# High-Dimensional Mahalanobis Distances of Complex Random Vectors 

Deliang Dai ${ }^{1, *}$ (D) and Yuli Liang ${ }^{2(1)}$

1 Department of Economics and Statistics, Linnaeus University, 35195 Växjö, Sweden
2 Department of Statistics, Örebro Univeristy, 70281 Örebro, Sweden; yuli.liang@oru.se

* Correspondence: deliang.dai@lnu.se

Citation: Dai, D.; Liang, Y. High-Dimensional Mahalanobis Distances of Complex Random
Vectors. Mathematics 2021, 9, 1877.
https://doi.org/10.3390/
math9161877

Academic Editor: Lev Klebanov

Received: 27 May 2021
Accepted: 4 August 2021
Published: 7 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:/ / creativecommons.org/licenses/by/ 4.0/).


#### Abstract

In this paper, we investigate the asymptotic distributions of two types of Mahalanobis distance (MD): leave-one-out MD and classical MD with both Gaussian- and non-Gaussian-distributed complex random vectors, when the sample size $n$ and the dimension of variables $p$ increase under a fixed ratio $c=p / n \rightarrow \infty$. We investigate the distributional properties of complex MD when the random samples are independent, but not necessarily identically distributed. Some results regarding the F-matrix $\boldsymbol{F}=S_{2}^{-1} S_{1}$-the product of a sample covariance matrix $S_{1}$ (from the independent variable array $\left(\boldsymbol{b e}\left(Z_{i}\right)_{1 \times n}\right)$ with the inverse of another covariance matrix $S_{2}$ (from the independent variable array $\left.\left(\boldsymbol{Z}_{j \neq i}\right)_{p \times n}\right)$-are used to develop the asymptotic distributions of MDs. We generalize the F-matrix results so that the independence between the two components $S_{1}$ and $S_{2}$ of the F-matrix is not required.


Keywords: Mahalanobis distance; complex random vector; moments of MDs

## 1. Introduction

Mahalanobis distance (MD) is a fundamental statistic in multivariate analysis. It is used to measure the distance between two random vectors or the distance between a random vector and its center of distribution. MD has received wide attention since it was proposed by Mahalanobis [1] in the 1930's. After decades of development, MD has been applied as a distance metric in various research areas, including propensity score analysis, as well as applications in matching [2], classification and discriminant analysis [3]. In the former case, MD is used to calculate the propensity scores as a measurement of the difference between two objects. In the latter case, the differences among the clusters are investigated based on MD. The applications of MD are also related to multivariate calibration [4], psychological analysis [5] and the construction of multivariate process control charts [6]; it is a standard method to assess the similarity between the observations.

As an essential scale for distinguishing objects from each other, the robustness of MD is important for guaranteeing the accuracy of an analysis. Thus, the properties of MD have been investigated in the last decade. Gath and Hayes [7] investigated the bounds for extreme MDs. Dai et al. [8] derived a number of marginal moments of the sample MD and showed that it behaves unexpectedly when $p$ is relatively larger than $n$. Dai and Holgersson [9] examined the marginal behavior of MDs and determined their density functions. With the distribution derived by Dai and Holgersson [9], one can set the inferential boundaries for the MD itself and acquire robust criteria. Beside the developments above, some attention has been paid to the Fisher matrix (F-matrix) and Beta matrix which can be considered as the generalizations of MD.

We provide the definition of the Fisher matrix (see [10]) which is essential here. The Fisher matrices, or simply $F$-matrices, are an ensemble of matrices with two components each. Let $\left\{Z_{j k}, j, k=1,2, \ldots\right\}$ be either both real or both complex random variable arrays. For $p \geq 1$ and $n \geq 1$, let

$$
\begin{aligned}
& \mathbf{Z}=\left(Z_{j k}: 1 \leq j \leq p, 1 \leq k \leq n\right)=\left(\mathbf{Z}_{\cdot 1}, \ldots, \boldsymbol{Z}_{\cdot n}\right) \\
& \boldsymbol{z}_{(i)}=\left(z_{i}: 1 \leq i \leq n, i \neq k\right)=\boldsymbol{Z}_{\cdot i}
\end{aligned}
$$

be the arrays, with column vectors $\left(\boldsymbol{Z}_{\cdot k}\right)$. The $\boldsymbol{z}_{(i)}$ corresponds to the $i$ th observation in $\boldsymbol{Z}$. This observation is isolated from the rest of the observations, so that $\boldsymbol{z}_{(i), i \neq k,(p \times 1)}$ and $Z_{p \times(n-1)}$ are two independent samples. We determine the following matrices:

$$
\begin{aligned}
& \boldsymbol{S}_{1}=\boldsymbol{z}_{(i), i \neq k} \boldsymbol{z}_{(i), i \neq k}^{*}=\boldsymbol{z}_{(i)} \boldsymbol{z}_{(i)}^{*}, \\
& \boldsymbol{S}_{2}=\frac{1}{n} \sum_{k=1}^{n} \boldsymbol{Z}_{\cdot k} \boldsymbol{Z}_{\cdot k}^{*}=\frac{1}{n} \mathbf{Z} \mathbf{Z}^{*}
\end{aligned}
$$

where * stands for complex conjugate transpose. These two matrices are both of size $p \times p$. Then the Fisher matrix (or $F$-matrix) is defined as

$$
F:=S_{1} S_{2}^{-1}
$$

Pillai [11], Pillai and Flury [12] and Bai et al. [13] have studied the spectral properties of the F-distribution. Johansson [14] investigated the central limit theorem (CLT) for the random Hermitian matrices, including the Gaussian unitary ensemble. Guionnet [15] established the CLT for the non-commutative functional of Gaussian large random matrices. Bai et al. [16] obtain the CLT for the linear spectral statistics of the Wigner matrix. Zheng [10] extends the moments of the F-matrix into non-Gaussian circumstances with the assumption that the two components $S_{1}$ and $S_{2}$ of an F-matrix $F=S_{2}^{-1} S_{1}$ are independent. Gallego et al. [17] extended the MD into the condition of a multidimensional Normal distribution and studied the properties of MD for this case.

However, to the best of our knowledge, the properties of MDs with non-Gaussian complex random variables have not been studied much in the literature. In this paper, we derive the CLT for an MD with complex random variables under a more general circumstance. The common restrictions on the random vector, such as independent and identically distributed (i.i.d.) for the random sample, are not required. The independence between the two components $S_{1}$ and $S_{2}$ of the F-matrix $F=S_{2}^{-1} S_{1}$ is relaxed. We investigate the distributional properties of complex MD without assuming normal distribution. The fourth moments of the MDs are allowed to be an arbitrary value.

This paper is organized as follows. In Section 2, we introduce the basic definitions concerning different types of complex random vectors, their covariance matrix, and the corresponding MDs. In Section 3, the first moments of different types of MDs and their distribution properties are given. The connection of leave-one-out MD and classical MD is derived, and their respective asymptotic distributions under general assumptions are investigated. We end up this paper by giving some concluding remarks and discussions in Section 4.

## Some Examples of Mahalanobis Distance in Signal Processing

MD has been applied in different research areas. We give two examples that are related to the applications of MD in signal processing.

Example 1. We illustrate our idea by using an example. In Figure 1, each point corresponds to the real and comlex part of an MD. The circle in red is an inferential boundary with the radius calculated based on MDs. The points that are outside the boundary are considered to be signals, while the points that are lying inside the circle are detected as noise.

