Proceedings of the Pakistan Academy of Sciences:
A: Physical and Computational Sciences 58(1): 61–70 (2021)

ISSN (print) 2518-4245; ISSN (Online) 2518-4253
https://doi.org/10.53560/PPASA(58-1)724

Pakistan Academy of Sciences

Research Article

Automatic Detection of Noisy Signals in sEMG Grids Using Statistical Thresholding

Khalil Ullah[1*], and Khalid Shah[2]

[1]Department of Software Engineering, University of Malakand, Chakdara, Pakistan
[2]Department of Computer Science, Bannu University of Science and Technology, Bannu, Pakistan

Abstract: Electromyogram (EMG) signal is often processed offline, after its acquisition, using digital signal processing algorithms to extract muscle anatomical and physiological information. As most of the signal processing algorithms work on an adequate quality of the signals, thus quality checking of the EMG in real-time during its acquisition is of immense importance. In multi-channel sEMG signals, usually there are some noisy or bad channels. If the noise is of low level, it is of little concern but high level of noise can limit the usefulness of the EMG. To make sure acquisition of a good quality EMG signal in terms of SNR, one way to detect noisy channels is through visual inspection by an expert human operator, however visual inspection of multiple electrodes in real-time is not possible and is also expensive both in terms of time and cost. In this research study, we propose a novel method for automatic detection of noisy channels in multi-channel surface EMG signals based on statistical thresholding of several parameters. The results of the proposed method are in perfect agreement with the ground truth for simulated EMG signals, with an accuracy of 98.6%.

Keywords: EMG, Noisy Channel, Power Line Interference, Statistical Thresholding.

1. INTRODUCTION

In a multi-channel signal detection system, some of the channels are often contaminated by various physiological and non-physiological sources of noise. Typically, these noises come from the acquisition system itself (amplifier noise, saturation, poor electrical contacts), the environment (mains interference, stimulation devices, body vibration etc.) and from the subject (bad skin electrode contact, movement of the skin under electrode, heartbeats, artefactual spikes, Baseline Wander) [1,2]. These noises are inherent in most signal acquisition protocols and are often the limiting factor in the performance of the post EMG signal processing algorithms. In most signal processing techniques, an adequate quality of the sEMG signal is assumed, if this assumption is not met, may lead to invalid extraction of the muscle physiological and anatomical features. Thus, sEMG quality checking is required to have valid signals for further processing and analysis. sEMG quality checking can be performed by an expert human operator, however human expert involvement is sometime impractical in real applications as it is time consuming and not cost effective. Thus, an algorithm is required for automatic detection of these artifacts.

Literature review reveals significant research activities for automatic detection of artifacts in sEMG signals. Wavelet analysis has been used to detect and remove the artifacts from bio-signals [3-6], however due to its computationally intensive nature; wavelet transform can’t be used in real time situations [7]. Independent Component Analysis (ICA) is also being used to detect and isolate artifacts from EMG signals [8-10]. Higher Order Statistics (HOS), and Empirical mode decomposition (EMD) are some other approaches used to detect and remove artifacts from bio- signals [11,13]. Due to high computational cost and error in case of signal

Received: December 2020; Accepted: March 2021
*Corresponding Author: Khalil Ullah <khalil.ullah@uom.edu.pk>

distorted by Colored-Gaussian noise, these methods can't be used. Some of the above methods focus on detection of a specific artifact in EMG signals while others are application dependent. Most of these methods are also based on supervised learning for classification of artifacts. As supervised methods require training of the algorithm for different scenarios, it is time consuming and may not be real time applicable [14,15]. In [16], the authors used a supervised method to detect artifacts in EEG signals, however this method detects time epochs of the EEG signals and is also not applied to multi-channel EMG.

Due to the random nature of sEMG signals, an alternate method, to detect automatically the bad channels, is to use a thresholding of various statistical parameters. Our proposed method is an unsupervised method and it uses various statistical parameters for automatic detection of bad channels. As it is an unsupervised method, it could be applied in real- time without need of training of any classifier. The statistical parameters used in our method are Mean Correlation, Root Mean Square, Hurst Exponent, and Complexity Coefficient. Adaptive thresholding of these parameters has been used to classify a channel as "Good" or "Bad". Our proposed method is a high accuracy of 99.2% by detecting both the noisy and clean signals automatically. The primary contribution of this research work are: 1) the selection of a new set of parameters which can differentiate noise from EMG signal, 2) Use of statistical thresholding to distinguish between a clean signal and the one contaminated by some noise sources.

