Graph-based Denoising of EEG Signals in Impulsive Environments

Anastasia Pentari[1,2], George Tzagkarakis[2], Kostas Marias[2,3], and Panagiotis Tsakalides[1,2]
[1]*Department of Computer Science, University of Crete, Heraklion, Greece*
[2]*Institute of Computer Science, Foundation for Research and Technology-Hellas, Heraklion, Greece*
[3]*Department of Electrical and Computer Engineering, Hellenic Mediterranean University, Greece.*
E-mails: apentari@csd.uoc.gr, gtzag@ics.forth.gr, kmarias@ics.forth.gr, tsakalid@ics.forth.gr

***Abstract*—As the fields of brain-computer interaction and digital monitoring of mental health are rapidly evolving, there is an increasing demand to improve the signal processing module of such systems. Specifically, the employment of electroencephalogram (EEG) signals is among the best non-invasive modalities for collecting brain signals. However, in practice, the quality of the recorded EEG signals is often deteriorated by impulsive noise, which hinders the accuracy of any decision-making process. Previous methods for denoising EEG signals primarily rely on second order statistics for the additive noise, which is not a valid assumption when operating in impulsive environments. To alleviate this issue, this work proposes a new method for suppressing the effects of heavy-tailed noise in EEG recordings. To this end, the spatio-temporal interdependence between the electrodes is first modelled by means of graph representations. Then, the family of alpha-stable models is employed to fit the distribution of the noisy graph signals and design an appropriate adjacency matrix. The denoised signals are obtained by solving iteratively a regularized optimization problem based on fractional lower-order moments. Experimental evaluation with real data reveals the improved denoising performance of our algorithm against well-established techniques.**

***Index Terms*—Graph signal denoising, alpha-stable models, fractional lower order moments, impulsive noise, EEG signals**

I. Introduction

Recently, there is an ever increasing demand to further improve the processing capabilities of advanced systems in the fields of brain-computer interaction and digital monitoring of mental health. Electroencephalogram (EEG) signals, that is, the electrical signals recorded directly from the scalp, giving supplementary information about the neural activity of the human brain [1], play a key role in the development of such systems. However, in practice, the quality of the recorded EEG signals is often deteriorated by additive observation noise, which hinders the accuracy of any decision-making process. The majority of previous methods for denoising EEG signals primarily rely on second-order statistics for the corrupting noise, which is not a valid assumption when operating in impulsive environments [2], [3].

This work is funded by the Hellenic Foundation for Research and Innovation (HFRI) and the General Secretariat for Research and Technology (GSRT), under (i) grant agreement No. 2285 (neuronXnet) and (ii) faculty grant no. 1725, and by the Stavros Niarchos Foundation within the framework of the project ARCHERS.

On the other hand, the spatio-temporal interdependence between the electrodes is naturally modelled by means of graph representations. The basic advantage of a *graph signal processing* (GSP) framework is that it exploits both the local and global interrelations between signals captured by the nodes of irregular graph structures. This is especially important for the denoising of graph signals, where the joint consideration of inter-/intra-node correlations often improves the overall noise suppression [4]. As in the case of individual signal denoising, graph signal denoising techniques are typically based on light-tailed, finite-variance, assumptions for the statistics of the noise generating process.

Despite the analytical tractability and practical appeal, these assumptions may yield a dramatic degradation of the denoising performance when we operate in highly impulsive environments, which give rise to heavy-tailed processes with infinite variance. To alleviate the effects of gross errors that mask the information conveyed by the recorded signals, new graph filtering paradigms should emerge.

Furthermore, *alpha-stable* distributions [5] have been proven very powerful in accurately modeling impulsive phenomena. However, their intractability due to the lack of closed-form expressions for the density functions of all except for a few stable distributions (Gaussian and Cauchy) has prevented their exploitation in the framework of GSP. To address this problem, whilst also revealing the advantages of alpha-stable models in designing efficient graph signal filtering algorithms, this paper proposes a graph filtering method for denoising EEG signals under heavy-tailed observation noise, by modeling the noise statistics via *symmetric alpha-stable* ($S\alpha S$) distributions.