This example has been used in some research; for example, in evaluating the capacity of multiple-input, multiple-output (MIMO) wireless communication systems (see [18] for more details). Denote the number of inputs (or transmitters) and the number of outputs (or receivers) of the MIMO wireless communication system by $n_{t}$ and $n_{r}$ respectively, and assume that the channel coefficients are correlated at both the transmitter and the receiver ends. Then the MIMO channel can be represented by an $n_{r} \times n_{t}$ complex random matrix $\boldsymbol{H}$ with corresponding covariance matrices
$\boldsymbol{\Sigma}_{r}$ and $\boldsymbol{\Sigma}_{t}$. For most of the cases, $\boldsymbol{\Sigma}_{r}$ and $\boldsymbol{\Sigma}_{t}$ should be replaced by the sample covariance matrices $\boldsymbol{S}_{r}$ and $\boldsymbol{S}_{t}$ to represent the channel correlations at the receiver and transmitter ends, respectively. The information processed by this random channel is a random quantity which can be measured by MD.


Figure 1. Inferential on the signals based on MD.
Example 2. Zhao et al. [19] use MD in a fuzzy clustering algorithm in order to reduce the effect of noise on image segmentation. They compare the boundaries of clusters calculated by both the Euclidean distance and MD. The boundaries calculated by the Euclidean distance are straight lines which misclassify the observations that diverge far from their cluster center, while the boundaries based on the MDs are curves that fit better with the covariance of the cluster. This example implies that the MD is more accurate with regards to the measure of dissimilarity for image segmentation. It can be extended to the application of clustering when the channel signals are complex random variables.

## 2. Preliminaries

In this section, some important definitions of complex random variables and related concepts are introduced, including the covariance matrix of a complex random vector and the corresponding MDs.

We first define the covariance matrix of a general complex random vector on an $n \times p$ dimension complex random variable $z$ where $p$ is the number of variables and $n$ is the sample size of the dataset, as given in [20]:

Definition 1. Let $z_{j}=\left(z_{1}, \ldots, z_{p}\right) \in \mathbb{C}^{p}, j=1, \ldots, n$ be a complex random vector with known mean $E\left[z_{j}\right]=\boldsymbol{\mu}_{z, j}$ where $\boldsymbol{z}_{j}=\boldsymbol{x}_{j}+i \boldsymbol{y}_{j,}, i=\sqrt{-1}$ and $\boldsymbol{x}_{j}, \boldsymbol{y}_{j} \in \mathbb{R}^{p}$. Let $\boldsymbol{\Gamma}_{p \times p}$ be the covariance matrix and $\boldsymbol{C}_{p \times p}$ be the relation matrix of $\boldsymbol{z}_{j}$ respectively. Then $\boldsymbol{\Gamma}_{p \times p}$ and $\boldsymbol{C}_{p \times p}$ are defined respectively as follows:

$$
\begin{gathered}
\boldsymbol{\Gamma}=E\left[\left(z_{j}-\boldsymbol{\mu}_{z, j}\right)\left(z_{j}-\boldsymbol{\mu}_{z, j}\right)^{*}\right], \\
\boldsymbol{C}=E\left[\left(\boldsymbol{z}_{j}-\boldsymbol{\mu}_{z, j}\right)\left(z_{j}-\boldsymbol{\mu}_{z, j}\right)^{\prime}\right],
\end{gathered}
$$

where' is the transpose.

Definition 1 shows a general presentation of the complex random variables without imposing any distributional assumption on $z_{j}$. If we set the components of $z_{j}$, the random variables $x_{j}$ and $y_{j}$, as normally distributed, then we are in a more familiar context in the complex space: circular symmetricity. The circular symmetricity of a complex normal random variable is an assumption that is used for the standardized form of complex Gaussian-distributed random variables. The definition of circular symmetricity is the following.

Definition 2. A p-dimensional complex random variable $\boldsymbol{z}_{p \times 1}=\boldsymbol{x}_{p \times 1}+i \boldsymbol{y}_{p \times 1}$ is a circularly symmetric complex normal one if the vector vec $[\boldsymbol{x} \boldsymbol{y}]^{\prime}$ is normally distributed, i.e.,

$$
\binom{x_{p \times 1}}{\boldsymbol{y}_{p \times 1}} \sim N\left(\left[\begin{array}{c}
\operatorname{Re} \boldsymbol{\mu}_{z, p \times 1} \\
\operatorname{Im} \boldsymbol{\mu}_{z, p \times 1}
\end{array}\right], \frac{1}{2}\left[\begin{array}{cc}
\operatorname{Re} \Gamma_{z, p \times 1} & -\operatorname{Im} \Gamma_{z, p \times 1} \\
\operatorname{Im} \Gamma_{z, p \times 1} & \operatorname{Re} \Gamma_{z, p \times 1}
\end{array}\right]\right)
$$

where $\boldsymbol{\mu}_{z, p \times 1}=E[\boldsymbol{z}]$ and $\boldsymbol{\Gamma}_{z, p \times p}=E\left[\left(\boldsymbol{z}-\boldsymbol{\mu}_{z}\right)\left(\boldsymbol{z}-\boldsymbol{\mu}_{z}\right)^{*}\right]$, " $R \boldsymbol{e}^{\prime \prime}$ stands for the real part and " $I m$ " stands for the imaginary part.

The circularly symmetric normally distributed complex random variable can be used in the circumstance of a standard normal distribution in real space to simplify the derivations on complex random variables. Based on this condition, we acquire a simplified probability density function (hereinafter p.d.f.) on a complex normal random vector which is presented in the following example.

Example 3. The circularly symmetric complex random vector $\boldsymbol{z}=\left(z_{1}, \ldots, z_{p}\right)^{\prime} \in \mathbb{C}^{p}$ with mean vector $\boldsymbol{\mu}_{z}=\mathbf{0}$ and relation matrix $\mathbf{C}=\mathbf{0}$ has the p.d.f. as follows:

$$
f(\boldsymbol{z})=\frac{1}{\pi^{p}|\boldsymbol{z}|} \exp \left(-\boldsymbol{z}^{*} \boldsymbol{\Gamma}_{z}^{-1} \boldsymbol{z}\right)
$$

where $\boldsymbol{\Gamma}_{z}$ is the covariance matrix of $z$ and " $|$.$| " is the determinant.$
Based on Example 3, one can see the possibility of transforming the expression of a complex random vector into a form consisting of vector multiplication between a complex constant vector and a real random vector. Let $z_{j}$ be a complex random variable; then, the transformation between a complex random variable and its real random variables can be presented as follows:

$$
z_{j}=J\binom{x_{j}}{y_{j}}, \text { where } J=\left(\begin{array}{ll}
1 & i
\end{array}\right)
$$

The transformation offers a different way of inspecting the complex random vector [21]. The complex random vector can be considered as the bivariate form of a real random vector pre-multiplied by a constant vector. This transformation offers another way to present the complex covariance matrix in the form of real random vectors. The idea is presented by the example as follows.

