[C_4] &= -\frac{\tau}{(1+\tau)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\right) \\ &\times\int\frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\} \\ &\times\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T(\boldsymbol{\Sigma}(\boldsymbol{\theta}))^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}d\boldsymbol{y} \\ &= -\frac{\tau}{(1+\tau)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\right) \\ &\times\frac{1}{(1+\tau)^{\frac{m}{2}}}\int\frac{1}{(2\pi)^{m/2}\left|\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\tau+1}\right|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\tau+1}\right)^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}d\boldsymbol{y} \\ &= -\frac{\tau}{(1+\tau)^{m+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\right). \end{aligned}$$

Related to $C_5$, we have

$$C_5 = b^2\ trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\right),$$

and

$$E[C_5] = \frac{\tau^2}{(1+\tau)^{m+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\right).$$

The expression of $C_6$ is given by

$$\begin{aligned} C_6 &= b\, trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\} \\ &\times\left[2\left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))+(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right], \end{aligned}$$

and

$$\begin{aligned} E[C_6] &= \frac{\tau}{(1+\tau)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\right)\frac{1}{(1+\tau)^{\frac{m}{2}+1}} \\ &= \frac{\tau}{(1+\tau)^{m+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\right). \end{aligned}$$

The expression of $C_7$ is

$$\begin{aligned} C_7 &= -trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)\exp\left\{-\frac{2\tau}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\} \\ &\times\left[2\left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))+(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right], \end{aligned}$$

and

$$E[C_7] = -\frac{\tau}{(1+2\tau)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\right).$$

Related to $C_8$,

$$\begin{aligned} C_8 &= b\, trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\} \\ &\times\left[2\left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))+(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right], \end{aligned}$$

and

$$\begin{aligned} E[C_8] &= \frac{\tau}{(1+\tau)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\right)\frac{1}{(1+\tau)^{\frac{m}{2}+1}} \\ &= \frac{\tau}{(1+\tau)^{m+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\right). \end{aligned}$$

Finally,

$$\begin{aligned} C_9 &= \exp\left\{-\frac{2\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\} \\ &\times\left[2\left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))+(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right] \\ &\times\left[2\left(\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))+(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial \theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right]. \end{aligned}$$

Therefore,

$$
\begin{aligned}
C_9 &= \exp\left\{-\frac{2\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\left\{4(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\right. \\
&\quad\times\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta})) \\
&\quad+2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta})) \\
&\quad+2(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta})) \\
&\quad\left.+(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\} \\
&= A_1+A_2+A_3+A_4,
\end{aligned}
$$

and

$$E[C_9]=E[A_1]+E[A_4]$$

because $E[A_2]=E[A_3]=0$. We have

$$
\begin{aligned}
E[A_1] &= E\left[\exp-\frac{2\tau}{2}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))4(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right. \\
&\quad\left.\times\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right] \\
&= 4\int\frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{2\tau+1}\right)^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T \\
&\quad\times\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))d\boldsymbol{y} \\
&= 4\frac{1}{(2\tau+1)^{m/2}}\int\frac{1}{(2\pi)^{m/2}\left|\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{2\tau+1}\right|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{2\tau+1}\right)^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\} \\
&\quad\times(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))d\boldsymbol{y} \\
&= 4\frac{1}{(2\tau+1)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{2\tau+1}\right) \\
&= \frac{4}{(2\tau+1)^{m/2+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right).
\end{aligned}
$$

Finally, we are going to obtain $E[A_4]$.

$$
\begin{aligned}
E[A_4] &= E\left[\exp\left\{-\frac{2\tau}{2}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)\right.\\
&\quad\left.\times(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right]\\
&= \int\frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{2\tau+1}\right)^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\\
&\quad\times\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\\
&\quad\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))d\boldsymbol{y}\\
&= \frac{1}{(2\tau+1)^{m/2}}\left(trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{2\tau+1}\right)\right.\\
&\quad\left.\times\left[\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}+\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right]\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{2\tau+1}\right)\\
&\quad+\frac{1}{(2\tau+1)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{2\tau+1}\right)\\
&\quad\times trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{2\tau+1}\right)\\
&= \frac{1}{(2\tau+1)^{\frac{m}{2}+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\\
&\quad+\frac{1}{(2\tau+1)^{\frac{m}{2}+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\left(\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)^T\right)\\
&\quad+\frac{1}{(2\tau+1)^{\frac{m}{2}+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\\
&= 2\frac{1}{(2\tau+1)^{\frac{m}{2}+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\\
&\quad+\frac{1}{(2\tau+1)^{\frac{m}{2}+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right).
\end{aligned}
$$

