A Novel Multivariate Goodness-of-Fit Test based on Mahalanobis Distance and Its Application in Denoising

Khuram Naveed
Department of Electrical and Computer Engineering,
COMSATS University Islamabad (CUI),
Islamabad, Pakistan
khurram.naveed@comsats.edu.pk

Naveed ur Rehman
Department of Electrical and Computer Engineering,
Aarhus University,
Aarhus, Denmark
naveed.rehman@ece.au.dk

***Abstract*—Existing multivariate goodness of fit (GoF) tests, to check for multivariate data normality, are cumbersome or intractable. Yet, they garner considerable interest in many practical applications. Mostly, the current multivariate GoF approaches are trivial extensions of their univariate counterparts thereby ignoring inter channel signal dependencies that are inherent in multivariate data sets. To address that, we develop a novel multivariate goodness of fit (GoF) test that uses Mahalanobis distance (MD) as a transformation to map multivariate data into a univariate time series. This way, a novel multivariate GoF test is defined based on the premise that EDF of MD computed for multichannel data are distinct. To test for normality, CDF of quadratic transformation of multivariate normal random variables is used as the reference model within the Anderson Darling (AD) statistic. Finally, this test is used on multiple scales to reject multivariate coefficients fitting the normal distribution leading to a novel multivariate signal denoising method.**

***Index Terms*—Multivariate signals, Denoising, Mahalanobis distance, Goodness-of-fit test**

I. Introduction

Goodness-of-fit (GoF) tests measure how well an observed data coincides with the assumed data model [1]. The tests require a measure of fit, known as test statistic, to quantify the difference between the empirical distribution functions of observed data from the assumed model. Then, hypothesis testing framework are employed to check (statistically) whether the observed data belongs to the assumed model or not.

For univariate (single-channel) data, GoF tests employing test statistics based on empirical distribution functions (EDF) are popular in many engineering applications [2] due to their ease of implementation, e.g., Kolmogrov-Smirnov (KS) [3] statistic, Cramer-Von-Mises (CVM) [4] statistic and Anderson Darling (AD) [5] statistic. Since, cumulative distribution function (CDF) for multivariate data are not uniquely defined [10], their extension to multivariate case is not realized. Nevertheless, tests for multivariate normality are in abundance, e.g., chi-square test [6], SW test [7], skewness and kurtosis test [7] etc., yet the following bottlenecks limit their widespread use: i) computational expense or even intractability [11]; ii) extension to distributions other than normal are not available.

Striving for an implementable yet effective multivariate GoF test, an empirical approach is proposed in [11], which computed Mahalanobis distances (MD) [12] from multivariate data and used the global deviation statistic on their EDF. The resulting test was computationally expensive owing to the intense empirical procedure used to estimate reference EDF. Further, the test was highly sensitive to sample size, mostly under performing for smaller data sets thus rendering it problematic in many real world settings.

To address these issues, we propose a theoretically sound yet easy to implement multivariate GoF test that essentially maps the data in multidimensional space of positive real numbers $\mathcal{R}^M$, with $M > 1$, to the unidimensional space $\mathcal{R}_+$ by using the squared Mahalanobis distance (MD) measure. The premise that MDs from multivariate data follow a distinct distribution allows us to define a unique EDF for multivariate data. Based on that, a novel procedure to test multivariate normality in data is formulated by specifying the reference model as a CDF of quadratic transformation of multivariate normal random variables. To quantify the statistical differences between the test and the reference model, within the proposed GOF test, we use the Anderson Darling (AD) statistic which is known for its robustness for shorter data segments as compared to other EDF statistics. In addition to the aforementioned test, we develop a novel multivariate signal denoising method which tests the normality of multivariate coefficients at multiple wavelet scales to reject noise.