Define matrices $\Gamma_{x x}=E\left[(x-\operatorname{Re} \boldsymbol{\mu})(x-\operatorname{Re} \boldsymbol{\mu})^{\prime}\right], \Gamma_{y y}=E\left[(y-\operatorname{Im} \boldsymbol{\mu})(y-\operatorname{Im} \boldsymbol{\mu})^{\prime}\right]$, $\boldsymbol{\Gamma}_{x y}=E\left[(\boldsymbol{x}-\operatorname{Re} \boldsymbol{\mu})(\boldsymbol{y}-\operatorname{Im} \boldsymbol{\mu})^{\prime}\right]$ and $\boldsymbol{\Gamma}_{y x}=E\left[(\boldsymbol{y}-\operatorname{Im} \boldsymbol{\mu})(\boldsymbol{x}-\operatorname{Re} \boldsymbol{\mu})^{\prime}\right]$. The covariance matrix $\Gamma$ of a $p$-dimensional complex random vector can be represented in the form of real random vectors $x$ and $y$, as follows:

$$
\boldsymbol{\Gamma}_{z, 2 p \times 2 p}=\left(\begin{array}{cc}
\boldsymbol{\Gamma}_{x x} & \boldsymbol{\Gamma}_{x y} \\
\boldsymbol{\Gamma}_{y x} & \boldsymbol{\Gamma}_{y y}
\end{array}\right) .
$$

## 3. MD on Complex Random Variables

In this section, we introduce two types of MDs given in Definitions 5 and 6. We start with the classical MD which is based on a known mean and a known covariance matrix. The definition is given as follows.

Definition 3. The MD of the complex random vector $\boldsymbol{z}_{j} \in \mathbb{C}^{p}, j=1, \ldots, n$ with known mean $\boldsymbol{\mu}$ and known covariance matrix $\boldsymbol{\Gamma}_{z}$ is defined as follows:

$$
\begin{equation*}
D\left(\Gamma_{z}, z_{j}, \mu\right)=\left(z_{j}-\mu\right)^{*} \Gamma_{z}^{-1}\left(z_{j}-\mu\right) \tag{1}
\end{equation*}
$$

There are both real and complex parts for a complex random vector. Thus, we define the corresponding MDs of the two components in a complex random variable.

Definition 4. The MDs on the real part and imaginary part $\boldsymbol{x}_{j}, \boldsymbol{y}_{j} \in \mathbb{R}^{p}, j=1, \ldots, n$ of a complex random vector $z_{j} \in \mathbb{C}^{p}, j=1, \ldots, n$ with known mean $\boldsymbol{\mu}$ and known covariance matrix $\boldsymbol{\Gamma}$.. are defined as follows:

$$
\begin{align*}
& D\left(\boldsymbol{\Gamma}_{x x}, x_{j}, \operatorname{Re} \mu\right)=\left(x_{j}-\operatorname{Re} \mu\right)^{*} \boldsymbol{\Gamma}_{x x}^{-1}\left(x_{j}-\operatorname{Re} \mu\right)  \tag{2}\\
& D\left(\boldsymbol{\Gamma}_{y y}, \boldsymbol{y}_{j}, \operatorname{Im} \boldsymbol{\mu}\right)=\left(y_{j}-\operatorname{Im} \mu\right)^{*} \boldsymbol{\Gamma}_{y y}^{-1}\left(y_{j}-\operatorname{Im} \mu\right) . \tag{3}
\end{align*}
$$

Under the definitions above, we can derive the distribution of the MDs on a complex random vector with known mean and covariance. We present the distribution as follows.

Proposition 1. Let $\mathbf{Z}$ be an $n \times p$ row-wise double array of i.i.d. complex normally distributed random variables with $E\left[\boldsymbol{z}_{. j}\right]=\mathbf{0}, j=1, \ldots, p$ and complex positive definite Hermitian covariance matrix $\boldsymbol{\Gamma}_{z}$. Then we have the distribution of the MD in (1) as $D\left(\boldsymbol{\Gamma}_{z}, \boldsymbol{z}_{i}, \mathbf{0}\right) \sim \chi_{2 p}^{2}$.

Proof. For the proof of this proposition, the reader can refer to reference [22].
The result of Proposition 1 is employed here to derive the moments of the MDs on the real and complex parts below.

Proposition 2. The first moment of $D\left(\boldsymbol{\Gamma}_{z}, \boldsymbol{z}_{j}, \mathbf{0}\right)$ defined in (1) is given as follows:

$$
E\left[D\left(\boldsymbol{\Gamma}_{z}, \boldsymbol{z}_{j}, \mathbf{0}\right)\right]=2 p
$$

Proof. The result follows Proposition 1.
Theorem 1. Set the random variables $\boldsymbol{x}_{i}$ and $\boldsymbol{y}_{i}$ to be normally distributed, and the covariance of $D\left(\boldsymbol{\Gamma}_{x x}, \boldsymbol{x}_{i}, \mathbf{0}\right)$ and $D\left(\boldsymbol{\Gamma}_{y y}, \boldsymbol{y}_{i}, \mathbf{0}\right)$ as given in (2) and (3). Then their distributions are given as follows:

$$
\begin{gathered}
\operatorname{Cov}\left[D\left(\boldsymbol{\Gamma}_{x x}, \boldsymbol{x}_{i}, \mathbf{0}\right), D\left(\boldsymbol{\Gamma}_{y y}, \boldsymbol{y}_{i}, \mathbf{0}\right)\right]=\operatorname{tr}\left[\boldsymbol{\Gamma}_{x x}^{-1} \otimes \boldsymbol{\Gamma}_{y y}^{-1} \Phi\right]-p^{2}, \\
\binom{D\left(\boldsymbol{\Gamma}_{x x}, \boldsymbol{x}_{j}, \mathbf{0}\right)}{D\left(\boldsymbol{\Gamma}_{y y}, \boldsymbol{y}_{j}, \mathbf{0}\right)}
\end{gathered}
$$

where $\otimes$ is the tensor product, $\Phi=E\left(\boldsymbol{x}_{j} \boldsymbol{x}_{j}{ }^{\prime} \otimes \boldsymbol{y}_{j} \boldsymbol{y}_{j}{ }^{\prime}\right)$ and $\operatorname{tr}($.$) stands for trace.$

Proof. Proposition 1 shows $D\left(\boldsymbol{\Gamma}_{y y}, \boldsymbol{y}_{j}, \mathbf{0}\right) \sim \chi_{p}^{2}$ and $D\left(\boldsymbol{\Gamma}_{x x}, \boldsymbol{x}_{j}, \mathbf{0}\right) \sim \chi_{p}^{2}$. In order to derive the covariance of these two MDs, we need to first derive their cross moments.

$$
\begin{aligned}
& E\left(x^{\prime} \boldsymbol{\Gamma}_{x x}^{-1} x y^{\prime} \boldsymbol{\Gamma}_{y y}^{-1} \boldsymbol{y}\right)=\operatorname{Etr}\left[\left(x^{\prime} \otimes \boldsymbol{y}^{\prime}\right)\left(\boldsymbol{\Gamma}_{x x}^{-1} \otimes \boldsymbol{\Gamma}_{y y}^{-1}\right)(\boldsymbol{x} \otimes \boldsymbol{y})\right] \\
= & \operatorname{Etr}\left[\left(\boldsymbol{\Gamma}_{x x}^{-1} \otimes \boldsymbol{\Gamma}_{y y}^{-1}\right)(\boldsymbol{x} \otimes \boldsymbol{y})\left(\boldsymbol{x}^{\prime} \otimes \boldsymbol{y}^{\prime}\right)\right]=\operatorname{Etr}\left[\left(\boldsymbol{\Gamma}_{x x}^{-1} \otimes \boldsymbol{\Gamma}_{y y}^{-1}\right)\left(x x^{\prime} \otimes \boldsymbol{y} \boldsymbol{y}^{\prime}\right)\right] \\
= & \operatorname{tr}\left[\left(\boldsymbol{\Gamma}_{x x}^{-1} \otimes \boldsymbol{\Gamma}_{y y}^{-1}\right) E\left(x x^{\prime} \otimes \boldsymbol{y} y^{\prime}\right)\right] .
\end{aligned}
$$