Based on the previous results, we have

$$
\begin{aligned}
E\left[Y_\tau^i(\boldsymbol{\theta})Y_\tau^j(\boldsymbol{\theta})\right] &= a^2\left(\frac{\tau}{2}\right)^2|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau} \\
&\left\{\frac{1}{(2\tau+1)^{\frac{m}{2}}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\right. \\
&-\frac{\tau}{(\tau+1)^{m+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
&-\frac{1}{(2\tau+1)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
&-\frac{\tau}{(\tau+1)^{m+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
&+\frac{\tau^2}{(\tau+1)^{m+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
&+\frac{\tau}{(\tau+1)^{m+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
&-\frac{1}{(2\tau+1)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
&+\frac{\tau}{(\tau+1)^{m+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
&+\frac{4}{(2\tau+1)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
&+\frac{2}{(2\tau+1)^{\frac{m}{2}+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
&\left.+\frac{1}{(2\tau+1)^{\frac{m}{2}+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\right\}.
\end{aligned}
$$

The previous expression can be written as

$$
\begin{aligned}
E\left[Y_\tau^i(\boldsymbol{\theta})Y_\tau^j(\boldsymbol{\theta})\right] &= a^2\left(\frac{\tau}{2}\right)^2\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-\tau}\left\{trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\right. \\
&\times\left[\frac{1}{(2\tau+1)^{\frac{m}{2}}}-\frac{1}{(2\tau+1)^{\frac{m}{2}+1}}-\frac{1}{(2\tau+1)^{\frac{m}{2}+1}}+\frac{1}{(2\tau+1)^{\frac{m}{2}+2}}\right] \\
&+trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
&\times\left[-\frac{\tau}{(\tau+1)^{m+1}}-\frac{\tau}{(\tau+1)^{m+1}}+\frac{\tau^2}{(\tau+1)^{m+2}}+\frac{\tau}{(\tau+1)^{m+2}}+\frac{\tau}{(\tau+1)^{m+2}}\right] \\
&+\frac{4}{(2\tau+1)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
&\left.+\frac{2}{(2\tau+1)^{\frac{m}{2}+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\right\}.
\end{aligned}
$$

Therefore,

$$\begin{aligned} E\left[Y_\tau^i(\boldsymbol{\theta})Y_\tau^j(\boldsymbol{\theta})\right] &= \left(\frac{\tau+1}{\tau(2\pi)^{m\tau/2}}\right)^2\left(\frac{\tau}{2}\right)^2|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau} \\ &\times\left\{\frac{4\tau^2}{(2\tau+1)^{\frac{m}{2}+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\right. \\ &-\frac{\tau^2}{(1+\tau)^{m+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\ &+4\frac{4}{(2\tau+1)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\ &\left.+\frac{2}{(2\tau+1)^{\frac{m}{2}+2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\right\}. \end{aligned}$$

Finally,

$$\begin{aligned} E\left[Y_\tau^i(\boldsymbol{\theta})Y_\tau^j(\boldsymbol{\theta})\right] &= (\tau+1)^2\left\{\left(\frac{1}{(2\pi)^{m\tau/2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{1/2}}\right)^{2\tau}\frac{1}{(2\tau+1)^{\frac{m}{2}+2}}\right. \\ &\times\left(\Delta_{2\tau}^i\Delta_{2\tau}^j+(2\tau+1)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right)\right. \\ &\left.+\frac{1}{2}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\right) \\ &\left.-\left(\frac{1}{(2\pi)^{m\tau/2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{1/2}}\right)^{2\tau}\frac{1}{(1+\tau)^{m+2}}\Delta_\tau^i\Delta_\tau^j\right\} \\ &= (\tau+1)^2K_\tau^{ij}(\boldsymbol{\theta}), \end{aligned}$$

where $K_\tau^{ij}(\boldsymbol{\theta})$ was defined in (10). Then,

$$\sqrt{n}\frac{\partial}{\partial\boldsymbol{\theta}}H_n(\boldsymbol{\theta}) \underset{n\longrightarrow\infty}{\xrightarrow{\mathcal{L}}} \mathcal{N}\left(\mathbf{0},(\tau+1)^2\boldsymbol{K}_{\boldsymbol{\tau}}(\boldsymbol{\theta})\right),$$

and

$$\sqrt{n}\left(\frac{1}{\tau+1}\frac{\partial}{\partial\boldsymbol{\theta}}H_n(\boldsymbol{\theta})\right) \underset{n\longrightarrow\infty}{\xrightarrow{\mathcal{L}}} \mathcal{N}(\mathbf{0},\boldsymbol{K}_{\boldsymbol{\tau}}(\boldsymbol{\theta})).$$