2. MATERIALS AND METHODS

The proposed algorithm depicted as a flow chart in Figure 1, computes and checks various statistical parameters to verify the quality of the EMG signal in each channel of a multi-channel acquisition system. The surface EMG recordings are affected by several artifacts such as power line interference (PLI), abrupt baseline drift due to movement of patient and skin-electrode impedance, ECG signal, electrode artifacts and amplifier saturation [17,18]. A channel of an electrode array that has any of the above-mentioned artifacts larger in amplitude than EMG with SNR less than 15dB is termed as Bad Channel. A Flag of each parameter is determined for each channel. The value of the Flag of each parameter is set either to 0 (for bad channel) or 1 (for good channel). As we have four parameters, so four flags are obtained for each channel. The values of all these flags are added together and if this sum is greater than two (majority voting) then it means that more than two parameters have identified this channel as good, thus the channel is classified as good channel. Similarly, if a channel is detected as bad by majority of the parameters then the sum of its parameter's Flags is less than two and this channel is classified as bad channel. We use an adaptive thresholding for each parameter after analysis of 320 simulated EMG channels. The selection of threshold for each parameter is discussed in next subsections.

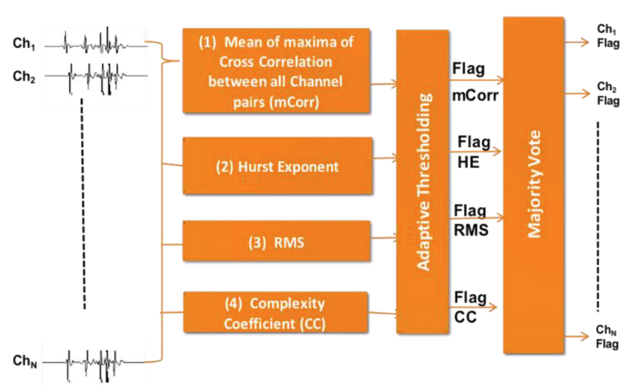

Fig. 1. Flow chart of the proposed method for identification of bad channels based on statistical parameters 1) mean of the maxima of cross correlation of each channel with all other channels, 2) Hurst exponent of each channel, 3) Root mean square value, and 4) The complexity Coefficient of each channel

2.1 Statistical Parameters

The parameters computed and analyzed for each channel are 1) mean of the maxima of the cross-correlation functions (mCorr) of each channel with all the other channels, 2) Hurst Exponent (HE), 3) Root mean square value (RMS) and 4) Complexity Coefficient(CC). The algorithm calculates all the above four parameters for each channel and apply a threshold for each parameter for each channel.

2.1.1 Mean of Maxima of Cross-Correlation

The first parameter used for automatic detection of the bad sEMG channels is the mean correlation (mCorr), which is the average of the maxima of cross correlation of each channel with all other channels. For EMG signal of channel 1 i.e. X¬1, out of the total 16 Channels, the mCorr is computed as follow.

$$mCorr(1) = \frac{1}{16}\sum_{i=1}^{16}(\max\left(R_{X_1 X_i}(\tau)\right)) \qquad (1)$$

Where $R_{X_1 X_i}(\tau)$ is the cross-correlation of channel X1 with channel Xi. This mCorr is thus computed for each channel 2,3 and so on up to 16.

For muscles, parallel to the skin, the sEMG Channels in a high-density system are highly correlated with neighboring channels therefore a channel with artifacts will have a lower correlation with other channels [19] and thus, the mean correlation of that channel will also be lower. Also, if in worst case most of the channels are noisy then still there will be low correlation between them as the noises are mostly uncorrelated and random. An example of a 16-channel single differential (SD) simulated EMG signals (simulated using the planer Model described in [1]) are shown in Figure 2a. The channels 3 and 14 are respectively contaminated by ECG artifact and electrode movement artifact with SNR of 8dB and 5dB respectively. The cross-correlation of all the channels with all other channels and the mCorr of each channel are shown respectively in Figure 2b and c. From the output in Figure 2b and c it is evident that the cross-correlation and the mean correlation i.e. mCorr of the contaminated channels (channel 3 and 14) is lower than the other channels and can easily be detected automatically by applying a proper threshold. The threshold is determined by finding a boundary condition between good and bad (noisy) channels for a total of 320 simulated channels (100 bad and 220 good channels) as shown in Figure 3. The optimal boundary condition obtained in our case was mCorr = 0.5. The channels having mCorr below this value are detected as bad channels.

2.1.2 Root Mean Square Value

Root Mean Square (RMS) value is commonly used as EMG amplitude indicator. RMS of the EMG signal usually ranges from 0 to 1.5 mV [20]. As most of the noises are additive in nature thus the noisy EMG channels have a higher RMS value. For example, due to sudden spikes, additive white Gaussian noise and movement artifacts the RMS value of the EMG channels significantly increases. The RMS of the EMG signal of a channel i is calculated as follow.

$$RMS = \left(\frac{1}{N}\sum_{1}^{N} X_i^{\,2}(n)\right)^{\frac{1}{2}} \qquad (2)$$

Here, N is the total number of samples of the channel i. As in our case the signal is simulated for 3 seconds with sampling frequency of 2048 samples/s, so the total number of samples (N) are 6144.

For a simulated SD EMG signal with four channels 2, 4, 8 and 13 contaminated with ECG artifact, electrode movement artifact, PLI and a mixture of PLI and low frequency noise respectively with SNR of 5dB, the RMS value of each channel are computed using eq. 2 as shown in Figure 4. It is found that the noisy channels appear as outlier in the corresponding histogram of the RMS values. Thus the channels which have RMS higher than the mean+2σ of the RMS of all the channels, is classified as bad channel i.e. outlier (see Figure 4b).