Thus, the covariance is given as follows:

$$
\begin{aligned}
& \operatorname{Cov}\left[D\left(\boldsymbol{\Gamma}_{x x}, \boldsymbol{x}_{j}, \mathbf{0}\right), D\left(\boldsymbol{\Gamma}_{y y}, \boldsymbol{y}_{j}, \mathbf{0}\right)\right] \\
& =E\left[D\left(\boldsymbol{\Gamma}_{x x}, \boldsymbol{x}_{j}, \mathbf{0}\right) D\left(\boldsymbol{\Gamma}_{y y}, \boldsymbol{y}_{j}, \mathbf{0}\right)\right]-E\left[D\left(\boldsymbol{\Gamma}_{x x}, \boldsymbol{x}_{j}, \mathbf{0}\right)\right] E\left[D\left(\boldsymbol{\Gamma}_{y y}, \boldsymbol{y}_{j}, \mathbf{0}\right)\right] \\
& =E\left[D\left(\boldsymbol{\Gamma}_{x x}, \boldsymbol{x}_{j}, \mathbf{0}\right) D\left(\boldsymbol{\Gamma}_{y y}, \boldsymbol{y}_{j}, \mathbf{0}\right)\right]-p^{2} \\
& =\operatorname{tr}\left[\left(\boldsymbol{\Gamma}_{x x}^{-1} \otimes \boldsymbol{\Gamma}_{y y}^{-1}\right) \boldsymbol{\Phi}\right]-p^{2},
\end{aligned}
$$

which completes the proof.
The results presented so far concern a complex random vector with known mean and known covariance matrix. In practice, the mean $\mu$ and covariance matrix $\Gamma_{z}$ are not available all the time. Thus, some alternative statistics such as sample mean $\overline{\boldsymbol{z}}=n^{-1} \sum_{j=1}^{n} \boldsymbol{z}_{j}$ and sample covariance $S_{z}=n^{-1} \sum_{j=1}^{n}\left(\boldsymbol{z}_{j}-\overline{\boldsymbol{z}}\right)\left(\boldsymbol{z}_{j}-\overline{\boldsymbol{z}}\right)^{*}$ are used as substitutions of the population mean and variance when building the MDs.

We introduce the definitions of MDs with sample mean and sample covariance matrix in the following.

Definition 5. Let $\boldsymbol{z}_{j} \in \mathbb{C}^{p}, j=1, \ldots, n$, be a complex normally distributed random sample. The $M D$ on the complex random vector $\boldsymbol{z}_{j}$ with sample mean $\overline{\boldsymbol{z}}=n^{-1} \sum_{j=1}^{n} z_{j}$ and sample covariance matrix $\boldsymbol{S}_{z}=(n-1)^{-1} \sum_{j=1}^{n}\left(\boldsymbol{z}_{j}-\overline{\boldsymbol{z}}\right)\left(z_{j}-\overline{\boldsymbol{z}}\right)^{*}$ is defined as follows:

$$
\begin{equation*}
D_{j}=D\left(S_{z}, z_{j}, \bar{z}\right)=\left(z_{j}-\bar{z}\right)^{*} S_{z}^{-1}\left(z_{j}-\bar{z}\right) \tag{4}
\end{equation*}
$$

For the purpose of deriving further distributional properties, we give an alternative definition of an MD: the leave-one-out case. Leave-one-out here means that we remove one observation each time and use the rest of the observations for constructing the MD.

Definition 6. Let the leave-one-out sample mean of a complex random vector be $\overline{\boldsymbol{z}}_{(i)}=(n-1)^{-1}$ $\sum_{j=1, j \neq i}^{n} \boldsymbol{z}_{j}$ and the corresponding leave-one-out sample covariance matrix $\boldsymbol{S}_{(i)}=(n-2)^{-1}$ $\sum_{j=1, j \neq i}^{n}\left(z_{j}-\bar{z}_{(i)}\right)\left(z_{j}-\bar{z}_{(i)}\right)^{*}$. The leave-one-out MD on the complex random vector $z_{i}$ is defined as follows:

$$
\begin{equation*}
D_{(i)}=D\left(S_{(i)}, z_{j}, \bar{z}_{(i)}\right)=\left(z_{i}-\bar{z}_{(i)}\right)^{*} S_{(i)}^{-1}\left(z_{i}-\bar{z}_{(i)}\right) . \tag{5}
\end{equation*}
$$

The advantage of leaving one observation out of the dataset is the achieved independence between the removed observation and the sample covariance matrix of the other observations. The similarity in structure between the estimated and the leave-one-out MDs can be explored in theorems on their distributions. The distribution of the leave-one-out MD is derived as in the following theorem.

Theorem 2. $(n-1)^{2} n^{-1} D_{j}$ given in (4) follows a Beta distribution $B(p / 2,(n-p-1) / 2)$.

The proofs of this theorem and others are relegated to Appendix A, so that readers can perceive the results more smoothly. The distribution of a leave-one-out MD can be used to derive the distribution of an estimated MD. We show this in Theorems 3 and 4.

We present the main results of this paper in the following two theorems. The assumption of a normal distribution on the complex random variable is released. Instead, we introduce two more general assumptions. Assume:
(i) The entries of complex random matrix $\boldsymbol{Z}_{n \times p}$ are independent complex random variables, but not necessarily identically distributed with mean 0 and variance 1. Let the 4th moment of the entries have arbitrary value $\beta$. The limiting ratio of their dimensions is $p / n \rightarrow c \in(0,1)$.
(ii) For any $\eta$,

$$
\lim _{n \rightarrow \infty} n^{-1} \eta^{-1} \sum_{j k} E\left|z_{j k}^{(n)}\right|^{4} I\left(\left|z_{j k}^{(n)}\right| \geq \eta \sqrt{n}\right)=0
$$

where $I($.$) is the indicator function. This assumption is a standard Linderberg-type$ condition that guarantees the convergence of the random variable without the assumption of identical distribution.