The multivariate signal denoising problem is formulated as follows: Let $\mathbf{x}_i \in \mathcal{R}^M$ denote a real multivariate observations with M number of channels where index $i = 1, \cdots, N$ denotes the number of observations (in time). Assuming additive multivariate noise observations are denoted by $\boldsymbol{\psi} \in \mathcal{R}^M$ with positive definite covariance matrix Σ, $\mathbf{x}_i$ is given by

$$\mathbf{x}_i = \mathbf{s}_i + \boldsymbol{\psi}_i, \tag{1}$$

where $\mathbf{s}_i \in \mathcal{R}^M$ denotes the true (free from noise) multivariate observations. Denoising aims to estimate $\mathbf{s}_i$, given $\mathbf{x}_i$ and the multivariate distribution governing noise observations $\boldsymbol{\psi}_i$.

To compute $\tilde{\mathbf{s}}_i$ - an estimate of $\mathbf{s}_i$, a class of multiscale methods exist where the multivariate wavelet denoising (MWD) method performs signal decomposition via Discrete Wavelet Transform (DWT), followed by channel-wise thresholding on resulting scales [13]. More recent multiscale approaches are based on Synchrosqueezed Transform (SST) [14] and Multivariate Empirical Mode Decomposition (MEMD) [15], [16] but these methods ignore inter-channel correlation of noise while thresholding each channel separately. In this work, we address this issue through the joint processing of multiple channels of input data within the proposed multivariate GOF framework that utilizes squared MD measure. The resulting denoising method fully caters for interchannel signal dependencies, thereby outperforming existing denoising approaches. While the efficacy of our new multivariate signal denoising methodology has already been demonstrated in [17], this work (a) seeks to introduce our new methodology to wider signal and image processing community through this flagship conference (a standard practice in the signal processing community); (b) presents an extension of the proposed multivariate denoising framework to multichannel (or color) images. The next couple of sections lay the theoretical foundations for our proposed method(s) in Section 4.

II. Multivariate Goodness-of-Fit Test

Definition 1 (Multivariate EDF)**.** *Let $\mathbf{x}_i \in \mathcal{R}^M$ denote M-variate observations where $i = 1, \cdots, N$, then multivariate EDF $\tilde{\mathcal{E}}(\cdot) : \mathcal{R}^M \rightarrow \mathcal{R}_+$ may be defined as follows:*

$$\begin{aligned}&\tilde{\mathcal{E}}(t_1,\ t_2, \ldots, t_M) = \\ &\frac{1}{T}\sum_{i=1}^{T} \mathbf{1}.(\mathrm{x}_i^{(1)} \leq t_1, \mathrm{x}_i^{(2)} \leq t_2,\ \ldots\ \mathrm{x}_i^{(M)} \leq t_M),\end{aligned} \tag{2}$$

where $\mathbf{x}_i = [\mathrm{x}_i^{(1)}\ \mathrm{x}_i^{(2)}\ \ldots\ \mathrm{x}_i^{(N)}]'$ denotes the ith observation.

The above definition of multivariate EDF (2) is one of the multiple definitions possible for an M-variate data [10]. Owing to this non-uniqueness of multivariate EDF, an approximation of multivariate CDF is devised in [10], [18] in which a symmetric kernel functions $\mathcal{K}(\cdot,\cdot)$ was used to obtain a unique localised cumulative distribution (LCD) $\mathcal{L}(\mathbf{m}, b)$ from the multivariate pdf $g(\mathbf{x})$, as follows

$$\mathcal{L}(\mathbf{m}, b) = \int_{\mathcal{R}^N} g(\mathbf{x})\mathcal{K}(\mathbf{x} - \mathbf{m}, b)d\mathbf{x}, \tag{3}$$

where $\mathcal{K}(\cdot,\cdot) \in \mathcal{R}_+ \rightarrow [0, 1]$ is a symmetric and integrable kernel located at position $\mathbf{m} \in \mathcal{R}^M$ having width b.