2.1.3 Hurst Exponent

Hurst Exponent (HE) is a parameter used to check the randomness of a signals and is also a measure of the long range dependence with in a signal [21], [22]. It is also considered as a measure of self-similarity. Self-similarity means that the random signals like EMG looks similar if it is zoomed in time in and out [23], like fractal index. Various algorithms are available for the estimation of the HE.

In this study, the Hurst Exponent is estimated using absolute moment method described in [23] and shown in Figure 5. Let X is one channel EMG signal of length N that is divided into M subseries each of length k such that the total number of subseries K = N/k. From each subseries an aggregate series is calculated as,

$$X_m^{(k)} = \frac{1}{k} \sum_{j=(m-1)k+1}^{mk} X_j, \quad m = 1,2,\dots,K \tag{3}$$

The Hurst Exponent is then approximated as,

$$H_m = \frac{1}{K} \sum_{m=1}^{K} \left| X_m^k - \overline{X^k} \right| \tag{4}$$

Here, $\overline{X^k}$ is the mean of the subseries.

To obtain a threshold for distinguishing between a good and a bad channel, we compute H for a total of 20 sets of simulated signals, each containing a total of 16 channel SD EMG signals. Randomly selected 5 out of 16 channels from all 20 sets are contaminated with one of the noises from PLI, ECG artifact, movement artifact, low frequency noise, Guassian and Colored Gaussian, with their SNRs varying from 10dB down to

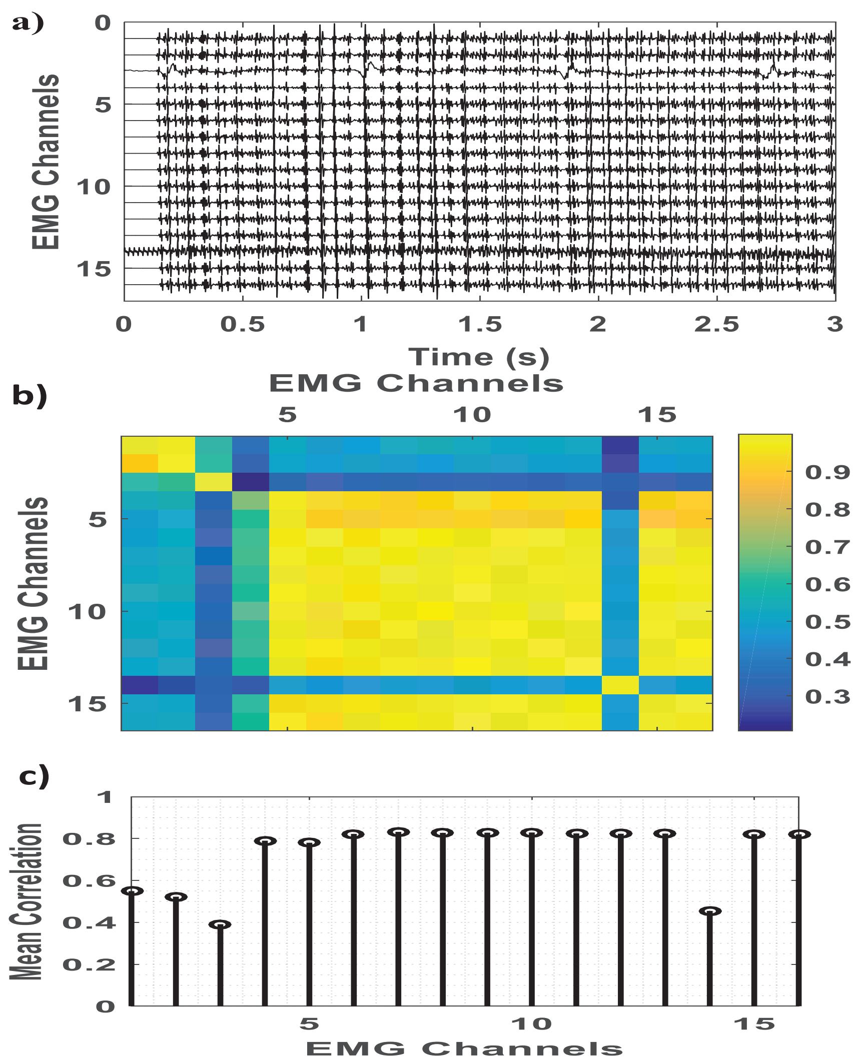

Fig 2. a) Simulated EMG signals with channels 3 and 14 contaminated with ECG artifact and electrode movement with SNR of 8dB and 5dB respectively, b) The maximum of cross correlation of each channel with all other 16 channels i.e. the correlation matrix. It can be seen from the correlation.