Theorem 3. Under the assumptions (i) and (ii), set the 4th moment of the complex random vector $E(z)^{4}=\beta<\infty$. Then the asymptotic distribution of $D_{(i)}$ in (5) is:

$$
\lim _{n, p \rightarrow \infty} \sqrt{p} v^{-1 / 2}\left(p^{-1} D_{(i)}-\tau\right) \xrightarrow{\ell} N(0,1)
$$

where $h=\sqrt{p+c-p c}$,

$$
\begin{aligned}
& \tau=\frac{\beta \cdot p}{h^{2}}\left[\sum_{i_{1}+i_{2}+i_{3}=2} \frac{i_{3}!(-1)^{i_{3}} h^{2}(1-c)^{1-i_{1}}}{2\left(1-i_{1}\right)!\left(1-i_{2}\right)!}\left(h^{2}-c\right)^{1-i_{2}}(-c)^{-1-i_{3}}\right. \\
& \left.+\sum_{i_{1}+i_{2}+i_{3}=0} \frac{\left(2+i_{3}\right)!(-1)^{i_{3}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!2} \cdot h^{1+i_{1}-i_{2}}\left(\frac{h}{c}\right)^{3+i_{3}}\right]+\frac{\beta}{2(1-c)} \\
& \times \sum_{i_{1}+\ldots+i_{4}=1} \frac{i_{4}!(-1)^{i_{4}} h^{i_{1}}(1-c)^{1-i_{1}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!}\left(\frac{h^{2}-c}{h}\right)^{1-i_{2}}\left(\frac{\sqrt{c}-c}{h}\right)^{1-i_{3}}\left(-\frac{h}{c}\right)^{1+i_{4}} \\
& +\sum_{i_{1}+\ldots+i_{4}=1} \frac{i_{4}!(-1)^{i_{4}} h^{i_{1}}(1-c)^{1-i_{1}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!}\left(\frac{h^{2}-c}{h}\right)^{1-i_{2}}\left(\frac{-\sqrt{c}-c}{h}\right)^{1-i_{3}}\left(-\frac{h}{c}\right)^{1+i_{4}} \\
& +\sum_{i_{1}+\ldots+i_{4}=0} \frac{\left(i_{4}+1\right)!(-1)^{i_{4}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!} \cdot h^{1+i_{1}-i_{2}}\left(\left(\frac{\sqrt{c}}{h}\right)^{1-i_{3}}+\left(-\frac{\sqrt{c}}{h}\right)^{1-i_{3}}\right)\left(\frac{h}{c}\right)^{2+i_{4}} \\
& -\sum_{i_{1}+\ldots+i_{5}=2} \frac{i_{5}!(1-c)^{1-i_{1}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!(-1)^{i_{4}+2 i_{5}}}\left(h^{2}-c\right)^{1-i_{2}}(\sqrt{c}-c)^{1-i_{3}}(\sqrt{c}+c)^{1-i_{4}} c^{-1-i_{5}} \\
& \left.-\sum_{i_{1}+\ldots+i_{5}=0} \frac{\left(i_{5}+2\right)!(-1)^{i_{5}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!} \cdot h^{1+i_{1}-i_{2}}\left(\frac{\sqrt{c}}{h}\right)^{1-i_{3}}\left(-\frac{\sqrt{c}}{h}\right)^{1-i_{4}}\left(\frac{h}{c}\right)^{3+i_{5}}\right]
\end{aligned}
$$

and

$$
\begin{aligned}
& v=\frac{1}{(1-c)^{4}} \sum_{i_{1}+i_{2}+i_{3}=0} \sum_{j_{1}+j_{2}=2+i_{3}} \frac{\left(i_{3}+1\right)!h^{2+i_{1}-i_{2}+j_{1}-j_{2}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!\left(2+i_{3}\right)!\left(1-j_{1}\right)!\left(1-j_{2}\right)!} \\
& +\frac{\beta(p+c)(1-c)^{2}}{h^{2}}\left[\sum_{i_{1}+i_{2}+i_{3}=1} \frac{\left(i_{3}\right)!(-1)^{i_{3}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!} h^{i_{1}}(1-c)^{1-i_{1}}\left(\frac{h^{2}-c}{h}\right)^{1-i_{2}}\left(-\frac{h}{c}\right)^{1+i_{3}}\right. \\
& \left.+\sum_{i_{1}+i_{2}+i_{3}=0} \frac{\left(1+i_{3}\right)!(-1)^{i_{3}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!} h^{1+i_{1}-i_{2}}\left(\frac{h}{c}\right)^{2+i_{3}}\right]^{2}
\end{aligned}
$$

By using the results above, we derive the asymptotic distribution of the estimated MD in (4).

Theorem 4. The asymptotic distribution of $D_{j}$ in (4) is given as follows:

$$
\sqrt{p} v^{-1 / 2} p^{-1}(n+\tau)^{2} /(n \sqrt{n-1})\left(D_{j}-(n-1) \tau /(n+\tau)^{2}\right) \xrightarrow{\ell} N(0,1)
$$

where $\tau$ and $v$ are given in Theorem 3.
In the F-matrix $F=S_{2}^{-1} S_{1}=S^{-1}\left(z_{j}-\bar{z}\right)\left(z_{j}-\bar{z}\right)^{*}$, the two component $S^{-1}$ and $\left(z_{j}-\bar{z}\right)\left(z_{j}-\bar{z}\right)^{*}$ are assumed to be independent. Theorem 3 is derived under this restriction, while the results in Theorem 4 extend the circumstance and release the assumption of independence in Theorem 3.

## 4. Summary and Conclusions

This paper defines different types of MDs on complex random vectors with either known or unknown mean and covariance matrix. The MDs' first moments and the distributions of MD with known mean and covariance matrix are derived. Further, some asymptotic distributions of the estimated and leave-one-out MDs under complex nonnormal distribution are investigated. We have relaxed several assumptions from our previous work [9]. The random variables in the MD are required to be independent but not necessarily identically distributed. The fourth moment of our random variables can be of arbitrary value. The independence between the two components $S_{2}$ and $S_{1}$ of the F-matrix $\boldsymbol{F}=\boldsymbol{S}_{2}^{-1} \boldsymbol{S}_{1}$ can be ignored in our results.

In conclusion, the MDs on complex random vectors are useful tools when dealing with complex random vectors in many situations, for example, robust analysis with signal processing [23,24]. The asymptotic properties of MDs can be used in some inferential studies for finance theory [25]. Further studies could develop upon the real and imaginary parts of MDs over a complex random sample.

Author Contributions: Conceptualization, D.D.; Methodology, D.D. and Y.L.; Writing-original draft, D.D. and Y.L. All authors have read and agreed to the published version of the manuscript.
Funding: This research received no external funding.
Institutional Review Board Statement: Not applicable.
Informed Consent Statement: Not applicable.
Data Availability Statement: Not applicable.
Acknowledgments: The authors would like to thank the referees and the Academic Editor for their constructive suggestions on how to improve the quality and the presentation of this paper. Yuli Liang's research is supported by the strategic funding of Örebro University.
Conflicts of Interest: The authors declare no conflict of interest.

## Appendix A

## Appendix A.1. Proof of Theorem 2

Proof. Follow the setting in (4). Then Hardin and Rocke [26] shows that $n^{-1}(n-1)^{2} D_{j}$ with a scalar follows a Beta distribution $\operatorname{Beta}(p / 2,(n-p-1) / 2)$.