□

*Appendix A.3*

**Proof of Proposition 2.** Note that

$$\begin{aligned}
\frac{\partial}{\partial\theta_i}H_n^\tau(\boldsymbol{\theta}) &= -a\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
&\times\left[\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}-b\right] \\
&+a\,|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}\left[\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\left(-\frac{\tau}{2}\right)\right. \\
&\times\left(2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right. \\
&\left.\left.-(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right)\right] \\
&= -a\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
&\times\left[\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}-b\right] \\
&+a\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}\frac{\tau}{2}\left[\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\right. \\
&\times\left(2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right. \\
&\left.\left.\times(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right)\right].
\end{aligned}$$

Therefore,

$$\frac{\partial^2}{\partial\theta_i\partial\theta_j}H_n^\tau(\boldsymbol{\theta}) = \frac{\partial}{\partial\theta_j}L_1^\tau(\boldsymbol{\theta}) + \frac{\partial}{\partial\theta_j}L_2^\tau(\boldsymbol{\theta}),$$

being

$$\begin{aligned}
L_1^\tau(\boldsymbol{\theta}) &= -a\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
&\times\left[\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}-b\right]
\end{aligned}$$

and

$$\begin{aligned}
L_2^\tau(\boldsymbol{\theta}) &= a\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2}\frac{\tau}{2}\left[\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\right. \\
&\times\left(2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right. \\
&\left.\left.\times(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right)\right].
\end{aligned}$$

We are going to obtain $\frac{\partial}{\partial\theta_j}L_1^\tau(\boldsymbol{\theta})$.

$$\begin{aligned}
\frac{\partial}{\partial\theta_j}L_1^\tau(\boldsymbol{\theta}) &= -a\frac{\tau}{2}\left(-\frac{\tau}{2}\right)|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
&\times\left[\frac{1}{n}\sum_{i=1}^n \exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}-b\right] \\
&-a\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2} trace\left(-\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}+\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial^2\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right) \\
&\times\left[\frac{1}{n}\sum_{i=1}^n \exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}-b\right] \\
&-a\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
&\times\left[\frac{1}{n}\sum_{i=1}^n \exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\left(-\frac{\tau}{2}\right)\right. \\
&\times\left(-2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right. \\
&\left.\left.-(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right)\right] \\
&= D_1+D_2+D_3,
\end{aligned}$$

being

$$\begin{aligned}
D_1 &= a\frac{\tau^2}{4}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
&\times\left[\frac{1}{n}\sum_{i=1}^n \exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}-b\right],
\end{aligned}$$

$$\begin{aligned}
D_2 &= -a\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2} trace\left(-\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right. \\
&\left.+\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial^2\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right) \\
&\times\left[\frac{1}{n}\sum_{i=1}^n \exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}-b\right],
\end{aligned}$$

and

$$\begin{aligned}
D_3 &= a\frac{\tau^2}{4}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\tau/2} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
&\times\left[\frac{1}{n}\sum_{i=1}^n \exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\right. \\
&\times\left(-2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right. \\
&\left.\left.-(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right)\right].
\end{aligned}$$

Now, we are going to see some result that will be important in order to obtain convergence in probability of $D_1, D_2$ and $D_3$. □

**Lemma A3.** *We have*

$$l = \frac{1}{n}\sum_{i=1}^{n}\exp\Big\{-\frac{\tau}{2}(\boldsymbol{Y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{Y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))\Big\} \xrightarrow[n\longrightarrow\infty]{\mathcal{P}} \frac{1}{(1+\tau)^{m/2}}.$$