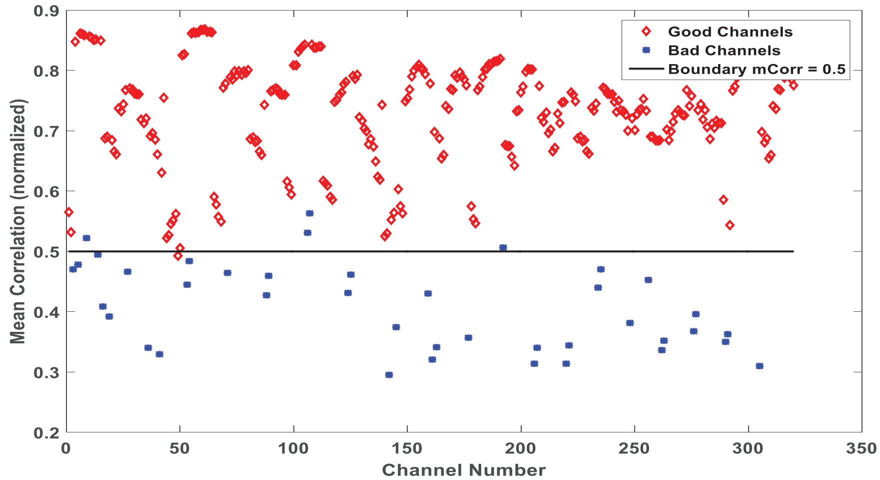

Fig. 3. Mean Correlation coefficient computed for a total of 320 single differential EMG channels. It is clear from the Figure that the threshold for Mean correlation coefficient between bad and good channels is 0.5.

5dB. The threshold for detection of bad and good channels is obtained by computing the mean of the maximum and minimum values of H for all the 16 channels from the total 20 sets of signals. By doing so, initial clusters of both good and bad channels are obtained. Now, to convert this static boundary into a dynamic boundary, another threshold is computed as the average of the minimum value of one cluster with that of maximum of the other cluster. By doing so, the difference of the distance between the threshold and the two clusters is enlarged. This process is repeated until there is no further change in the threshold value. This optimization of the threshold is shown in Figure 6. Once this threshold is optimized, then the channels having H values greater than this threshold are classified as bad channels and the channels having H value lower than this threshold are classified as good channels.

2.1.4 Complexity Coefficient

Complexity coefficient (CCx) is another feature to characterize a signal. It is defined as the ratio of the mobility coefficient ($MC_x = \sigma_{x'}^2/\sigma_x^2$) of the derivative of the signal to the mobility coefficient of the signal itself. For a signal x(t) the complexity coefficient is computed as,

$$CC_x = \frac{MC_{x'}}{MC_x} \tag{5}$$

CCx is constant for a single sinusoid and is independent of the frequency and peak amplitude. We also empirically observed that CCX is almost constant for simulated EMG signals (see Figure 7a). If the EMG signal is contaminated

It is clear from the histogram that the RMS of the four noisy channels appear as outlier i.e. the RMS are away from mean (0.142) by more than 2σ. with a noise, the value of CCX changes due to change in the corresponding mobility coefficient of the signal and its derivative. Thus, the CCX is sensitive to noise and is chosen.

As one parameter for the detection of bad channels in EMG signals. From Figure 7a, the CC for clean EMG signals is 1.5. As noise is introduced to the simulated signals the CC value deviates from 1.5 (See Figure 7b). The threshold selected after analysis of 320 channels of simulated EMG signals is 2. Any channel with CC value greater than this threshold is marked as bad channel.

3. RESULTS AND DISCUSSION

In this study a new method for automatic detection of bad channels, based on thresholding of four statistical parameters as discussed in section 2, is proposed. To check the performance of the proposed method, 2 sets of simulated single differential EMG signal each consisting of multi-channels are prepared. The first set of EMG signals, consists of 20 simulated single differential EMG signals (each consisting of 16 channels and total of 320 channels), is generated using the planar model developed by Merletti et al [1]. The 2nd set

