A Study on the Effect of Distinct Adjacency Matrices for Graph Signal Denoising

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 field of brain monitoring is evolving rapidly, there is an increasing demand of finding innovative ways to handle relevant signals. Especially electroencephalogram (EEG) signals provide a non-invasive way of diagnostic inference of brain's functionality. Nevertheless, EEG signals are often corrupted by impulsive noise, thus prior denoising is required for accurate analysis and decision making. On the other hand, EEG signals admit naturally a representation in the form of graphs, with the electrodes corresponding to the nodes of the graph and the edges expressing the connectivity strength. To this end, graph signal processing (GSP) is a versatile tool, which enables the representation and analysis of graph-structured signals, whose interdependencies are encoded in the form of an appropriate adjacency matrix. To address the denoising of graph-structured signals, under impulsive noise conditions, this work introduces a regularized graph filtering scheme based on fractional lower order moments, coupled with distinct adjacency matrices inspired both by statistical approaches and visibility graphs that are better capable of capturing the topological and functional connectivity between the distinct nodes. The experimental evaluation on real EEG signals recorded in epileptic and non-epileptic seizures, reveals the effects of the adjacency matrix choice on the denoising performance.**

***Index Terms*—EEG, graph signal filtering, fractional lower order moments, visibility graph, topological connectivity, functional connectivity**

I. Introduction

In recent years, brain computer interface has attracted a great interest, ranging from medicine to economics. An efficient way of monitoring brain electrical activity is by the EEG signals. The electroencephalogram is a non-invasive technique, which takes records from the human scalp and helps clinicians measure and thus study these electrical activities of the human brain. Clinicians exploit this information for the diagnosis of many neurological diseases, such as the epilepsy. Specifically, epilepsy is a neurological disease characterized by epileptic seizures [1]. It is estimated that 1.5% of the global population suffers from some kind of epilepsy [2]. Still, most of the cases can be controlled by surgery or anti-epileptic medication. An epileptic seizure occurs when the human brain produces four times larger signals than it can normally produce. As a consequence, a group of brain cells results in electrical discharges. EEG recordings give clinicians the opportunity to have a diagnostic insight of the brain and conclude to three facts: (a) if a patient has epilepsy or not, (b) which brain zone is epileptic, and (c) how a specific medication affects (controls) the epilepsy.

Due to head motions in epilepsy and the environmental influences, EEG records often suffer from impulsive noise fluctuations, that is, random spikes with greater amplitude than what is expected. Moreover, the noise which is caused by the eye blink and the muscle activities resembles more a heavy-tailed distribution [3]. Denoising and processing such signals is a highly challenging task, since the noise must be suppressed, whilst maintaining the critical fluctuations of the signals that are due to the epileptic seizures. Motivated by this, in this work, the corrupting noise distribution is modelled by members of the *alpha-stable* family [4].

Recently, *graph signal processing* (GSP) has emerged as a powerful tool to process and analyze graph-structured signals. Regarding EEG signal ensembles, they are naturally represented as graphs, where the nodes correspond to the electrodes and the edges encode the strength of spatial connectivity. A main advantage of GSP is that it allows for the exploitation of both local and global interrelations between the signals. This is important, as the incorporation of both inter- and intra-node correlations typically improves the suppression of noise effects. Specifically, in a recent work [5], the denoised EEG signals' quality was improved significantly, when compared against state-of-the-art method, by employing a graph-based filtering scheme.

Choosing an appropriate adjacency matrix is a critical step towards designing high-performance graph-based denoising methods. Typical choices include weighted adjacency matrices based on the Pearson correlation and the visibility graph theory [6], [7]. In this work, we aim at combining our recent graph-based filter, which better adapts to heavy-tailed noise, with a more efficient weighted adjacency matrix, which is better capable of capturing both topological and functional connectivity among the signals. Designing an appropriate adjacency matrix is especially important for epilepsy signals, due to their unique behavior, which is characterized by multiple fluctuations of the signal's amplitude that must be preserved.

Motivated by this, this work tries to examine how multiple adjacency matrices will react in the problem of graph-signals denoising, using the graph-based filtering method proposed in [5]. Specifically, the examining cases are based on two categories with the one being the statistical-based adjacency

matrices, including the correlation-based and covariation-based ones, while the other concerns the topological/structure exploitation and is based on the state-of-the-art visibility graph theory. However, our proposed adjacency matrix is a combination of both information, as it presents below. Hence, different adjacency matrices are computed, they replace the adjacency matrix used in the filter proposed in [5], and through the performance of the denoised signals the evaluation of such adjacency matrices is held. In summary, the main contribution of this paper is twofold: (i) an innovative adjacency matrix is introduced, encoding both topological and functional connectivity information; (ii) the problem of comparing the adjacency matrices addressed through the use of an existing ℓ_p-regularized graph-based filter, proposed in [5].

The rest of the paper is organized as follows: in Section II, the two building blocks of our method, namely, graph signal basic theory, the symmetric alpha-stable models and the analysis of the graph-based filter used, are briefly overviewed. The jointly adjacency matrix along with its counterparts constructions are analysed in Section III. Section IV evaluates the performance of our method on a real EEG dataset, whilst Section V summarizes the main outcomes and gives directions for further work.

II. GSP and Alpha-Stable Models

This section introduces the main building blocks of our method, namely, the basic graph signal processing concepts and the family of symmetric alpha-stable distributions for modelling the statistics of impulsive 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 (N, K denote the number of electrodes and samples per electrode, respectively), 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.

As mentioned above, EEG signal ensembles can be naturally represented in the form of a graph, $G = (V, E)$ with V, E denoting the set of nodes and edges, respectively. The interrelations among the nodes (i.e., the electrodes) are represented by an adjacency matrix $\mathbf{A} \in \mathbb{R}^{N\times N}$, which, in our analysis, is calculated from the noisy data $\mathbf{X}$.