**Proof.** It is clear that

$$\begin{aligned} & l \xrightarrow[n\longrightarrow\infty]{\mathcal{P}} E_{\mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}),\boldsymbol{\Sigma}(\boldsymbol{\theta}))}\left[\exp\left\{-\frac{\tau}{2}(\boldsymbol{Y} - \boldsymbol{\mu}(\boldsymbol{\theta}))^T \left(\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\tau}\right)(\boldsymbol{Y} - \boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\right] \\ = & \int \exp\left\{-\frac{\tau}{2}(\boldsymbol{Y} - \boldsymbol{\mu}(\boldsymbol{\theta}))^T \left(\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\tau}\right)(\boldsymbol{Y} - \boldsymbol{\mu}(\boldsymbol{\theta}))\right\} f_{\mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}),\boldsymbol{\Sigma}(\boldsymbol{\theta}))}(\boldsymbol{y})d\boldsymbol{y} \\ = & \frac{1}{(1+\tau)^{m/2}} \int \frac{1}{(2\pi)^{m/2}} \frac{1}{\left|\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\tau+1}\right|^{1/2}} \exp\left\{-\frac{\tau+1}{2}(\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{\theta}))^T \left(\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\tau}\right)(\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{\theta}))\right\} d\boldsymbol{y} \\ = & \frac{1}{(1+\tau)^{m/2}}. \end{aligned}$$

□

**Lemma A4.** *We have*

$$m = \frac{1}{n}\sum_{i=1}^{n}\exp\Big\{-\frac{\tau}{2}(\boldsymbol{Y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{Y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))\Big\} - b \xrightarrow[n\longrightarrow\infty]{\mathcal{P}} \frac{1}{(1+\tau)^{\frac{m}{2}+1}}.$$

**Proof.** Applying the previous Lemma

$$m \xrightarrow[n\longrightarrow\infty]{\mathcal{P}} \frac{1}{(1+\tau)^{m/2}} - \frac{\tau}{(1+\tau)^{\frac{m}{2}+1}} = \frac{1}{(1+\tau)^{\frac{m}{2}+1}}.$$

□

**Lemma A5.** *If we denote*

$$n = \frac{1}{n}\sum_{i=1}^{n}\exp\Big\{-\frac{\tau}{2}(\boldsymbol{Y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{Y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))\Big\}(\boldsymbol{Y}_i - \boldsymbol{\mu}(\boldsymbol{\theta}))^T \boldsymbol{A}(\boldsymbol{Y}_i - \boldsymbol{\mu}(\boldsymbol{\theta})),$$

*we have*

$$n \xrightarrow[n\longrightarrow\infty]{\mathcal{P}} \frac{trace(\boldsymbol{A}\boldsymbol{\Sigma}(\boldsymbol{\theta}))}{(1+\tau)^{\frac{m}{2}+1}}.$$

**Proof.** It is clear that

$$
\begin{aligned}
& n \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} E_{\mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}),\boldsymbol{\Sigma}(\boldsymbol{\theta}))}\left[\exp\left\{-\frac{1}{2}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\tau}\right)^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\right. \\
& \left.\times(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{A}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right] \\
= \; & \frac{1}{(1+\tau)^{m/2}}\int\frac{1}{(2\pi)^{m/2}}\frac{1}{\left|\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\tau+1}\right|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\tau+1}\right)^{-1}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\} \\
& \times(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{A}(\boldsymbol{y}-\boldsymbol{\mu}(\boldsymbol{\theta}))d\boldsymbol{y} \\
= \; & \frac{1}{(1+\tau)^{m/2}}E_{\mathcal{N}\left(\boldsymbol{\mu}(\boldsymbol{\theta}),\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{1+\tau}\right)}\left[(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{A}(\boldsymbol{Y}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right] \\
= \; & \frac{1}{(1+\tau)^{\frac{m}{2}+1}}trace(\boldsymbol{A}\boldsymbol{\Sigma}(\boldsymbol{\theta})).
\end{aligned}
$$

□

Based on the previous results we have in relation to $D_1$,

$$
D_1 \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} \frac{\tau}{4}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}(1+\tau)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right).
$$

With respect to $D_2$,

$$
\begin{aligned}
& D_2 \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} \frac{1}{(2\pi)^{m/2}}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}\frac{1}{2}\frac{1}{(1+\tau)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
& -\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m/2}}\frac{1}{2}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial^2\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right).
\end{aligned}
$$

In a similar way, we obtain for $D_3$ that

$$
D_3 \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} -\frac{\tau}{4}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)\frac{1}{(1+\tau)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right).
$$