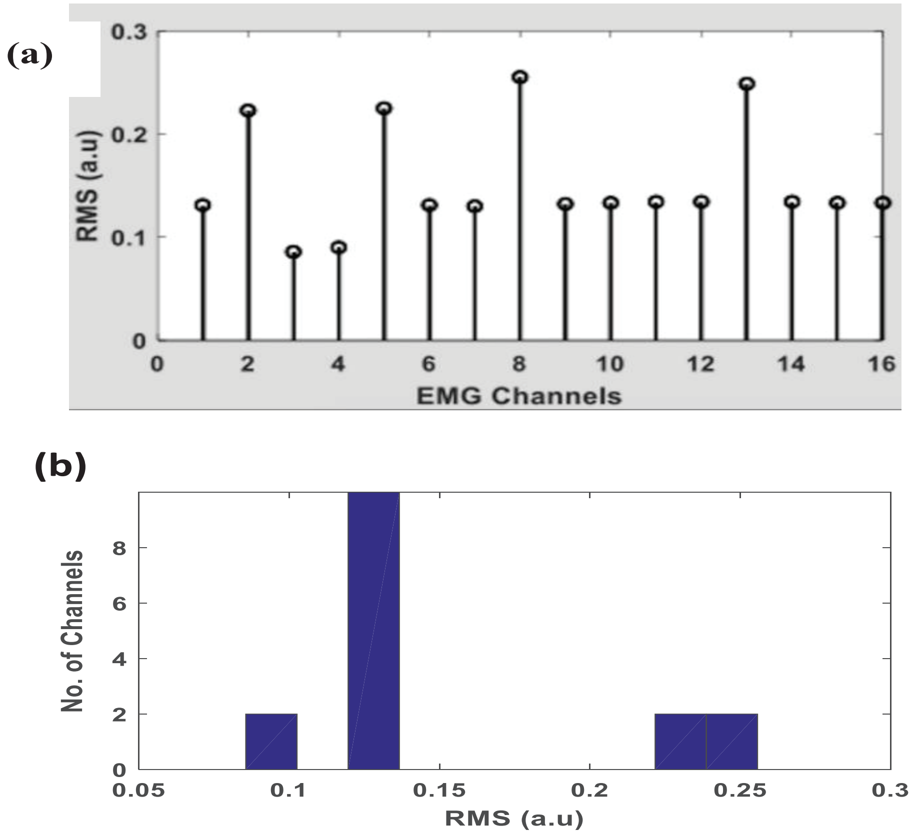

Fig. 4. a) RMS value of each channel of Simulated EMG signal with channel 2 contaminated by ECG artifact, channel 5 with Electrode Moment artifact, channel 8 with PLI and 13 with PLI and low frequency noise with an SNR of 5dB, b)

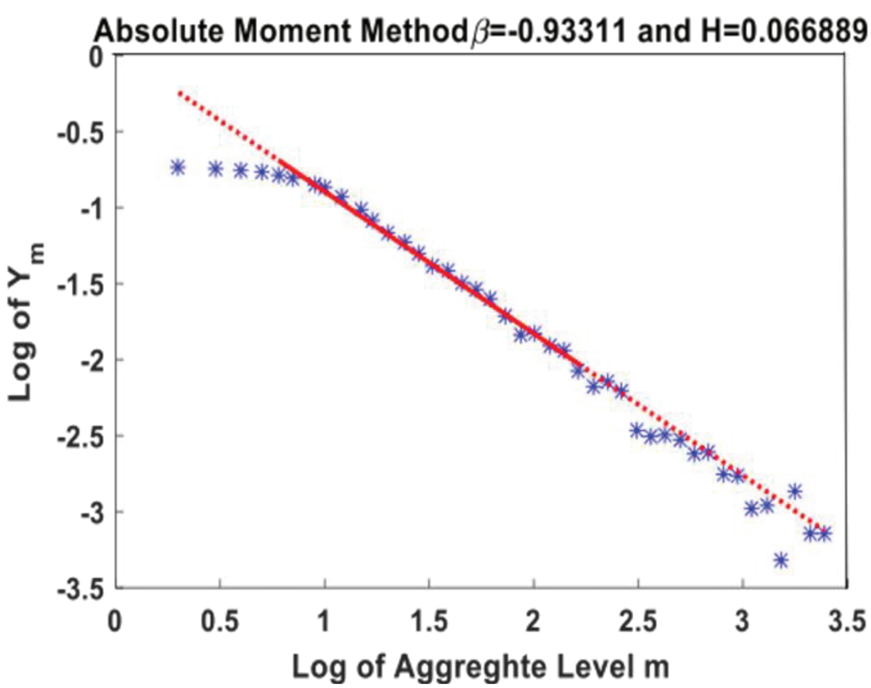

Fig. 5. The logarithmic plot between m and Ym and the fitted line. It should be noted that the slope of this line added with one gives an estimate of the H, the aggregate level m is the size of the non-overlapping blocks which could be from 1 to K.

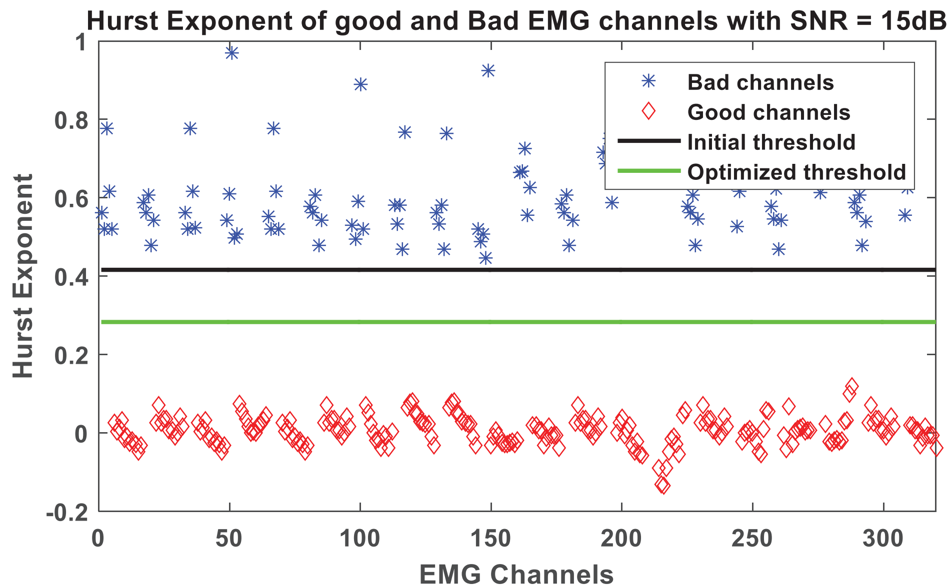

Fig. 6. Hurst Exponent for 320 channels (220 good, 100 bad (contaminated randomly with one of the noises PLI, Electrode movement artifact, ECG artifact and bad contact, WN and CN, with SNR=10dB and 15dB). The initial threshold (-) and the adaptive threshold (-) for the detection of good and bad channels.