A modified CVM statistic $\tilde{\gamma}$ based on LCD (3) was introduced in [10] to devise a normality test, as follows

$$\begin{aligned}\mathcal{H}_0 : &\quad \mathcal{L}(\mathbf{m}, b) \cong \mathcal{L}_0(\mathbf{m}, b) \Rightarrow \tilde{\gamma} < \lambda \\ \mathcal{H}_1 : &\quad \mathcal{L}(\mathbf{m}, b) \ncong \mathcal{L}_0(\mathbf{m}, b) \Rightarrow \tilde{\gamma} \geq \lambda\end{aligned} \tag{4}$$

where $\mathcal{L}_0(\mathbf{m}, b)$ denotes the LCD corresponding to the multivariate normal distribution, $\mathcal{H}_0$ denotes null hypothesis that normal multivariate sample is detected and $\mathcal{H}_0$ denotes the alternate hypothesis. However, the downside of this test is its enormous computational cost required to compute the LCD in (3) that prohibits its practical use.

III. Mahalanobis Distance (MD)

Definition 2. *Let $\mathbf{x}_i = [\mathrm{x}_i^{(1)}, \cdots, \mathrm{x}_i^{(M)}] \in \mathcal{R}^M$ denote a random multivariate observation, having M channels, from a set of N number of observations. The MD for $\mathbf{x}_i$, given mean $\boldsymbol{\mu} \in \mathcal{R}^M$ and covariance matrix Σ, is defined as*

$$\Delta_i = \sqrt{(\mathbf{x}_i - \boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}_i - \boldsymbol{\mu})}. \tag{5}$$

where $'$ denotes vector transpose.

Remark 1. *MD measures the distance of a point in multidimensional space from the mean of its distribution in terms of the number of standard deviations [19].*

Remark 2. *MDs corresponding to data from a multivariate probability distribution function follow a distinct probability distribution [11].*

Remark 3. *Square of the MD* (5) *can be seen as the quadratic transformation of multivariate random observations $\mathbf{x}_i$ through the covariance matrix Σ.*

Definition 3 (Quadratic Transformation of Random Variables)**.** *Let $\mathbf{x}$ denote a real vector of P random observations $\{x_1, \cdots, x_P\}$ with mean $\boldsymbol{\mu} = \mathbf{0}$, then the quadratic transformation of random variables $Q(\mathbf{x})$ is defined as*

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} \tag{6}$$

where $A \in \mathcal{R}^{P \times P}$ is a real, symmetric and positive definite matrix such that $A = A^T > 0$.

IV. Proposed Methodology

In this section, we present a novel signal denoising method for multivariate data; a new multivariate GoF test based on MD and AD statistic which underpins the proposed denoising method is explained first.

A. Multivariate GOF test based on MD and AD statistic

A multivariate GoF test requires a unique definition of multivariate EDF. That is challenging since the typical EDF reprsentation as defined in (2) is not unique for a given multivariate data. While some kind of averaging over all possible EDFs is an option [10], the resulting representation is cumbersome and computationally expensive. To address that, we propose to utilise the distribution of MD as an alternative to the non-unique multidimensional EDF (2) and the unique but cumbersome LCD (3) based on the premise that given a multivariate pdf, the corresponding distributions of MD will be distinct [11]. The EDF based on MD is defined as

Definition 4 (Mahalanobis EDF). *Let $\mathbf{x}_i \in \mathcal{R}^M$ denotes zero-mean multivariate measurements such that $i = 1, \cdots, N$, then the unique EDF $\mathcal{E}(t)$ based on squared Mahalanobis distance is defined as follows*

$$\mathcal{E}(t) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}.(\mathbf{x}_i'\Sigma^{-1}\mathbf{x}_i \leq t), \tag{7}$$

where Σ is the covariance matrix of the observations $\mathbf{x}_i$ and $\mathcal{E}(t) : \mathcal{R}^N \rightarrow \mathcal{R}_+$ denotes the EDF of multivariate data defined over support t.