The main contribution of this paper is twofold: (i) a statistical weighted adjacency matrix is proposed based on *fractional lower-order moments* (FLOMs), which better adapts to the underlying heavy-tailed noise distribution; (ii) the denoising problem is expressed as a regularized ℓ_p optimization problem, which is related to the FLOMs in a natural way, and a modified iterative reweighted algorithm is designed for its solution.

The rest of the paper is organized as follows: Section II refers to the differences between our proposed method and prior studies. Section III introduces briefly the main concepts of graph signal representations and $S\alpha S$ models, which constitute the main building blocks of our method. Our proposed graph filtering scheme for denoising EEG signals by solving iteratively an ℓ_p-regularized optimization problem is analyzed in Section IV. Section V evaluates the performance of our method on real EEG data and compares its efficiency with well-established denoising methods, namely, wavelet-based denoising and independent component analysis (ICA). Finally, Section VI summarizes the main outcomes of this work and gives directions for further extensions.

II. Relation to Prior Work

Denoising of EEG signals is a well-established topic in the signal processing and bioinformatics communities [1], [6]. Previous methods typically denoise each EEG channel separately by employing transform-based techniques, such as the wavelet transform and its variants [3], [6], [7]. An alternative approach applies an independent component analysis (ICA) [1] towards distinguishing between the signal and noise terms. In all the previous studies, the noise distribution is restricted in the Gaussian case. In order to suppress the noise effects when recording signals in impulsive environments, the methods proposed in [8], [9] exploit the efficiency of $S\alpha S$ models to fit the noise distribution, resulting in improved denoising filters. However, these works do not exploit the potential correlations which often exist in graph-structured data, as is the case with EEG signals. To alleviate this issue, graph filtering techniques have been introduced recently [4], for suppressing the effects of light-tailed (mainly Gaussian) observation noise. Although these methods can treat Gaussian-distributed noise efficiently, however, their performance deteriorates dramatically in the case of impulsive noise, which can not be handled properly. Our proposed method combines the efficiency of graph representations with the power of $S\alpha S$ distributions, in order to design a graph filtering technique for denoising EEG signal ensembles corrupted by heavy-tailed observation noise.

III. GSP and Alpha-Stable Models

This section introduces the main building blocks of our method, namely, the graph representation of EEG signal ensembles, along with the basics of alpha-stable distributions for modelling the statistics of heavy-tailed, possibly of infinite variance, observation noise.

A. Graph representation of EEG signal ensembles

Let $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N] \in \mathbb{R}^{K \times N}$ be the data matrix, where $\mathbf{x}_i \in \mathbb{R}^K$ is the signal recorded by the ith electrode. We adopt an additive observation noise assumption, i.e.,

$$\mathbf{X} = \mathbf{S} + \mathbf{W} \ , \tag{1}$$

where $\mathbf{S} \in \mathbb{R}^{K \times N}$ denotes the noiseless data matrix and $\mathbf{W} \in \mathbb{R}^{K \times N}$ is the noise matrix.

EEG signals admit a natural representation in the form of a graph. Specifically, by considering that each electrode corresponds to a node of the graph, then the EEG signals ensemble can be expressed as a graph $G = (V, E)$ with V, E denoting the set of nodes and edges, respectively. The interdependencies among the distinct nodes (i.e, electrodes) are encoded via an *adjacency matrix* $\mathbf{A} \in \mathbb{R}^{N \times N}$, which is calculated directly from the noisy data $\mathbf{X}$. The naive case is to use a binary matrix $\mathbf{A}$, with $a_{mn} = 1$ if an edge between the nodes m, n exists, otherwise $a_{mn} = 0$ $(m, n = 1, \ldots, N)$.

In order to capture more complex relations between the nodes of a graph, a weighted version of $\mathbf{A}$ is employed, which encodes not only the existence but also the strength of a connection between two nodes. A typical example of a *weighted adjacency matrix* is the correlation matrix, where $a_{mn} = \mathrm{corr}(\mathbf{x}_m, \mathbf{x}_n)$. Notice that for undirected graphs, as is the case considered herein, $\mathbf{A}$ is symmetric.