Therefore, we have

$$
\begin{aligned}
& \frac{\partial}{\partial\theta_j}L_1^{\tau}(\boldsymbol{\theta}) \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} \frac{1}{2}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
& -\frac{1}{2}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}}\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial^2\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right).
\end{aligned}
$$

Now, we have

$$\begin{aligned}
\frac{\partial}{\partial\theta_j}L_2^{\tau}(\boldsymbol{\theta}) &= -a\frac{\tau^2}{4}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
&\times\left[\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\left(2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\right.\right. \\
&\left.\left.\times\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))+(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right)\right] \\
&+a\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}\left[\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\right. \\
&\times\left(-\frac{\tau}{2}\right)\left(\frac{\partial}{\partial\theta_j}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right)\left(2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right. \\
&\left.\left.\times(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right)\right] \\
&+a\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}\left[\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\right. \\
&\times\frac{\partial}{\partial\theta_j}\left(2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right) \\
&\left.\left.+(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right)\right] \\
&= C_1+C_2+C_3.
\end{aligned}$$

Now,

$$\begin{aligned}
C_1 &= -a\frac{\tau^2}{4}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)\left[\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\right. \\
&\times\left(2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right. \\
&\left.\left.\times(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right)\right].
\end{aligned}$$

It is clear that

$$C_1 \xrightarrow[n\longrightarrow\infty]{\mathcal{P}} -\frac{\tau}{4}\frac{1}{(1+\tau)^{m/2}}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right).$$

Next,

$$\begin{aligned}
C_2 &= -a\frac{\tau^2}{4}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}(S_1+S_2+S_3+S_4) \\
&= L_1^*+L_2^*+L_3^*+L_4^*.
\end{aligned}$$

First,

$$
\begin{aligned}
L_1^* &= -a\frac{\tau^2}{4}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\} \\
&\quad \times\left(-4(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right),
\end{aligned}
$$

and

$$
L_1^* \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} \frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{\tau}{(1+\tau)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\right).
$$

It is clear that

$$
L_2^* \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} \mathbf{0} \text{ and } L_3^* \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} \mathbf{0}.
$$

Finally,

$$
\begin{aligned}
&L_4^* \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} \frac{\tau+1}{(2\pi)^{m\tau/2}}\frac{\tau}{4}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}\frac{1}{(1+\tau)^{m/2}} \\
&\times\left\{trace\left\{\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)\left[\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}+\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right]\right.\right. \\
&\left.\left.\times\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{1+\tau}\right\}+trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{1+\tau}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\boldsymbol{\Sigma}(\boldsymbol{\theta})}{1+\tau}\right)\right\} \\
&=2\frac{\tau}{4}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\boldsymbol{\Sigma}(\boldsymbol{\theta})\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
&+\frac{\tau}{4}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right).
\end{aligned}
$$

Therefore,

$$
C_2 = L_1^* + L_2^* + L_3^* + L_4^* \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} R,
$$

where

$$
\begin{aligned}
R &= \frac{\tau|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}(1+\tau)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right)^T\right) \\
&\quad +\frac{\tau}{2}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}(1+\tau)^{m/2+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
&\quad +\frac{\tau}{4}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}(1+\tau)^{m/2+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right).
\end{aligned}
$$

Finally,

$$\begin{aligned}
C_3 &= a\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\\
&\times\left\{\left(2\frac{\partial}{\partial\theta_j}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right)\right.\\
&+2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}}{\partial\theta_i}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\\
&-2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right)^T\\
&-2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right)(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\\
&+(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\{-\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\\
&\left.+\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial^2\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}-\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right\}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\\
&= A_1^*+A_2^*+A_3^*+A_4^*+A_5^*.
\end{aligned}$$

It is clear that

$$A_1^*\xrightarrow[n\longrightarrow\infty]{\mathcal{P}}0\ A_2^*\xrightarrow[n\longrightarrow\infty]{\mathcal{P}}0\text{ and }A_4^*\xrightarrow[n\longrightarrow\infty]{\mathcal{P}}0.$$

On the other hand,

$$\begin{aligned}
A_3^* &= -a\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\\
&\left(2\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right),
\end{aligned}$$

and

$$A_3^*\xrightarrow[n\longrightarrow\infty]{\mathcal{P}}-(\tau+1)\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}(1+\tau)^{m/2}}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}.$$