Given the unique Mahalanobis EDF (7) for multivariate data, the formulation of a multivariate normality test requires the specification of a reference CDF based on the same (MD) transformation of the multivariate normal distribution. Since, squared MD is essentially a quadratic transformation (6) based on the covariance matrix Σ, we propose to use the CDF of the quadratic transformation of multivariate normal random variables as the reference distribution model in our test for normality, given in *Theorem 1*.

Theorem 1. *Let $\boldsymbol{\psi}_i$ denote a vector-valued Gaussian random variable, i.e., $\boldsymbol{\psi}_i \sim \mathcal{N}_M(\mathbf{0}, \Sigma)$, where Σ is symmetric and positive definite (and therefore Σ^{-1} is also symmetric and positive definite). Given that eigenvalues λ of Σ^{-1} are distinct, the CDF $\mathcal{F}_0(t)$ of the quadratic transformation $y = \boldsymbol{\psi}^T \Sigma^{-1} \boldsymbol{\psi}$ is given from [20], as follows*

$$\mathcal{E}_0(t) = \sum_{n=0}^{\infty} (-1)^n \; c_n \frac{t^{\frac{M}{2}+n}}{\Gamma(\frac{M}{2}+n+1)}, \quad 0 < t < \infty \tag{8}$$

where $\Gamma(\cdot)$ is Gamma function and the coefficients c_n for $n = 0$ and $n \geq 1$ are respectively given below

$$c_0 = \prod_{m=0}^{M} (2\lambda_m)^{-\frac{1}{2}}; \quad c_n = \frac{1}{n} \sum_{r=0}^{n-1} h_{n-r} \; c_r, \qquad n \geq 1,$$

where h_m are given as follows

$$h_m = \frac{1}{2} \sum_{m=0}^{M} (2\lambda_m)^{-m}, \qquad n \geq 1.$$

The traditional EDF like definitions for both observed EDF $\mathcal{E}(t)$ and assumed reference CDF $\mathcal{E}_0(t)$ respectively in (7) and (8) enables us to formulate a modified AD statistic for multivariate data, as follows

$$\gamma = \int_{-\infty}^{\infty} \frac{(\mathcal{E}_0(t) - \mathcal{E}(t))^2}{\mathcal{E}_0(t)(1 - \mathcal{E}_0(x))} d\mathcal{E}_0(t), \tag{9}$$

where γ the estimated AD distance. The proposed modifications due to multidimensional extension of AD statistic are apparent in the computable version of (9), as follows

$$\gamma = L - \sum_{i=1}^{L} \frac{(2i-1)}{L} (\ln(\mathcal{E}_0(\mathbf{x}_i^{'} \Sigma^{-1} \mathbf{x}_i)) - \ln(\mathcal{E}_0(\mathbf{x}_{M+1-i}^{'} \Sigma^{-1} \mathbf{x}_{M+1-i}))), \tag{10}$$

which we obtained by substituting the squared Mahalanobis distance (or quadratic form) of $\mathbf{x}_i$ into the numerical form of (9) suggested in [1] for 1D data of sample size L. Since the quadratic transformation $\mathbf{x}_i^{'} \Sigma^{-1} \mathbf{x}_i$ maps the multichannel data to a single channel dataset, the proposed test statistic (10) employs the powerful AD statistic, originally devised for univariate time series data, to design a robust multivariate normality test.