B. Symmetric alpha-stable models

Our proposed method assumes that the noise matrix $\mathbf{W}$ has independent and identically distributed (i.i.d.) entries drawn from a $S\alpha S$ distribution. In the following, we introduce briefly the family of $S\alpha S$ distributions exploited by our graph filtering method. A $S\alpha S$ distribution is best defined by its characteristic function [5], as follows,

$$\phi(t) = \exp(i\delta t - \gamma^\alpha |t|^\alpha) \ , \tag{2}$$

where α $(0 < \alpha \leq 2)$ is the *characteristic exponent*, which is a shape parameter controlling the "thickness" of the tails of the density function, $\delta \in \mathbb{R}$ is the *location parameter* and $\gamma > 0$ is the *dispersion*, which determines the spread of the distribution around its location parameter, similar to the variance of the Gaussian. The smaller the α, the heavier the tails of a $S\alpha S$ density function. In general, no closed-form expressions exist for $S\alpha S$ distributions except for the Cauchy $(\alpha = 1)$ and the Gaussian $(\alpha = 2)$. Without loss of generality, in the following we model the additive observation noise via $S\alpha S$ distributions with $\delta = 0$.

An important characteristic of $S\alpha S$ distributions (with $\alpha < 2$) is the lack of second-order moments. Instead, all moments of order $p < \alpha$ do exist and are called the *fractional lower-order moments* (FLOMs). In particular, the FLOMs of $X \sim f_\alpha(\gamma, \delta = 0)$ are given by [10]:

$$\mathbb{E}\{|X|^p\} = \big(C(p, \alpha) \cdot \gamma\big)^p \ , \quad 0 < p < \alpha \ , \tag{3}$$

where $\big(C(p, \alpha)\big)^p = \frac{\Gamma\left(1 - \frac{p}{\alpha}\right)}{\cos\left(\frac{\pi}{2} p\right) \Gamma(1-p)}$. The $S\alpha S$ model parameters (α, γ) can be estimated using the consistent maximum likelihood (ML) method described by Nolan [11], which gives reliable estimates and provides the tightest possible confidence intervals.

The concept of covariance is fundamental in the second-order moment theory. However, covariances do not exist for $S\alpha S$ random variables. Instead, a quantity called *covariation*, which plays an analogous role for $S\alpha S$ random variables to the one played by covariance in the Gaussian case, has been proposed. Let X, Y be jointly $S\alpha S$ random variables (i.e., with the same α), zero location parameters and dispersions γ_X and γ_Y, respectively. Then the covariation of X with Y is defined by

$$[X, Y]_\alpha = \frac{\mathbb{E}\{X Y^{<p-1>}\}}{\mathbb{E}\{|Y|^p\}} \gamma_Y^\alpha \ , \tag{4}$$

where $z^{<a>} = |z|^a \text{sign}(z)$ for $z \in \mathbb{R}$ and $a \geq 0$, and $\mathbf{z}^{<a>} = [|z_1|^a \text{sign}(z_1), \ldots, |z_N|^a \text{sign}(z_k)]$ for a vector $\mathbf{z} \in \mathbb{R}^K$. In the discrete case, let two vectors $\mathbf{x}_i$, $\mathbf{x}_j \in \mathbb{R}^K$ be two independent realizations of a $S\alpha S$ distribution. Their FLOM-based covariation estimator is defined by

$$c_{ij}^{\text{FLOM}} = \frac{\sum_{k=1}^K x_{i,k} |x_{j,k}|^{p-1} \text{sign}(x_{j,k})}{\sum_{k=1}^K |x_{j,k}|^p} \gamma_{\mathbf{x}_j}^\alpha \ , \qquad (5)$$