consists of simulated EMG signals generated using the multilayer cylindrical description of the volume conductor model, described in [24].

For the first set of 20 simulated EMG signals (320 channels), noise such as power line interference (PLI), electrode movement (EM), white noise (WN), colored noise (CN) and ECG artifacts (ECGA) are added to channel number 2, 5, 8, 13 and 16 of each signal with SNR varying from 15 dB to -5 dB (total of 61 SNR values) respectively. The accuracy, sensitivity and specificity is then computed for each signal of the set of 20 signals across all the SNR values, resulting in 61 values of accuracy, sensitivity and specificity for each signal.

The accuracy, sensitivity and specificity of the proposed algorithm for this set of 20 signals across all SNR values are shown in Figure 8. The results show that the proposed method has a high accuracy in detecting both bad and good channels.

The 2nd set of simulated EMG signals, generated using the cylindrical model, were detected with circular electrodes (diameter 1 mm), arranged in a grid with 5 columns and 40 rows (200 electrodes) with 5 mm inter-electrode distance in both the longitudinal and transverse directions. The center of the grid corresponded to the center of the muscle volume projected on the skin surface. The detection system covered both the muscle (approximately 20 electrodes corresponding to the central portion of each column) and the tendon regions (approximately 10 electrodes over each tendon).

A monopolar recording was simulated for each electrode of the detection system and is then converted to signal differential across the channels. The surface-recorded motor-unit potential was obtained by summing the action potentials of all muscle fibers belonging to individual motor units. EMG signals were simulated at 4096 samples/s. As in this study we investigate only an array of electrodes so only the central column of the detection system is taken. A total of 20 simulated signals with different level of contractions (ranging from 10% to 100% Maximum Voluntary Contractions) were taken.

In the first case all the clean signals were passed through the algorithm for quality checking. The behavior of each of the statistical parameter (quality indicator) are on one side of the threshold value which means that all the channels are identified as good channels. Various artifacts like PLI, movement artifact, real ECG artifact, white noise, colored noise etc. were then added to randomly selected channels with SNR ranging from 10 to -2dB to make some channels bad. For all these signals the accuracy and sensitivity of the proposed algorithm is also computed at various level of noises as shown in Figure 9. It is evident from the results that even at low level of noise i.e. higher SNR, the algorithm still detects the bad channels with a high accuracy, sensitivity and specificity. It is clear from the results that the parameters values are on opposite side of the threshold for good and bad channels in this case too.

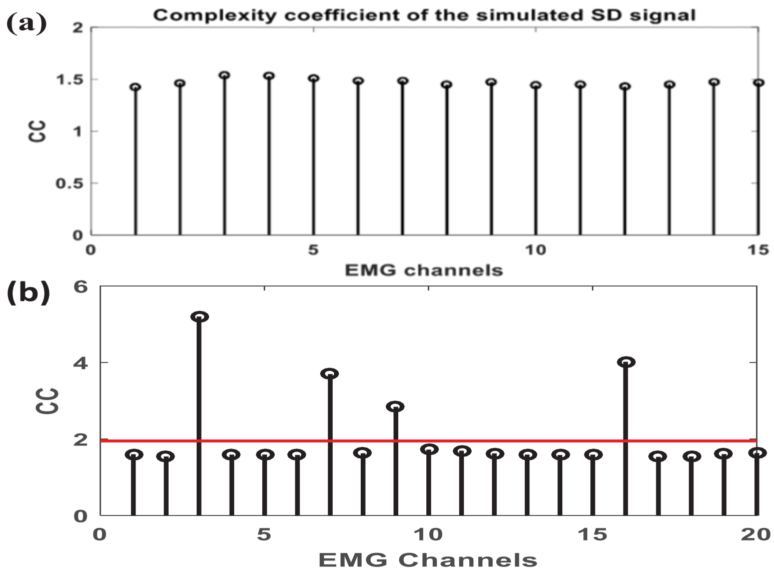

Fig. 7. a) Complexity coefficient of clean (without noise) 16 channel simulated EMG signal, b) Complexity coefficient of simulated EMG with channels 3, 7, 9 and 16 contaminated with different noises.