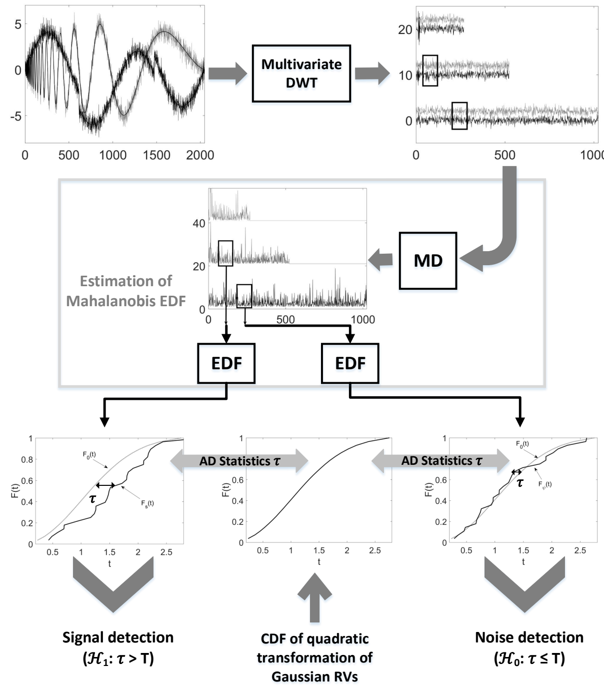


Fig. 1: Illustration of the proposed denoising framework.

Finally, the hypothesis testing framework given in (11) is used to check whether given Mahalanobis EDF $\mathcal{E}(t)$ corresponding to the multivariate observations fit the reference CDF $\mathcal{E}_0(t)$.

$$\begin{aligned} \mathcal{H}_0 : &\quad \mathcal{E}(t) \cong \mathcal{E}_0(t) \Rightarrow \tilde{\gamma} < \lambda \\ \mathcal{H}_1 : &\quad \mathcal{E}(t) \ncong \mathcal{E}_0(t) \Rightarrow \tilde{\gamma} \geq \lambda \end{aligned} \tag{11}$$

where operator $\cong$ and $\ncong$ respectively denote close fit and no fit using AD statistic, $\mathcal{H}_0$ denotes the null hypothesis that multivariate normal observations are detected while $\mathcal{H}_1$ denotes the alternate hypothesis. The threshold λ is selected based on the desired probability of false alarm P_{fa} [21].

B. Multivariate signal denoising based on MD

We now propose a novel multivariate signal denoising method based on multivariate empirical GoF test described in the previous section. GoF tests are typically used in signal detection applications, e.g., spectrum sensing in cognitive radio [21] and denoising of singal channel data [23]–[25] since their framework facilitates detection of noise ($\mathcal{H}_0$) and 'signal + noise' ($\mathcal{H}_1$). For this framework to be applicable to signal denoising, the binary hypothesis testing must be modified to make noise *only* ($\mathcal{H}_0$) and signal *only* ($\mathcal{H}_1$) decisions. That way, data observations corresponding to noise can be discarded while those associated with signal could be retained yielding the denoised signal.

To achieve that, we propose to decompose a signal at multiple scales using the discrete wavelet transform owing to its following properties: i) the distribution of noise in the transform domain is not altered; ii) sparse signal representation

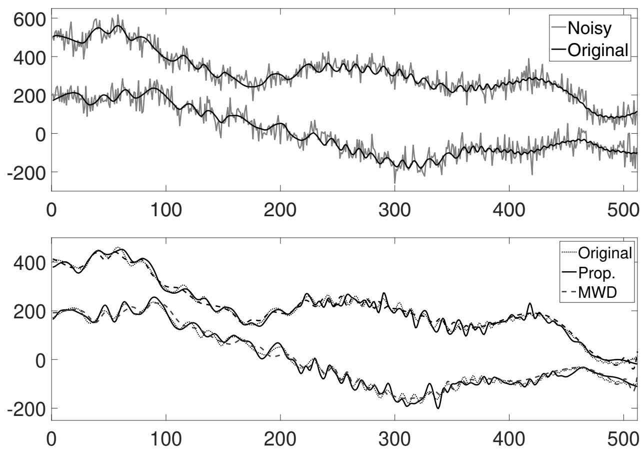


Fig. 2: Original and noisy bivariate SOFAR data (upper) and corresponding denoised signal from MWD and the proposed methods (lower).

at is obtained at multiple scales of DWT thus enabling suitable segregation between noise and signal.