Related to $A_5^*$, we have

$$\begin{aligned}
A_5^* &= a\frac{\tau}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}\frac{1}{n}\sum_{i=1}^{n}\exp\left\{-\frac{\tau}{2}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\}\\
&\times\left\{(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))^T\left[\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right.\right.\\
&\left.\left.+\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial^2\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}-\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right](\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{\theta}))\right\},
\end{aligned}$$

and

$$\begin{aligned}
A_5^*\xrightarrow[n\longrightarrow\infty]{\mathcal{P}}&-\frac{1}{(2\pi)^{m\tau/2}}\frac{1}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}\frac{1}{(1+\tau)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)\\
&+\frac{1}{(2\pi)^{m\tau/2}}\frac{1}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}\frac{1}{(1+\tau)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial^2\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i\partial\theta}\right)\\
&-\frac{1}{(2\pi)^{m\tau/2}}\frac{1}{2}|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}\frac{1}{(1+\tau)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right).
\end{aligned}$$

Therefore,

$$
\begin{aligned}
C_3 \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} & -(\tau+1)\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j} \\
& -\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}}\frac{1}{2}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
& +\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}}\frac{1}{2}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial^2\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right) \\
& -\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{2}\frac{1}{(1+\tau)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right).
\end{aligned}
$$

We are going to join all the previous expressions in order to obtain $\frac{\partial}{\partial\theta_j}L_2^\tau(\boldsymbol{\theta})$,

$$
\begin{aligned}
\frac{\partial}{\partial\theta_j}L_2^\tau(\boldsymbol{\theta}) &= C_1+C_2+C_3 \\
&= C_1+L_1^*+L_2^*+L_3^*+L_4^*+C_3 \\
&= C_1+L_1^*+L_2^*+L_3^*+L_4^*+A_1^*+A_2^*+A_3^*+A_4^*+A_5^*.
\end{aligned}
$$

Then,

$$
\begin{aligned}
\frac{\partial}{\partial\theta_j}L_2^\tau(\boldsymbol{\theta}) \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} & -\frac{\tau}{4}\frac{1}{(1+\tau)^{m/2}}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
& +\frac{\tau|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}(1+\tau)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
& +\frac{\tau}{2}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}(1+\tau)^{\frac{m}{2}+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
& +\frac{\tau}{4}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}(1+\tau)^{m/2+1}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right)trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
& -(\tau+1)\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j} \\
& -\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}}\frac{1}{2}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
& +\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}}\frac{1}{2}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial^2\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right) \\
& -\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{2}\frac{1}{(1+\tau)^{m/2}}trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right).
\end{aligned}
$$

Based on the previous results, we have

$$
\begin{aligned}
\frac{\partial^2}{\partial\theta_i\partial\theta_j}H_n^\tau(\boldsymbol{\theta}) &= \frac{\partial}{\partial\theta_j}L_1^\tau(\boldsymbol{\theta})+\frac{\partial}{\partial\theta_j}L_2^\tau(\boldsymbol{\theta}) \\
&= D_1+D_2+D_3+C_1+C_2+C_3 \\
&= D_1+D_2+D_3+C_1+L_1^*+L_2^*+L_3^*+L_4^* \\
&\quad +A_1^*+A_2^*+A_3^*+A_4^*+A_5^*
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial^2}{\partial\theta_i\partial\theta_j}H_n^\tau(\boldsymbol{\theta}) \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} & \frac{1}{2}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
& -\frac{1}{2}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}}\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial^2\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right) \\
& -\frac{\tau}{4}\frac{1}{(1+\tau)^{m/2}}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) \\
& +\frac{\tau|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}(1+\tau)^{m/2}} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
& +\frac{\tau}{2}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}(1+\tau)^{m/2+1}} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
& +\frac{\tau}{4}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}(1+\tau)^{m/2+1}} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\right) trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
& -(\tau+1)\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}}\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right)^T\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j} \\
& -\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}}\frac{1}{2} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right) \\
& +\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}}\frac{1}{2} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial^2\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right) \\
& -\frac{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-\frac{\tau}{2}}}{(2\pi)^{m\tau/2}}\frac{1}{(1+\tau)^{m/2}}\frac{1}{2} trace\left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})}{\partial\theta_i}\right).
\end{aligned}
$$