## Appendix A.2. Proof of Theorem 3

Proof. By using the Theorem 3.2 from [10], we can acquire the mean and covariance of the complex F matrix $F=S_{(i)}^{-1}\left(z_{i}-\bar{z}_{(i)}\right)\left(z_{i}-\bar{z}_{(i)}\right)^{*}$. Set $\beta$ to be arbitrary. Then we receive

$$
\begin{aligned}
& \tau=\lim _{r \downarrow 1} \frac{1}{4 \pi \mathrm{i}} \oint_{|\xi|=1} \frac{(1+h \xi)(\xi+h)}{(1-c)^{2} \cdot \xi}\left(\frac{1}{\xi+r^{-1}}+\frac{1}{\xi-r^{-1}}-\frac{2}{\xi+c / h}\right) \mathrm{d} \xi \\
& +\frac{\beta p(1-c)^{2}}{2 \pi \mathrm{i} \cdot h^{2}} \oint_{|\xi|=1} \frac{(1+h \xi)(\xi+h)}{(1-c)^{2} \cdot \xi} \frac{1}{(\xi+c / h)^{3}} \mathrm{~d} \xi \\
& +\frac{\beta(1-c)}{4 \pi \mathrm{i}} \oint_{|\xi|=1} \frac{(1+h \xi)(\xi+h)}{(1-c)^{2} \cdot \xi} \frac{\xi^{2}-c / h^{2}}{(\xi+c / h)^{2}} \\
& \times\left[\frac{1}{\xi-\sqrt{c} / h}+\frac{1}{\xi+\sqrt{c} / h}-\frac{2}{\xi+c / h}\right] \mathrm{d} \xi \\
& =\frac{\beta p}{h^{2}}\left[\sum_{i_{1}+i_{2}+i_{3}=2} \frac{i_{3}!(-1)^{i_{3}} h^{i_{1}}(1-c)^{1-i_{1}}}{2\left(1-i_{1}\right)!\left(1-i_{2}\right)!}\left(\frac{h^{2}-c}{h}\right)^{1-i_{2}}\left(-\frac{h}{c}\right)^{1+i_{3}}\right. \\
& \left.+\sum_{i_{1}+i_{2}+i_{3}=0} \frac{\left(2+i_{3}\right)!(-1)^{i_{3}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!2} \cdot h^{1+i_{1}-i_{2}}\left(\frac{h}{c}\right)^{3+i_{3}}\right]+\frac{\beta}{2(1-c)} \\
& \times\left[\sum_{i_{1}+\ldots+i_{4}=1} \frac{i_{4}!(-1)^{i_{4}} h^{i_{1}}(1-c)^{1-i_{1}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!}\left(\frac{h^{2}-c}{h}\right)^{1-i_{2}}\left(\frac{\sqrt{c}-c}{h}\right)^{1-i_{3}}\left(-\frac{h}{c}\right)^{1+i_{4}}\right. \\
& +\sum_{i_{1}+\ldots+i_{4}=1} \frac{i_{4}!(-1)^{i_{4}} h^{i_{1}}(1-c)^{1-i_{1}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!}\left(\frac{h^{2}-c}{h}\right)^{1-i_{2}}\left(\frac{-\sqrt{c}-c}{h}\right)^{1-i_{3}}\left(-\frac{h}{c}\right)^{1+i_{4}} \\
& +\sum_{i_{1}+\ldots+i_{4}=0} \frac{\left(i_{4}+1\right)!(-1)^{i_{4}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!} \cdot h^{1+i_{1}-i_{2}}\left(\left(\frac{\sqrt{c}}{h}\right)^{1-i_{3}}+\left(-\frac{\sqrt{c}}{h}\right)^{1-i_{3}}\right)\left(\frac{h}{c}\right)^{2+i_{4}} \\
& -\sum_{\substack{i_{1}+\ldots++i_{5}=2}} \frac{(-1)^{2-i_{4}} i_{5}!(1-c)^{1-i_{1}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!}\left(h^{2}-c\right)^{1-i_{2}}(\sqrt{c}-c)^{1-i_{3}}(\sqrt{c}+c)^{1-i_{4}} c^{-1-i_{5}} \\
& \left.-\sum_{i_{1}+\ldots+i_{5}=0} \frac{\left(i_{5}+2\right)!(-1)^{i_{5}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!} \cdot h^{1+i_{1}-i_{2}}\left(\frac{\sqrt{c}}{h}\right)^{1-i_{3}}\left(-\frac{\sqrt{c}}{h}\right)^{1-i_{4}}\left(\frac{h}{c}\right)^{3+i_{5}}\right] \\
& =\frac{\beta p}{h^{2}}\left[\sum_{i_{1}+i_{2}+i_{3}=2} \frac{i_{3}!(-1)^{i_{3}} h^{2}(1-c)^{1-i_{1}}}{2\left(1-i_{1}\right)!\left(1-i_{2}\right)!}\left(h^{2}-c\right)^{1-i_{2}}(-c)^{-1-i_{3}}\right. \\
& \left.+\sum_{i_{1}+i_{2}+i_{3}=0} \frac{\left(2+i_{3}\right)!(-1)^{i_{3}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!2} \cdot h^{1+i_{1}-i_{2}}\left(\frac{h}{c}\right)^{3+i_{3}}\right]+\frac{\beta}{2(1-c)} \\
& \times\left[\sum_{i_{1}+\ldots+i_{4}=1} \frac{i_{4}!(-1)^{i_{4}} h^{i_{1}}(1-c)^{1-i_{1}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!}\left(\frac{h^{2}-c}{h}\right)^{1-i_{2}}\left(\frac{\sqrt{c}-c}{h}\right)^{1-i_{3}}\left(-\frac{h}{c}\right)^{1+i_{4}}\right. \\
& +\sum_{i_{1}+\ldots+i_{4}=1} \frac{i_{4}!(-1)^{i_{4}} h^{i_{1}}(1-c)^{1-i_{1}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!}\left(\frac{h^{2}-c}{h}\right)^{1-i_{2}}\left(\frac{-\sqrt{c}-c}{h}\right)^{1-i_{3}}\left(-\frac{h}{c}\right)^{1+i_{4}} \\
& +\sum_{i_{1}+\ldots+i_{4}=0} \frac{\left(i_{4}+1\right)!(-1)^{i_{4}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!} \cdot h^{1+i_{1}-i_{2}}\left(\left(\frac{\sqrt{c}}{h}\right)^{1-i_{3}}+\left(-\frac{\sqrt{c}}{h}\right)^{1-i_{3}}\right)\left(\frac{h}{c}\right)^{2+i_{4}}
\end{aligned}
$$

$$
\begin{aligned}
& -\sum_{i_{1}+\ldots+i_{5}=2} \frac{(-1)^{2-i_{4}} i_{5}!(1-c)^{1-i_{1}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!}\left(h^{2}-c\right)^{1-i_{2}}(\sqrt{c}-c)^{1-i_{3}}(\sqrt{c}+c)^{1-i_{4}} c^{-1-i_{5}} \\
& \left.-\sum_{i_{1}+\ldots+i_{5}=0} \frac{\left(i_{5}+2\right)!(-1)^{i_{5}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!} \cdot h^{1+i_{1}-i_{2}}\left(\frac{\sqrt{c}}{h}\right)^{1-i_{3}}\left(-\frac{\sqrt{c}}{h}\right)^{1-i_{4}}\left(\frac{h}{c}\right)^{3+i_{5}}\right] .
\end{aligned}
$$