where $x_{i,k} \in \mathbb{R}$ is the kth element of $\mathbf{x}_i$, whilst α and $\gamma_{\mathbf{x}_j}$ are estimated directly from $\mathbf{x}_j$ using Nolan's ML estimator. Notice that, in general, the covariation matrix $\mathbf{C}^{\text{FLOM}}$ of the data matrix $\mathbf{X}$ is neither normalized nor symmetric. In order to convert it into a valid weighted adjacency matrix, first it is normalized as follows,

$$\mathbf{C}_{\text{norm}} = \frac{\mathbf{C}^{\text{FLOM}} - \min(\mathbf{C}^{\text{FLOM}})}{\max(\mathbf{C}^{\text{FLOM}}) - \min(\mathbf{C}^{\text{FLOM}})} \ , \qquad (6)$$

so as to bring its elements within the range $[0, 1]$, similarly to the conventional correlation adjacency matrix, which is typically used for EEG signals. Then, $\mathbf{C}_{\text{norm}}$ is further symmetrized, $\mathbf{C}_{\text{symm}} = (\mathbf{C}_{\text{norm}} + \mathbf{C}_{\text{norm}}^T)/2$, to enable its eigendecomposition. Finally, motivated by [4], our proposed weighted adjacency matrix is obtained by normalizing $\mathbf{C}_{\text{symm}}$ with its maximum eigenvalue,

$$\mathbf{A} = \mathbf{C}_{\text{symm}} / |\lambda_{\max}| \ . \qquad (7)$$

Estimation of p: Notice that the selection of an appropriate value for p is a critical step, which also depends on the characteristic exponent α, which is estimated from the noisy signals in our case. In our implementation, the optimal value of p is calculated as a function of α by minimizing the standard deviation of a FLOM-based covariation estimator, as described in [12], and specifically by linearly interpolating the entries of the lookup Table I in [12].

IV. ℓ_p-Regularized Graph Denoising of EEG Signals

This section analyzes our graph filtering method, which suppresses efficiently the effects of heavy-tailed observation noise, possibly of infinite variance, whilst achieving increased robustness to a broader range of impulsive noise behaviors, from near linear (i.e., $\alpha \to 2$) to extremely impulsive (i.e., $\alpha \to 1$). To this end, the denoising problem is first expressed as a regularized ℓ_p optimization problem, which is then solved by introducing a modified iterative reweighted algorithm.

Under a heavy-tailed assumption for the noise statistics, previously proposed ℓ_2-based formulations (ref. [4]) are not valid. This is because, as mentioned in Section III-B, heavy-tailed random variables of infinite variance do not posses second-order statistics. Motivated by this, we introduce an ℓ_p-regularized optimization scheme for denoising a signal ensemble, which is recorded by the nodes of a graph. In the current implementation, we consider the case of undirected graphs. The use of ℓ_p norms arises naturally in the framework of $S\alpha S$ models, since it is related directly to the FLOMs of $S\alpha S$ random variables (ref. (5)).

Doing so, the ℓ_p-regularized optimization problem for denoising a signal ensemble is expressed as follows,

$$\hat{\mathbf{S}} = \underset{\mathbf{S} \in \mathbb{R}^{K \times N}}{\text{argmin}} \left\{ \frac{1}{2} \|\mathbf{S} - \mathbf{X}\|_p^p + \frac{1}{2} b \|\mathbf{S} - \mathbf{A}\mathbf{S}\|_p^p \right\} \ , \qquad (8)$$

where b is a regularization parameter, which controls the smoothness of the denoised signals. As mentioned in Section III-B, FLOMs exist for $0 < p < \alpha$. On the other hand, the lookup table (ref. [12]) that is employed to choose the optimal p as a function of the estimated α shows that $p < \alpha/2$. This yields a $p < 1$, which makes the problem in (8) highly non-convex, thus necessitating an iterative solution, as described below. Furthermore, the use of an ℓ_p norm both at the data fidelity and regularization terms in (8) is justified by the fact that both the noisy data matrix $\mathbf{X}$ and the adjacency matrix $\mathbf{A}$, which is estimated from the noisy signals, are characterized by heavy-tailed statistics due to the presence of impulsive noise.