After some algebra, we have

$$
\frac{\partial^2}{\partial\theta_i\partial\theta_j}H_n^\tau(\boldsymbol{\theta}) \underset{n\longrightarrow\infty}{\overset{\mathcal{P}}{\longrightarrow}} -(\tau+1)J_\tau^{ij}(\boldsymbol{\theta}).
$$

## References


1. Zhang, T. General Gaussian estimation. *J. Multivar. Anal.* **2019**, *169*, 234–247. [CrossRef]
2. Kapur, J.N. *Maximum Entropy Models in Science and Engineering*; Wiley: New Delhi, India, 1989.
3. Castilla; E.; Zografos, K. On distance-type Gaussian estimation. *J. Multivar. Anal.* **2022**, *188*, 22. [CrossRef]
4. Basu, A.; Harris, I.R.; Hjort, N.L.; Jones, M.C. Robust and efficient estimation by minimising a density power divergence. *Biometrika* **1998**, *85*, 549–559. [CrossRef]
5. Pardo, L. *Statistical Inference Based on Divergence Measures*; CRC Press: Boca Raton, FL, USA, 2006.
6. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 2nd ed.; John Wiley and Sons, Inc.: Hoboken, NJ, USA, 2006.
7. Guiasu, S. *Information Theory with Applications;* McGraw Hill: New York, NY, USA,1977.
8. Basu, A.; Mandal, A.; Martin, N.; Pardo, L. Testing composite hypothesis based on the density power divergence. *Sankhya B* **2018**, *80*, 222–262.
9. Giraitis, L.; Taniguchi, M.; Taqqui, M.S. Asymptotic normality of quadratic forms of martingale difference. *Stat. Inference Stoch. Process.* **2017**, *20*, 315–327.
10. Götze, F.; Tikhomirov, A. Asymptotic distribution of quadratic forms and applications. *J. Theor. Probab.* **2002**, *15*, 423–475.
11. Wu, W.B.; Shao, X. A limit theorem for quadratic forms and its applications. *Econom. Theory* **2007**, *23*, 930–950.
12. Eubank, R.L.; Wang, S. A central limit theorem for the sum of generalized linear and quadratic forms. *Statistics* **1999**, *33*, 85–91.
13. Horváth, L.; Shao, Q.M. Limit theorems for quadratic forms with applications to Whittle's estimate. *Ann. Appl. Probab.* **1999**, *9*, 146–187.
14. Silvey, S D. *Reprinting, Monographs on Statistical Subjects*; Chapman and Hall: London, UK, 1975.
15. Sen, P.K.; Singer, J.M.; Pedroso De Lima, A.C. *From Finite Sample to Asymptotic Methods in Statistics. Cambridge Series in Statistical and Probabilistic Mathematics*; Cambridge University Press: Cambridge, UK, 2009.
16. Ferguson, T.S. *A Course in Large Sample Theory*; Chapman & Hall: London, UK, 1996.