Fig. 1 graphically explains the working of the proposed denoising method based on MD. Firstly, the multiscale decomposition of a multivariate signal is obtained via the DWT, see Fig. 1(top row). Subsequently, multivariate coefficients at multiple scales are divided into small windows leading to estimation of their *Mahalanobis EDF*, see Fig. 1 (middle row). Finally, modified AD statistic (10) is computed for each window using their EDFs estimated in previous steps and the *reference CDF* (8), see Fig. 1 (lower row) where examples of how noise and signal are detected using AD statistics are visually shown along with their corresponding hypothesis.

Let $\mathbf{c}_k$ denote multiscale coefficients at scale index k obtained by applying a transform $\mathcal{T}$ to input multivariate data $\mathbf{x}_i$. The proposed hypothesis testing framework given in (11), applied at multiple data coefficients $\mathbf{c}_k$, was implemented through the following multivariate thresholding function

$$\hat{\mathbf{c}}_k = \begin{cases} 0 & \tilde{\gamma} < \lambda_k, \\ \mathbf{c}_k & \tilde{\gamma} \geq \lambda_k. \end{cases} \tag{12}$$

where λ_k denotes the threshold at scale index k based on the desired probability of false alarm P_{fa} and $\hat{\mathbf{c}}_k$ denotes thresholded coefficients at scale index k.

Finally, inverse transform $\mathcal{T}^{-1}$ is applied to the thresholded coefficients $\hat{\mathbf{c}}_k$ to yield an estimate of the true signal.

V. Results and Discussion

We compare the performance of the proposed multivariate denoising method against the state of the art in multivariate denoising including MWD [13], MWSD [14] and MEMD-IT [15]. The input signals included a bivariate SOFAR data [22] and quadrivariate synthetic signal obtained by combining (1D) ‘Bumps’, ‘Blocks’, ‘Heavy sine’ and ‘Doppler’ signals. Multivariate Gaussian noise was added to the data and the cases of both *balanced noise* (channels corrrupted with noise having same power) and *unbalanced noise* (channels with different noise power) were considered. The proposed method was implemented using $K = 5$ decomposition levels of DWT and Daubechies mother wavelet with 8 vanishing moments. The probability of false alarm and window length were chosen to be $P_{fa} = 0.01$ and $M = 28$ respectively.

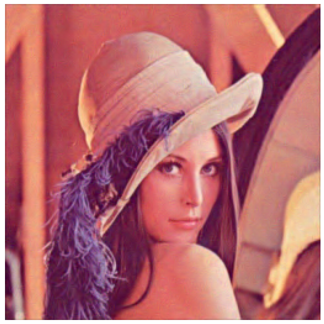
TABLE I: Comparison of the proposed method against comparative methods for SOFAR and synthetic signals for balanced and unbalanced input noise.

Inp. SNR	**-5 dB**	**0 dB**	**5 dB**	**10 dB**
Inp. Signal	**Bivariate Sofar Signal**			
Channel-wise SNR	-5,5 /-3,-7	0,0 /-2,2	5,5 /3, 7	10,10 /8, 12
MWD	7.10/6.48	11.68/11.34	15.26/14.30	17.83/**18.01**
MWSD	-0.33/-0.09	1.22/1.29	2.40/2.30	3.05/ 2.88
MEMD-IT	5.04/2.81	7.80/-2.36	5.11/5.88	7.28/ 8.23
Prop.	**9.15/8.23**	**13.62/12.32**	**15.95/14.55**	**18.02**/17.21
Inp. Signal	**Quadrivariate Synthetic Signal**			
Channel-wise SNR	-5,-5,-5,-5 /-3,-4,-6,-7	0,0,0,0 /-2,-1,1,2	5,5,5,5 /3,4,6,7	10,10,10,10 /12,11,9,8
MWD	6.32/2.86	10.02/6.90	13.18/11.29	16.57/14.12
MWSD	-2.58/-5.01	-1.42/-2.38	-0.92/-1.48	-0.73/-0.96
MEMD-IT	2.44/2.29	8.03/7.62	12.17/10.97	14.67/13.39
Prop.	**6.99/4.77**	**10.15/7.95**	**14.11/11.54**	**17.14/14.38**

We show the noisy and original SOFAR signals in Fig. 2 (upper) while its denoised versions by the proposed method and the MWD method are plotted in Fig. 2 (lower) for comparison. Note that the denoised signal from the proposed method closely follows the true signal (plotted in background) whereas MWD fails to capture the subtle variations of the SOFAR signal.