Choosing an appropriate adjacency matrix $\mathbf{A}$ is a critical step in modelling accurately the underlying graph structure of the given data. In order to adapt to the impulsive nature of the additive observation noise, our method employs the statistical weighted adjacency matrix defined by (7). Notice that the required parameters for calculating $\mathbf{C}_{\text{symm}}$, and thus $\mathbf{A}$, namely, α, $\gamma_{\mathbf{x}_j}$, and p, are estimated directly from the recorded (noisy) data $\mathbf{X}$, as described in Section III-B.

For the solution of (8), we propose a hybrid optimization scheme, namely, an iteratively reweighted least squares (IRLS)-based solution [13] over the samples of the EEG signals is applied first, followed by a gradient descent-based solution over the distinct channels. Both optimization steps terminate when a maximum number of iterations T has been reached. Regarding the sample-wise operation, our algorithm is initialized by $\mathbf{S}^{(0)} = \mathbf{X} + \epsilon$, where $\epsilon \ll 1$ is a small constant that prevents numerical errors. At each iteration t, the residual, $\mathbf{r} = \mathbf{S}(k,:)^{(t)} - \mathbf{X}(k,:)$, is calculated for the k-th row $(k = 1, \ldots, K)$ of the data matrices. Then, the update step is given by,

$$\mathbf{S}(k,:)^{(t)} = ((\mathbf{D}^T\mathbf{D}) + b(\mathbf{I} - \mathbf{A})^H (\mathbf{I} - \mathbf{A}))^{-1} (\mathbf{D}^T\mathbf{D}) \mathbf{X}(k,:)^T \ , \qquad (9)$$

where $\mathbf{I}$ is the identity matrix and

$$\mathbf{D} = diag(|\mathbf{r} + \epsilon|^{(p-2)}) \ . \qquad (10)$$

Having calculated the IRLS-based solution $\mathbf{S}^{(t)}$, it is used as an initial guess for the subsequent gradient descent scheme which is applied channel-wise. Specifically, the update step following the gradient descent method is given by

$$\mathbf{S}^{(t+1)} = \mathbf{S}^{(t)} + l \nabla Q(\mathbf{S}^{(t)}) \ , \qquad (11)$$

where $Q(\mathbf{S})$ denotes the objective function in (8). After some algebraic manipulation, the gradient in (11) is given by

$$\nabla Q(\mathbf{S}^{(t)}) = p\mathbf{W}_1^{(t)} (\mathbf{S}^{(t)} - \mathbf{X}) + p\mathbf{W}_2^{(t)} (\mathbf{S}^{(t)} - \mathbf{A}\mathbf{S}^{(t)}) \ , \qquad (12)$$

where $\mathbf{W}_1^{(t)}$ and $\mathbf{W}_2^{(t)}$ are diagonal matrices with elements

$$\mathbf{W}_1^{(t)}(n,n) = \sum \left(|\mathbf{I}(n,:) \mathbf{S}^{(t)T} - \mathbf{X}(:,n)^T|^{(p-2)} \right) \ , \qquad (13)$$

$$\mathbf{W}_2^{(t)}(n,n) = \sum \left(|(\mathbf{I}(n,:) - \mathbf{A}(n,:))\mathbf{S}^{(t)T}|^{(p-2)} \right) , \quad (14)$$

The elements of the diagonal matrices $\mathbf{W}_1^{(t)}$ and $\mathbf{W}_2^{(t)}$ correspond to each one of the n electrodes (channels) ($n = 1, \ldots, N$). Notice also that the $(p-2)$-power in (10), (13) and (14), as well as the summation in (13)-(14), are applied in an element-wise fashion.

The updating process is repeated until a maximum number of iterations T is reached. The final denoised EEG signal is given by $\hat{\mathbf{S}} = \mathbf{S}^{(T)}$, where $\mathbf{S}^{(T)}$ is the solution of the ℓ_p optimization problem. In our implementation, the learning rate l is set equal to 10^{-8}, the regularization parameter $b = 1$, and ϵ is at the order of 10^{-2}.