17. Aitchison, J.; Silvey, S.D. Maximum Likelihood Estimation of Parameters Subject to Restraints. *Ann. Math. Stat.* **1958**, *29*, 813–828.
18. Pardo, J.A.; Pardo, L.; Zografos, K. Minimum $\phi$-divergence estimators with constraints in multinomial populations. *J. Stat. Plan. Inference* **2002**, *104*, 221–237. [CrossRef]
19. Ghosh, A. Influence function analysis of the restricted minimum divergence estimators: A general form. *Electron. J. Stat.* **2015**, *9*, 1017–1040. [CrossRef]
20. Jaenada , M.; Pardo, L. Robust statistical inference in generalized linear models based on minimum Renyi's pseudodistance estimators. *Entropy* **2022**, *24*, 18. [CrossRef]
21. Jaenada, M.; Miranda, P.; Pardo, L. Robust test statistics based on restricted minimum Rényi's pseudodistance estimators. *Entropy* **2022**, *24*, 28. [CrossRef]
22. Martin, N. Rao's Score Tests on Correlation Matrices. *arXiv* **2021**, arXiv:2012.14238. [CrossRef]
23. Martin, N. Robust and efficient Breusch-Pagan test-statistic: An application of the beta-score Lagrange multipliers test for non-identically distributed individuals. *arXiv* **2023**, arXiv:2301.07245. [CrossRef]
24. Huang, C.; Liu, H.; Shi, H.; Chen, X.; Xiao, M.; Wang, Z.; Cao, J. Bifurcations in a fractional-order neural network with multiple leakage delays. *Neural Netw.* **2020**, *131*, 115–126. [CrossRef]
25. Huang, C.; Wang, J.; Chen, X.; Cao, J. Bifurcations in a fractional-order BAM neural network with four different delays. *Neural Netw.* **2021**, *141*, 344–354. [CrossRef]
26. Xu, C.; Mu, D.; Liu, Z.; Pang, Y.; Liao, M.; Aouiti, C. New insight into bifurcation of fractional-order 4D neural networks incorporating two different time delays. *Commun. Nonlinear Sci. Numer. Simul.* **2023**, *118*, 107043. [CrossRef]
27. Xu, C.; Zhang, W.; Aouiti, C.; Liu, Z.; Yao, L. Bifurcation insight for a fractional-order stage-structured predator-prey system incorporating mixed time delays. *Math. Methods Appl. Sci.* 2023, *early view*. [CrossRef]
28. Hampel, F.R. Contributions to the Theory of Robust Estimation. Ph.D. Thesis, University of California, Berkeley, CA, USA, 1968.
29. Basu, A.; Chakraborty, S.; Ghosh, A.; Pardo, L. Robust density power divergence based tests in multivariate analysis: A comparative overview of different approaches. *J. Multivar. Anal.* **2022**, *188*, 104846.
30. Basu, A.; Ghosh, A.; Martin, N.; Pardo, L. A robust generalization of the Rao test. *J. Bus. Econ. Stat.* **2022**, *40*, 868–879.
31. Castilla, E.; Jaenada, M.; Pardo, L. Estimation and testing on independent not identically distributed observations based on Rényi's pseudodistances. *IEEE Trans. Inf. Theory* **2022**, *68*, 4588–4609. [CrossRef]
32. Castilla, E.; Jaenada, M.; Martín, N.; Pardo, L. Robust approach for comparing two dependent normal populations through Wald-type tests based on Rényi's pseudodistance estimators. *Stat. Comput.* **2022**, *32*, 34. [CrossRef]
33. Ghosh, A.; Basu, A.; Pardo, L. Robust Wald-type tests under random censoring. *Stat. Med.* **2021**, *40*, 1285–1305. [CrossRef]
34. Menéndez, M.; Morales, D.; Pardo, L.; Vajda, I. Divergence-based estimation and testing of statistical models of classification. *J. Multivar. Anal.* **1995** *54*, 329–354. [CrossRef] [PubMed]
35. Toma, A.; Broniatowski, M. Dual Divergence Estimators and Tests: Robustness Results. *J. Multivar. Anal.* **2011**, *102*, 20–36. [CrossRef] [PubMed]
36. Toma, A.; Leoni-Aubin, S. Robust Tests Based on Divergence Estimators and Sadlepoint Approximations. *J. Multivar. Anal.* **2010**, *101*, 1143–1155. [CrossRef] [PubMed]
37. Qin, Y.; Priebe, C.E. Robust Hypothesis Testing via Lq-Likelihood. *Stat. Sin.* **2017**, *27*, 1793–1813. [CrossRef] [PubMed]
38. Cambanis, S.; Huang, S.; Simons, G. On the theory of elliptically contoured distributions. *J. Multivar. Anal.* **1981**, *11*, 368–385. [CrossRef] [PubMed]
39. Fang, K.T.; Zhang, K.T. *Generalized Multivariate Analysis*; Science Press Beiging: Beiging, China, 1990.
40. Fang, K.T.; Kotz, S.; Ng, K.W. *Symmetric Multivariate and Related Distributions*; Chapman & Hall: London, UK, 1987.
41. Gupta, A.K.; Varga, T. *Elliptically Contoured Models in Statistics*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1993. [CrossRef]
42. Veroniki, A.A. Multivariate Elliptically Contoured Distributions: Estimation-Testing Hypothesis. A Review and an Application. Master's Thesis, Department of Mathematics, University of Ioannina, Ioannina, Greece, 2010. [CrossRef]
43. Lehman, E.L. *Testing Statistical Hypothesis*; John Wiley and Sons: Hoboken, NJ, USA, 1959. [CrossRef]
44. Lindley, D.V. Fiducial distributions and Bayes' theorem. *J. R. Stat. Soc.* **1958**, *20*, 102–107. [CrossRef]
45. Harville, D.A. *Matrix Algebra From a Statistician's Perspective*; Springer: Berlin/Heidelberg, Germany, 1997. [CrossRef]