For the variance, we have

$$
\begin{aligned}
& v=-\frac{1}{4 \pi^{2}} \lim _{r \downarrow 1} \oint \oint_{\left|\xi_{1}\right|=\left|\xi_{2}\right|=1} \frac{\left(1+h \xi_{1}\right)\left(\xi_{1}+h\right)\left(1+h \xi_{2}\right)\left(\xi_{2}+h\right)}{(1-c)^{4}\left(\xi_{1}-r \xi_{2}\right)^{2} \xi_{1} \xi_{2}} \mathrm{~d} \xi_{1} \mathrm{~d} \xi_{2} \\
& -\frac{\beta(p+c)(1-c)^{2}}{4 \pi^{2} h^{2}} \oint_{\left|\xi_{1}\right|=1} \frac{\left(1+h \xi_{1}\right)\left(\xi_{1}+h\right)}{\left(\xi_{1}+c / h\right)^{2} \xi_{1}} \mathrm{~d} \xi_{1} \oint_{\left|\xi_{2}\right|=1} \frac{\left(1+h \xi_{2}\right)\left(\xi_{2}+h\right)}{\left(\xi_{2}+c / h\right)^{2} \xi_{2}} \mathrm{~d} \xi_{2}
\end{aligned}
$$

where

$$
\begin{aligned}
& -\frac{1}{4 \pi^{2}} \lim \oint \oint_{\left|\xi_{1}\right|=\left|\xi_{2}\right|=1} \frac{\left(1+h \xi_{1}\right)\left(\xi_{1}+h\right)\left(1+h \xi_{2}\right)\left(\xi_{2}+h\right)}{(1-c)^{4}\left(\xi_{1}-r \xi_{2}\right)^{2} \xi_{1} \xi_{2}} \mathrm{~d} \xi_{1} \mathrm{~d} \xi_{2} \\
= & -\frac{i}{2} \oint_{\left|\xi_{2}\right|=1} \frac{\left(1+h \xi_{2}\right)\left(\xi_{2}+h\right)}{\xi_{2}}\left(\oint_{\left|\xi_{1}\right|=1} \frac{\left(1+h \xi_{1}\right)\left(\xi_{1}+h\right)}{\xi_{1}\left(\xi_{1}-r \xi_{2}\right)^{2}} \mathrm{~d} \xi_{1}\right) \mathrm{d} \xi_{2} \\
= & -\frac{i}{2 \pi(1-c)^{4}} \sum_{i_{1}+i_{2}+i_{3}=0} \frac{\left(i_{3}+1\right)!h^{1+i_{1}-i_{2}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!} \oint_{\left|\xi_{2}\right|=1} \frac{\left(1+h \xi_{2}\right)\left(\xi_{2}+h\right)}{\xi_{2}^{3+i_{3}}} \mathrm{~d} \xi_{2} \\
= & \frac{1}{(1-c)^{4}} \sum_{i_{1}+i_{2}+i_{3}=0} \sum_{j_{1}+j_{2}=2+i_{3}} \frac{\left(i_{3}+1\right)!h^{2+i_{1}-i_{2}+j_{1}-j_{2}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!\left(2+i_{3}\right)!\left(1-j_{1}\right)!\left(1-j_{2}\right)!},
\end{aligned}
$$

and

$$
\begin{aligned}
& \oint_{\left|\xi_{1}\right|=1} \frac{\left(1+h \xi_{1}\right)\left(\xi_{1}+h\right)}{\left(\xi_{1}+c / h\right)^{2} \xi_{1}} \mathrm{~d} \xi_{1} \\
& =2 \pi \mathrm{i}\left[\sum_{i_{1}+i_{2}+i_{3}=1} \frac{\left(i_{3}\right)!(-1)^{i_{3}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!} h^{i_{1}}(1-c)^{1-i_{1}}\left(\frac{h^{2}-c}{h}\right)^{1-i_{2}}\left(-\frac{h}{c}\right)^{1+i_{3}}\right. \\
& \left.+\sum_{i_{1}+i_{2}+i_{3}=0} \frac{\left(1+i_{3}\right)!(-1)^{i_{3}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!} h^{1+i_{1}-i_{2}}\left(\frac{h}{c}\right)^{2+i_{3}}\right]
\end{aligned}
$$

So we obtain

$$
\begin{aligned}
& v=(1-c)^{-4} \sum_{i_{1}+i_{2}+i_{3}=0} \sum_{j_{1}+j_{2}=2+i_{3}} \frac{\left(i_{3}+1\right)!h^{2+i_{1}-i_{2}+j_{1}-j_{2}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!\left(2+i_{3}\right)!\left(1-j_{1}\right)!\left(1-j_{2}\right)!} \\
& +\beta(p+c)(1-c)^{2} h^{-2} \\
& \times\left[\sum_{i_{1}+i_{2}+i_{3}=1} \frac{\left(i_{3}\right)!(-1)^{i_{3}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!} h^{i_{1}}(1-c)^{1-i_{1}}\left(\frac{h^{2}-c}{h}\right)^{1-i_{2}}\left(-\frac{h}{c}\right)^{1+i_{3}}\right. \\
& \left.+\sum_{i_{1}+i_{2}+i_{3}=0} \frac{\left(1+i_{3}\right)!(-1)^{i_{3}}}{\left(1-i_{1}\right)!\left(1-i_{2}\right)!} h^{1+i_{1}-i_{2}}\left(\frac{h}{c}\right)^{2+i_{3}}\right]^{2},
\end{aligned}
$$

which completes the proof.

## Appendix A.3. Proof of Theorem 4

Proof. Regarding asymptotic distribution discussed, one may ignore the rank-1 matrix $\mathbf{Z Z}^{*}$ in the definition of the sample covariance matrix and define the sample covariance matrix to be $S=n^{-1} \sum_{j=1}^{n}\left(z_{j}-\bar{z}\right)\left(z_{j}-\bar{z}\right)^{*}$. This idea is given in [27]. Using the results from [9], set
$\boldsymbol{W}=n \boldsymbol{S}$ and $\boldsymbol{W}_{-i}=(n-1) \boldsymbol{S}_{-i}=\sum_{j=1, j \neq i}^{n} \boldsymbol{y}_{j} \boldsymbol{y}_{j}{ }^{*}$ where $\boldsymbol{y}_{j}=\left(\boldsymbol{z}_{j}-\overline{\boldsymbol{z}}\right)$. Let $D_{-i}=\boldsymbol{y}_{i}^{*} \boldsymbol{S}_{-i}^{-1} \boldsymbol{y}_{i}$. Then $\left|\boldsymbol{W}_{-i}+\boldsymbol{y}_{i} \boldsymbol{y}^{*}{ }_{i}\right|=\left|\boldsymbol{W}_{-i}\right|\left(1+\boldsymbol{y}^{*}{ }_{i} \boldsymbol{W}_{-i}^{-1} \boldsymbol{y}_{i}\right)$ and $\left|\boldsymbol{W}-\boldsymbol{y}_{i} \boldsymbol{y}^{*}{ }_{i}\right|=|\boldsymbol{W}|\left(1-\boldsymbol{y}^{*}{ }_{i} \boldsymbol{W}^{-1} \boldsymbol{y}_{i}\right)$ so that $\left|\boldsymbol{W}_{-i}+\boldsymbol{y}_{i} \boldsymbol{y}^{*}{ }_{i}\right| /\left|\boldsymbol{W}_{-i}\right|=1+(n-1)^{-1} D_{-i}$ and $\left|\boldsymbol{W}-\boldsymbol{y}_{i} \boldsymbol{y}^{*}{ }_{i}\right| /|\boldsymbol{W}|=1-n^{-1} D_{i}$. Hence $\left(1-c p^{-1} D_{i}\right)\left(1+c p^{-1} D_{-i}\right)=1$. Now $D_{-i}$ is not independent of observation $i$, since $\boldsymbol{X}_{i}$ is included in the sample mean vector $\bar{X}$, and hence does not fully represent the leave-oneout estimator of $D_{i}$. However, using the identity $\left(X_{i}-\bar{X}\right)=((n-1) / n)\left(X_{i}-\bar{X}_{(i)}\right)$ and substituting it into the above expressions, we find that $\left(1-n^{-1} D_{i}\right) \frac{\left(1+D_{(i)} n^{-1}\right)}{\left(1+D_{(i)} n^{-2}\right)}=1$. This can be used to derive properties of $D_{i}$ as a function of $D_{(i)}$, and vice versa. Note that $D_{i}$ is used in the proof to make the derivation easier to follow. This subscript can be replaced by $D_{j}$ in the theorem.