Table. 1 reports average output reconstructed SNR values (over $J = 10$ iterations) of different denoising methods for input SNRs $= -5, 0, 5, 10$ dB; the values in bold represent the highest reconstructed SNR. Different noise powers in each channel for the case of unbalanced noise are also given in the table; e.g., for the SOFAR signal, for unbalanced noise at averaged input SNR=5dB, the two channels had powers of 3 dB and 7 dB. From the table, it is clear that the proposed method outperforms comparative multivariate denoising methods by yielding highest output SNR at all noise levels for both input signals, except for a single case of unbalanced noise at input SNR $= 10$ dB for the SOFAR signal.

Moreover, we also demonstrate the efficacy of our approach for color image denoising in Fig. 3. For this purpose, the 3D extension of the proposed algorithm is obtained by (i) using the 2D redundant wavelet transform on each of the RGB channels and (ii) employing a 3D window (of size $5 \times 5 \times 3$) for local multiscale application of the proposed multivariate GoF test for normality without losing the spatial dependencies within the image. In this regard, Mahalanobis EDF of the 3D window is estimated by reshaping it into a trivariate segment thus allowing to incorporate the cross-channel-correlations of noise. Rest of the steps in our approach remain unchanged. Observe from the denoised imaged in the righ column of Fig. 3 that proposed method significantly enhances the quality of the noisy images (shown in left column). That is also apparent

(a) In. PSNR = 28.13 dB (b) Out. PSNR = 33.54 dB


(c) In. PSNR = 22.12 dB (d) Out. PSNR = 30.12 dB

Fig. 3: Denoising results on images using proposed denoising framework based on a novel multivariate GoF test.

from the significant improvement in peak SNR (PSNR) values mentioned below each figure.

VI. Conclusions

We have presented a novel multivariate goodness of fit (GoF) test for normality which employs empirical distribution of Mahalanobis distance (MD) as a substitute for the distribution of multivariate data. That is accomplished by specifying the reference distribution as a quadratic transformation of multivariate random variables. The resulting test that also employs a modified robust Anderson Darling test statistic fixes the long standing problem of lack of reliable yet practically convenient test for checking multivariate normality. While the theoretical and computational aspects of the proposed test fall outside the scope of this paper, we have illustrated the utility and potential of the proposed test in signal and image denoising applications.

References

[1] R. B. D'Agostino, Goodness-of-Fit-Techniques, *CRC Press*, 1986.

[2] M. A. Stephens, "EDF statistics for goodness-of-fit and some comparisons," *Journal of the American Statistic Association*, vol. 69, pp. 730-737, 1974.

[3] N. Smirnov, "Table for estimating the goodness of fit of empirical distributions," *Annals of Mathematical Statistics*, vol. 19, pp. 279281, 1948.

[4] H. Cramr, "On the composition of elementary errors," *Scandinavian Actuarial Journal*, 1928.

[5] T. W. Anderson and D. A. Darling, "A Test of Goodness-of-Fit," *Journal of the American Statistic Association*, vol. 49, pp. 765769, 1954.

[6] D. S. Moore and J. B. Stubblebine, "Chi-square tests for multivariate normality with applications to common stock prices," *Communications in Statistics-Theory and Methods*, vol. 10, pp. 713733, 1981.

[7] J. P. Royston, "Some techniques for assessing multivariate normality based on the ShapiroWilk," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 32, pp. 121133, 1983.