V. Experimental Evaluation

The performance of our graph filtering method is evaluated on a set of four real EEG signal ensembles from a publicly available repository[1]. Each ensemble corresponds to a patient, and consists of a 64-channel EEG signal, with approximately 10000 samples per channel. We note that these signals are considered to be already denoised. The performance of our method (hereafter denoted by ℓ_p) is compared against three well-established alternatives: (1) wavelet-based denoising (WT) applied on each channel separately, using the same parameters setting as in [7], i.e., 3-level decomposition with the "db8" wavelet; (2) independent component analysis (ICA), with a scaling factor equal to 0.64, i.e., proportional to the number of electrodes and (3) ℓ_2-regularized filtering proposed in [4]. The denoising performance of the above methods is evaluated and compared for a range of impulsive noise behaviors, from very impulsive ($\alpha \to 1$) to a Gaussian ($\alpha = 2$). Specifically, we vary $\alpha \in \{1.1, 1.4, 1.7, 2\}$ and $\gamma \in \{0.1, 1\}$. In the following, the values for all the performance metrics we employ are averaged over all the channels and all the patients, as well as over 100 Monte Carlo runs, where in each run a different random noise term is generated.

The performance of the methods compared herein is measured between the original and denoised signals in terms of: (i) the Signal-to-Error Ratio (SER), defined by $\mathrm{SER}(\mathbf{s}, \hat{\mathbf{s}}) = 10\log_{10}\left(\frac{\|\mathbf{s}\|_2^2}{\|\mathbf{s}-\hat{\mathbf{s}}\|_2^2}\right)$ (in dB), where $\mathbf{s}$ and $\hat{\mathbf{s}}$ denote the original and denoised signals, respectively; (ii) the structural similarity index (SSIM) [16], which takes values in $[0, 1]$ (the closer to 1 the higher the similarity); and (iii) the Spearman's correlation [14], [15].

As a first illustration, Fig. 1 shows parts of (a) the second channel of Patient 2 and (b) the tenth channel of Patient 3, along with a noisy version of them and their denoised output using the four methods compared herein. As it can be seen, the WT method yields the smoothest denoised signal, by also suppressing important details of the original (noiseless) signals. On the other hand, in contrast to the ICA and ℓ_2-based filters, which fail in suppressing the spikes due to the impulsive noise, our method (ℓ_p) achieves to adapt to and suppress the underlying heavy-tailed noise.

[1]Dataset available at www.physionet.org/content/eegmmidb/1.0.0/

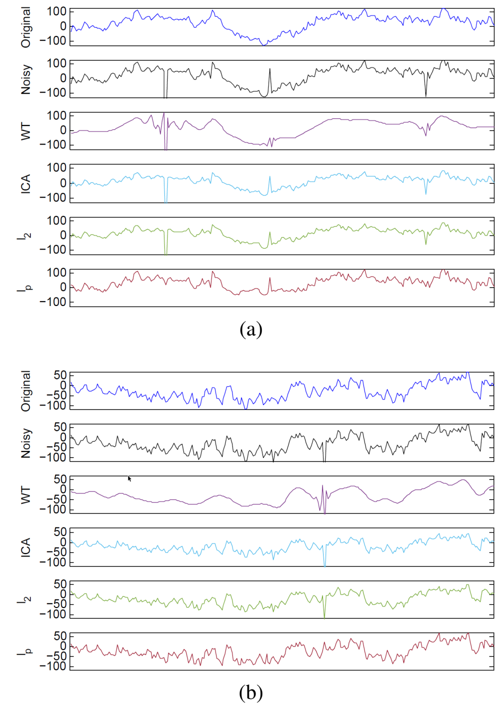

Fig. 1: Denoising performance for (a) 2nd channel, Patient 2; (b) 10th channel, Patient 3. Original, noisy, and denoised signals using WT, ICA, ℓ_2 and ℓ_p ($\alpha = 1.1$, $\gamma = 1$).