Set $g\left(p^{-1} D_{(i)}\right)=p^{-1}(n-1) D_{(i)} /\left(n+D_{(i)}\right)$; we acquire its first derivative that $\dot{g}=$ $\partial g\left(D_{(i)}\right) / \partial D_{(i)}=p^{-1}(n-1) n /\left(n+D_{(i)}\right)^{2}$. The same applies to the function $g(\tau)$, so that we can acquire $\dot{g}(\tau)=p^{-1}(n-1) n /(n+\tau)^{2}$. With the results in Theorem 3 and the identity from above, we apply Cramer's theorem [28] and derive $g\left(p^{-1} D_{(i)}\right)-g(\tau) \xrightarrow{\ell} N\left(0, \dot{g}^{2}(\tau)\right)$.

$$
\text { As a result, } \sqrt{p} v^{-1 / 2}(n+\tau)^{2} /(n \sqrt{n-1})\left(p^{-1} D_{i}-p^{-1}(n-1) \tau /(n+\tau)^{2}\right) \xrightarrow{\ell} N(0,1)
$$ as $n, p \rightarrow \infty$, which completes the proof.

## References

1. Mahalanobis, P. On tests and measures of group divergence. Int. J. Asiat. Soc. Bengal 1930, 26, 541-588.
2. Stuart, E.A. Matching methods for causal inference: A review and a look forward. Stat. Sci. A Rev. J. Inst. Math. Stat. 2010, $25,1$. [CrossRef] [PubMed]
3. McLachlan, G. Discriminant Analysis and Statistical Pattern Recognition; John Wiley \& Sons: New York, NY, USA, 2004.
4. De Maesschalck, R.; Jouan-Rimbaud, D.; Massart, D.L. The Mahalanobis distance. Chemom. Intell. Lab. Syst. 2000, 50, 1-18. [CrossRef]
5. Schinka, J.A.; Velicer, W.F. Research Methods in Psychology. In Handbook of Psychology; John Wiley \& Sons: Hoboken, NJ, USA, 2003; Volume 2.
6. Cabana, E.; Lillo, R.E. Robust multivariate control chart based on shrinkage for individual observations. J. Qual. Technol. 2021, 1-26. [CrossRef]
7. Gath, E.G.; Hayes, K. Bounds for the largest Mahalanobis distance. Linear Algebra Appl. 2006, 419, 93-106. [CrossRef]
8. Dai, D.; Holgersson, T.; Karlsson, P. Expected and unexpected values of individual Mahalanobis distances. Commun. Stat. Theory Methods 2017, 46, 8999-9006. [CrossRef]
9. Dai, D.; Holgersson, T. High-dimensional CLTs for individual Mahalanobis distances. In Trends and Perspectives in Linear Statistical Inference; Springer: Berlin/Heidelberg, Germany, 2018; pp. 57-68.
10. Zheng, S. Central Limit Theorems for Linear Spectral Statistics of Large Dimensional F-Matrices. Annales de l'IHP Probabilités et Statistiques 2012, 48, 444-476. [CrossRef]
11. Pillai, K.S. Upper percentage points of the largest root of a matrix in multivariate analysis. Biometrika 1967, 54, 189-194. [CrossRef] [PubMed]
12. Pillai, K.; Flury, B.N. Percentage points of the largest characteristic root of the multivariate beta matrix. Commun. Stat. Theory Methods 1984, 13, 2199-2237. [CrossRef]
13. Bai, Z.D.; Yin, Y.Q.; Krishnaiah, P.R. On limiting spectral distribution of product of two random matrices when the underlying distribution is isotropic. J. Multivar. Anal. 1986, 19, 189-200. [CrossRef]
14. Johansson, K. On fluctuations of eigenvalues of random Hermitian matrices. Duke Math. J. 1998, 91, 151-204. [CrossRef]
15. Guionnet, A. Large deviations upper bounds and central limit theorems for non-commutative functionals of Gaussian large random matrices. In Annales de l'Institut Henri Poincare (B) Probability and Statistics; Elsevier: Amsterdam, The Netherlands, 2002; Volume 38, pp. 341-384.
16. Bai, Z.; Yao, J. On the convergence of the spectral empirical process of Wigner matrices. Bernoulli 2005, 11, 1059-1092. [CrossRef]
17. Gallego, G.; Cuevas, C.; Mohedano, R.; Garcia, N. On the Mahalanobis distance classification criterion for multidimensional normal distributions. IEEE Trans. Signal Process. 2013, 61, 4387-4396. [CrossRef]
18. Ratnarajah, T.; Vaillancourt, R. Quadratic forms on complex random matrices and multiple-antenna systems. IEEE Trans. Inf. Theory 2005, 51, 2976-2984. [CrossRef]
19. Zhao, X.; Li, Y.; Zhao, Q. Mahalanobis distance based on fuzzy clustering algorithm for image segmentation. Digit. Signal Process. 2015, 43, 8-16. [CrossRef]
20. Goodman, N.R. Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). Ann. Math. Stat. 1963, 34, 152-177. [CrossRef]
21. Chave, A.D.; Thomson, D.J. Bounded influence magnetotelluric response function estimation. Geophys. J. Int. 2004, 157, 988-1006. [CrossRef]
22. Giri, N. On the complex analogues of T2-and R2-tests. Ann. Math. Stat. 1965, 36, 664-670. [CrossRef]
23. Zoubir, A.M.; Koivunen, V.; Ollila, E.; Muma, M. Robust Statistics for Signal Processing; Cambridge University Press: Cambridge, UK, 2018.
24. Ma, Q.; Qin, S.; Jin, T. Complex Zhang neural networks for complex-variable dynamic quadratic programming. Neurocomputing 2019, 330, 56-69. [CrossRef]
25. Bai, Z.; Silverstein, J.W. Spectral Analysis of Large Dimensional Random Matrices, 2nd ed.; Springer: New York, NY, USA, 2010.
26. Hardin, J.; Rocke, D.M. The distribution of robust distances. J. Comput. Graph. Stat. 2005, 14, 928-946. [CrossRef]
27. Yao, J.; Zheng, S.; Bai, Z. Sample Covariance Matrices and High-Dimensional Data Analysis; Cambridge University Press Cambridge: Cambridge, UK, 2015.
28. Birke, M.; Dette, H. A note on testing the covariance matrix for large dimension. Stat. Probab. Lett. 2005, 74, 281-289. [CrossRef]