[8] J. F. Malkovich and A.A. Afifi, "On tests for multivariate normality," *Journal of the american statistical association*, vol. 68, pp. 176179, 1973.

[9] S. Velilla, "Diagnostics and robust estimation in multivariate data transformation," *Journal of the American Statistical Association*, vol. 90, pp. 945951, 1995.

[10] U. D. Hanebeck, and V. Klumpp, "Localized cumulative distributions and a multivariate generalization of the cramr-von mises distance," *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pp. 33-39, 2008.

[11] M. P. McAssey, "An empirical goodness-of-fit test for multivariate distributions,"*Journal of Applied Statistics*, vol. 40, no. 5, pp. 1120-1131, 2013.

[12] P. C. Mahalanobis, "On the generalised distance in statistics," *Proceedings of the National Institute of Sciences of India*, vol. 2, no. 1, pp. 4955.

[13] M. Aminghafari, N. Cheze and J. M. Poggi, "Multivariate denoising using wavelets and principal component analysis," *Computational Statistics & Data Analysis*, vol. 50, no. 9, pp. 2381-2398, 2006.

[14] A. Ahrabian, and D. P. Mandic, "A Class of Multivariate Denoising Algorithms Based on Synchrosqueezing," *IEEE Transactions on Signal Processing*, vol. 63, no. 9, pp. 2196-2208, 2015.

[15] H. Hao, H. L. Wang, and N. U. Rehman, "A joint framework for multivariate signal denoising using multivariate empirical mode decomposition," *Signal Processing*, vol. 135, pp. 263-273, 2017.

[16] N. ur Rehman, B. Khan, and K. Naveed, "Data-driven multivariate signal denoising using Mahalanobis distance," *IEEE Signal Processing Letters*, vol. 26, no. 9, pp. 14081412, Sep. 2019.

[17] K. Naveed and N. u. Rehman, "Wavelet Based Multivariate Signal Denoising Using Mahalanobis Distance and EDF Statistics,"*IEEE Transactions on Signal Processing*, vol. 68, pp. 5997-6010, 2020.

[18] U. D. Hanebeck, M. F. Huber, V. Klumpp, "Dirac mixture approximation of multivariate gaussian densities," *IEEE International Conference on Decision and Control*, pp. 3851-3858, 2009.

[19] R. D. Maesschalck, D. J. Rimbaud, and D. L. Massart, *The mahalanobis distance*," *Chemometrics and intelligent laboratory systems*, vol. 50, no. 1, pp. 1-18, 2000.

[20] A. M. Mathai and S. B. Provost, Quadratic forms in random variables: Theory and applications. Dekker, 1992.

[21] S. Rostami, K. Arshad, K. Moessner, "Order-Statistic based Spectrum Sensing for Cognitive Radio", *IEEE Communication Letters*, vol. 16, no. 5, pp. 592-595, 2012.

[22] G. Rilling, P. Flandrin, P. Goncalves, J. M. Lilly "Bivariate Empirical Mode Decomposition", *IEEE Signal Processing Letters*, vol. 14, pp. 936-939, 2007.

[23] N. ur Rehman, S. Z. Abbas, A. Asif, A. Javed, K. Naveed, and D. P. Mandic, "Translation invariant multi-scale signal denoising based ongoodness-of-fit tests," *Signal Processing*, vol. 131, pp. 220234, 2017.

[24] K. Naveed, B. Shaukat, S. Ehsan, K. D. Mcdonald-Maier, and N. U. Rehman, "Multiscale image denoising using goodness-of-fit test based on EDF statistics," *PLOS One*, vol. 14, no. 5, pp. e0216197, 2019.

[25] K. Naveed, M. T. Akhtar, M. F. Siddiqui, and N. ur Rehman, "A statistical approach to signal denoising based on data-driven multiscale representation," *Digital Signal Processing*, vol. 108, pp. 102896.