Fig. 2 shows the SER values between the original and denoised signals, averaged over all channels, patients, and Monte Carlo runs, as a function of noise characteristic exponent, for $\gamma \in \{0.1, 1\}$. As expected, for highly impulsive noise (i.e., small α) the performance of all methods decreases. This is also the case as the noise dispersion γ increases. On the other hand, for lower-dispersion noise and for noise statistics tending to a Gaussian (i.e., $\alpha \to 2$) the denoising performance of all methods improves. Most importantly, in all cases, our ℓ_p-based graph filter outperforms significantly its counterparts, for the whole range of α, revealing an increased robustness to heavy-tailed noise.

Since the SER can be sensitive to isolated outliers in the denoised signals, the performance of the four methods is also compared in terms of SSIM, which quantifies the structural similarity between two signals. To this end, the left plot in Fig. 3 shows the SSIM between the original and denoised signals, averaged over all channels, patients and Monte Carlo runs, as a function of the noise characteristic exponent. As it can be seen, the WT method yields the lowest denoising performance, due to the smoothing effects of wavelet filtering, which suppress both the noise term and the high-frequency part of the original signals. Regarding the other three methods, the performance of ℓ_2 is inferior against ICA and ℓ_p, since it is not capable in adapting to lower-order moments implied by the impulsive noise. Notably, our ℓ_p method yields the best performance for a broad range of impulsive environments. Furthermore, motivated by [14], [15], the right plot in Fig. 3

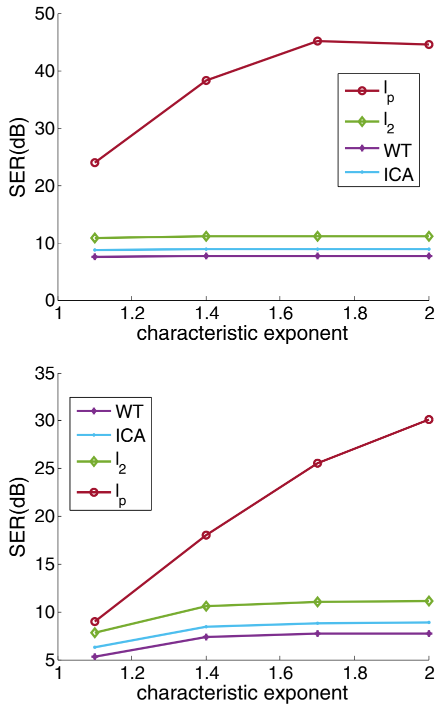

Fig. 2: Average SER, over all channels, patients and Monte Carlo runs, as a function of α for the four denoising methods: (a) Top: $\gamma = 0.1$, (b) Bottom: $\gamma = 1$.

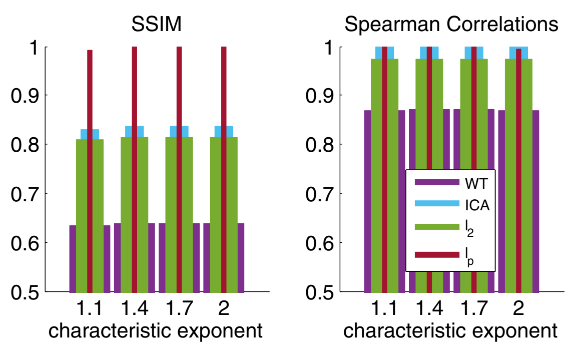

Fig. 3: Average SSIM (left) and Spearman's correlations (right), over all channels, patients and Monte Carlo runs, as a function of α, for the four denoising methods ($\gamma = 0.1$).

shows the average Spearman correlation between the original and denoised signals, over all channels, patients and Monte Carlo runs. The largest the Spearman correlation, the better the denoising performance. Clearly, our method together with the ICA yield the best performance, in terms of this measure, followed by ℓ_2 and WT methods.

Overall, our ℓ_p-regularized graph filter demonstrated the best performance against its alternatives, for a broad range of impulsive environments.