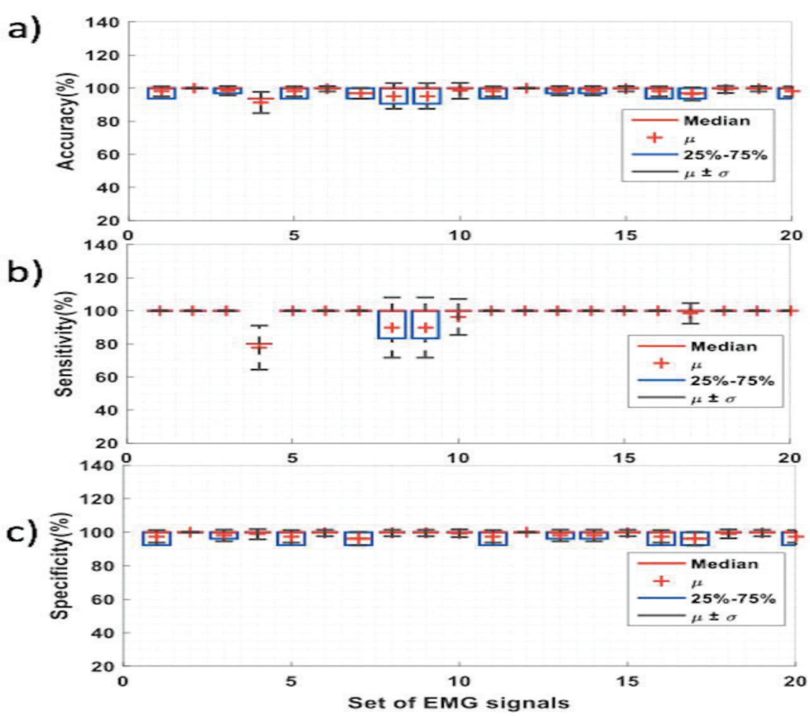

Fig. 8. Accuracy, sensitivity, and specificity of the proposed algorithm for a set of 20 simulated SD EMG signals with SNR varying from 15dB to -5dB. The higher specificity value (98.5492±2.681), shows that the algorithm can preserve most of the good channel as 'good'.

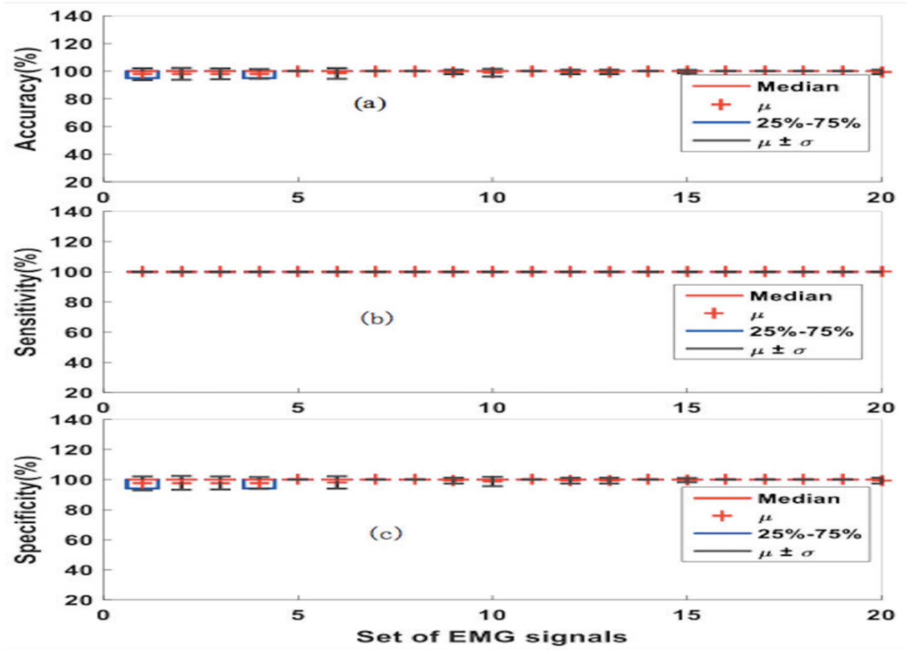

Fig. 9. Accuracy, sensitivity, and specificity of the proposed algorithm for a set of 20 simulated SD EMG signals generated using the cylindrical model [25], each set consisted of 20 channels, with noisy channels of SNR varying from 15dB to -5dB

4. CONCLUSION

This study presents a novel method for the automatic detection of noisy and clean EMG signals which will be helpful to the experimenter while recording EMG signals from a subject. If there are too many noisy signals in an experiment then it is useless for future, so the experimenter will know in realtime to clean the electrodes and make other necessary actions to record clean signals. From the performance of the algorithm on simulated signals, it is concluded that it will be a best choice for automatic detection of bad channels at the time of the acquisition of the signals. As a future study, a classification algorithm can be used after the detection of the noisy channel to also identify the type of the noise and then attenuate it by using a proper filter. As future work this research may be enhanced by recording more experimental signals and use of some machine learning techniques.