VI. Conclusions and Future Work

In this work, an iterative method was introduced for denoising EEG signals recorded in impulsive environments, by exploiting their graph structure. Specifically, the FLOMs of the signals were employed to design an appropriate weighted adjacency matrix that best adapts to the underlying heavy-tailed statistics of the observation noise. Then, an ℓ_p-regularized optimization problem was solved based on the IRLS algorithm for approximating the true noiseless signals. The experimental evaluation with real EEG signals revealed a significantly improved denoising performance, when compared against the state-of-the-art wavelet-based, ICA and ℓ_2-regularized graph filtering methods, for a broad range of impulsive environments.

As a further extension, we will test the proposed filter on distinct EEG analysis tasks with more channels and alternative adjacency matrices. In addition, the optimization problem in (8) can be further modified by introducing appropriate sparsity constraints on the EEG signals in appropriate transform domains. The joint consideration of a graph representation with the efficiency of sparse representations could further suppress the effects of non-sparse terms due to the additive noise.

References

[1] G. Kaushal, A. Singh and V. K. Jain, "Better approach for denoising EEG signals", *in Proc. Intl' Conf. Wireless Networks & Embedded Syst. (WECON)*, Rajpura, India, 2016.

[2] R. Vigario, *et al.*, "Independent component approach to the analysis of EEG and MEG recordings", *IEEE Trans. Biomed. Eng.*, vol. 47, no. 5, pp. 589–593, May 2000.

[3] A. Turnip and J. Pardede, "Artefacts removal of EEG signals with wavelet denoising", *in Proc. Intl' Conf. Mechan. & Manuf. Eng. (ICME)*, 2017.

[4] S. Chen, *et al.*, "Signal denoising on graphs via graph filtering", *IEEE Global Conf. Signal & Inf. Process.*, Atlanta, GA, pp. 872–876, 2014.

[5] G. Samorodnitsky and M. Taqqu, Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance. New York: Chapman & Hall, 1994.

[6] Harender and R. K. Sharma, "EEG signal denoising based on wavelet transform", *Intl' Conf. Electronics, Commun. and Aerospace Tech. (ICECA)*, Coimbatore, India, pp. 758–761, 2017.

[7] M. Mamun, M. Al-Kadi and M. Marufuzzaman, "Effectiveness of wavelet denoising on electroencephalogram signals," *J. Appl. Res. & Tech.*, vol. 11, no. 1, pp. 156–160, 2013.

[8] N. Bassiou, C. Kotropoulos and E. Koliopoulou, "Symmetric alpha-stable sparse linear regression for musical audio denoising," *in Proc. Intl' Symp. Image & Sig. Process. & Anal. (ISPA)*, Trieste, Italy, 2013.

[9] J. G. Gonzalez and G. R. Arce, "Statistically-efficient filtering in impulsive environments: Weighted myriad filters," *EURASIP J. Appl. Signal Process.*, pp. 4-20, 2002.

[10] G. Tzagkarakis, B. Beferull-Lozano and P. Tsakalides, "Rotation-invariant texture retrieval with Gaussianized steerable pyramids," *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2702–2718, Sept. 2006.

[11] J. Nolan, "Numerical calculation of stable densities and distribution functions," *Commun. Statist.-Stoch. Models*, vol. 13, pp. 759–774, 1997.

[12] G. Tzagkarakis, J. P. Nolan and P. Tsakalides, "Compressive sensing using symmetric alpha-stable distributions for robust sparse signal reconstruction," *IEEE Trans. Signal Process.*, vol. 67, no. 3, pp. 808–820, Feb. 2019.

[13] A. Bjorck, Numerical Methods for Least Squares Problems. SIAM, 1996.

[14] J. Bonita, A. M. Albano, L. Cristobal, C. Ambolode and P. E. Rapp. "Study of correlations of multichannel human EEG," *in Proc. 11th SPVM National Physics Conf. & Work.*, 2009.

[15] M. Lai, *et al.*, "A comparison between scalp- and source-reconstructed EEG networks," *Scient. Rep.*, vol. 8, 12269, Aug. 2018.

[16] Z. Wang, *et al.*, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.