5. REFERENCES

1. D. Farina, and R. Merletti. A novel approach for precise simulation of the EMG signal detected by surface electrodes. *IEEE T Bio-Med Eng.*48(6):637–646 (2001).
2. A. J. Fuglevand, D. A. Winter, A. E. Patla. Models of recruitment and rate coding organization in motor-unit pools. *J Neurophysiol*.70(6):2470–2488 (1993).
3. Y. Zhang, B. Liu, X. Ji, D. Huang. Classification of EEG Signals Based on Autoregressive Model and Wavelet Packet Decomposition. *Neural Process Lett.* 1–14 (2016).
4. R. Imoto, M. Migita, M. Toda, S. Sakurazawa, J. Akita, K. Kondo, et al. Preliminaly study on coordinated movement mechanism of multiple muscle using wavelet coherence analysis. *IEEE 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, 605–608 (2016).
5. F. Y. Wu, F. Tong, Z. Yang. EMG signal enhancement based on ICA decomposition and wavelet transform. *Appl Soft Comput.*43:561–571 (2016).
6. A. R. Al-Qawasmi, K. Daqrouq. ECG signal enhancement using wavelet transform. *WSEAS Transactions on Biology and Biomedicine.*7:62–71 (2010).
7. C. K. Chui. *An introduction to wavelets.* Elsevier; 2016.
8. S. Fitzgibbon, D. DeLos Angeles, T. Lewis, D. Powers, T. Grummett , E. Whitham, et al. Automatic determination of EMG-contaminated components and validation of independent component analysis using EEG during pharmacologic paralysis. *Clin Neurophysiol.* 127(3):1781–1793 (2016).
9. D. Dharmaprani, H. K. Nguyen, T. W. Lewis,

D. DeLos Angeles, J. O. Willoughby, K. J. Pope. A comparison of independent component analysis algorithms and measures to discriminate between EEG and artifact components. *IEEE 38^{th} International Conference of the Medicine and Biology Society (EMBC),* 825–828 (2016).

10. T. He, G. Clifford, L. Tarassenko. Application of independent component analysis in removing artefacts from the electrocardiogram. *Neural Comput Appl.* 15(2):105–116 (2006).
11. G. Tsolis, T. D. Xenos. Signal denoising using empirical mode decomposition and higher order statistics. *International Journal of Signal Processing, Image Processing and Pattern Recognition.*4(2):91–106 (2011).
12. V. K. Mishra, V. Bajaj, A. Kumar, G. K. Singh. Analysis of ALS and normal EMG signals based on empirical mode decomposition. *Iet Sci Meas Technol.*10(8):963–971 (2016).
13. A. Siddiqi, S.P. Arjunan, D. K. Kumar. Age related neuromuscular changes in sEMG of m. Tibialis Anterior using higher order statistics (Gaussianity & linearity test). *IEEE 38^{th} Annual International Conference of the Medicine and Biology Society (EMBC),* 3638–3641 (2016).
14. A. Keller, U. Blumenthal, G. Kar. Classification, and computation of dependencies for distributed management. *Fifth IEEE Symposium on Computers and Communications*; 78–83 (2000).
15. D. F. Specht. Probabilistic neural networks. *Neural networks.* 3(1):109–118 (1990).
16. M. Li, Y. Liang, L. Yang, H. Wang, Z. Yang, K. Zhao, Z. Shang, H. Wan, Automatic bad channel detection in implantable brain-computer interfaces using multimodal features based on local field potentials and spike signals, *Computers in Biology and Medicine* (2020), doi: https:// doi.org/10.1016/j.compbiomed.2019.103572
17. R. H. Chowdhury, M. B. Reaz, A. MABM, A. A. Bakar, K. Chellappan, T. G. Chang. Surface electromyography signal processing and classification techniques. *Sensors*. 13(9):12431–12466 (2013).
18. A. S. Khaing, Z. M. Naing. Quantitative Investigation of Digital Filters in Electrocardiogram with Simulated Noises. *IJIEE.*1(3):210 (2011).
19. H. Nolan, R. Whelan, R. Reilly. FASTER: fully automated statistical thresholding for EEG artifact rejection. *J Neurosci Meth.* 192(1):152–162 (2010).
20. E. A. Clancy, N. Hogan. Probability density of the surface electromyogram and its relation to amplitude detectors. *IEEE T Bio-Med Eng*.46(6):730–739 (1999).
21. A. Baˇrbulescu, C. Serban, C. Maftei. Evaluation of Hurst exponent for precipitation time series. *Latest Trends on Computers*. 2:590– 595 (2010).
22. E. Lukhanina, I. Karaban, N. Berezetskaya. *Diagnosis of Parkinson's Disease by Electrophysiological Methods.* INTECH (2011).
23. R. Adler, R. Feldman , M. Taqqu . *A practical guide to heavy tails: statistical techniques and applications.* Springer Science & Business Media, USA (1998).
24. C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, A. L. Goldberger. Mosaic organization of DNA nucleotides. *Phys Rev E*.49(2):1685 (1994).
25. D. Farina, L. Mesin, S. Martina, R. Merletti. A surface EMG generation model with multilayer cylindrical description of the volume conductor. *IEEE T Bio-Med Eng*.51(3):415– 426 (2004).