The conventional way for computing the adjacency matrix is to create a binary matrix $\mathbf{A}$, with $a_{mn} = 1$ if an edge between the nodes m, n exists, otherwise $a_{mn} = 0$ ($m, n = 1, \ldots, N$). On the other hand, it is common to use weights instead of ones, in order to capture more complex relations between the nodes of a graph. As a consequence, a weighted adjacency matrix not only encodes the existence of an edge, but also the strength of the relation.

Concerning EEG signals, there exist many ways of creating an adjacency matrix to describe their interdependencies. The naive way is to calculate the correlation between two nodes, i.e., $a_{mn} = \mathrm{corr}(\mathbf{x}_m, \mathbf{x}_n)$. Notice that, in the case of undirected graphs, the adjacency matrix is symmetric. Although the correlation-based approach is robust enough, it only encodes the functional connectivity of the system. This motivates the design of a more representative adjacency matrix, which could encode the topological information, too, as it will be described in Section III-B.

B. Symmetric alpha-stable models

In the following, we introduce briefly the family of *symmetric alpha-stable* (SαS) distributions. In our study, we assume that the additive noise matrix $\mathbf{W}$ has independent and identically distributed (i.i.d.) entries drawn from a SαS distribution. SαS distributions are better described by their characteristic function defined by [4],

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

where α, the *characteristic exponent*, varies in $(0, 2]$ and controls the "thickness" of the tails of the density function. The smaller the characteristic exponent, the heavier the density function's tails. The *location parameter* $\delta \in \mathbb{R}$ defines the location of the density function. Without loss of generality, in our proposed method, we assume that the noise follows a SαS distribution with $\delta = 0$. The *dispersion* $\gamma > 0$ determines the spread of the distribution around its location, and plays an analogous role to the variance in the case of Gaussian distributions.

Notice that in the case of SαS distributions second order statistics (i.e., covariances) do not exist. Instead, all moments of order $p < \alpha$ do exist, the so-called *fractional lower order moments* (*FLOMs*). In particular, the FLOMs of $X \sim f_\alpha(\gamma, \delta = 0)$ are given by [4]:

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

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 [9], which gives reliable estimates and provides the tightest possible confidence intervals.

Covariances are fundamental in the second-order-moment theory. Nevertheless, for SαS distributions a quantity called *covariation* plays an analogous role. Specifically, let X, Y be jointly SαS random variables (i.e., with the same α), with dispersions γ_X and γ_Y, respectively, and zero location parameters. Then, the covariation of X and Y is defined by

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

where $z^{<a>} = |z|^a \mathrm{sign}(z)$ for $z \in \mathbb{R}$ and $a \geq 0$, and $\mathbf{z}^{<a>} = [|z_1|^a\mathrm{sign}(z_1), \ldots, |z_N|^a\mathrm{sign}(z_k)]$ 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 SαS distribution. Their FLOM-based covariation estimator is defined by

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

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 $\mathbf{x}_j$ based on 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 matrix, analogous to the correlation matrix, first we have to normalize its absolute values as follows,

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

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 be symmetrized by

$$\mathbf{F} = \frac{(\mathbf{C}_{\text{norm}} + \mathbf{C}_{\text{norm}}^T)}{2} \ . \tag{7}$$

Estimation of p: We emphasize that the selection of an appropriate value for p is a critical step. Its computation depends on the characteristic exponent α, which is estimated directly from the noisy signals. 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 [10], and specifically by linearly interpolating the entries of the lookup Table I in [10].

C. ℓ_p-regularized graph-based filter

This section analyzes the graph filtering method, which aims to suppress the effects of heavy-tailed noise $\mathbf{W}$. Hereafter we assume that the noise follows a SαS distribution. The recently introduced ℓ_2 graph-based formulation, proposed in [14], is not valid in the case of SαS noise due to the lack of second-order statistics. To alleviate this issue, the use of ℓ_p norms emerges naturally from the FLOMs of SαS random variables. To this end, the optimization formulation for the graph signal denoising problem is expressed as follows,

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

where b is a regularization parameter which controls the smoothness of the denoised signals. As mentioned before, the FLOMs exist for $0 < p < \alpha$. Moreover, the Table I in [10] shows that $p < \alpha/2$ which is equivalent to $p < 1$, thus the problem in (8) is highly non-convex. This motivates the design of an iterative method for its solution. The choice of ℓ_p norms in both terms of (8) is justified by the facts that, on the one hand, in the data fidelity term the data $\mathbf{X}$ are corrupted by impulsive noise, whilst, on the other hand, the regularization term contains the adjacency matrix, which is calculated directly from the noisy data $\mathbf{X}$. Finally, the use of an appropriate adjacency matrix $\mathbf{A}$ is a critical issue towards achieving a robust denoising performance.

Specifically, for the solution of (8), our method combines a modified iterative reweighted least squares (IRLS) algorithm [15] with an update step which is based on a gradient descent-based formulation, summarized in Algorithm 1.

Let Q_1 denote the objective function to be minimized in (8). For computational convenience, a first approximate solution is obtained in closed form by applying the IRLS algorithm, proposed by [15], on the ℓ_2-based counterpart of our optimization problem, namely,

$$Q_2(\mathbf{S}) = \underset{\mathbf{S} \in \mathbb{R}^{K \times N}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{D}(\mathbf{S} - \mathbf{X})\|_2^2 + \frac{1}{2} b \|\mathbf{S} - \mathbf{A}\mathbf{S}\|_2^2 \right\} \ . \tag{9}$$

The proper initialization of the algorithm is a critical step. In our case, we initialize the solution by adding a small perturbation ϵ at the order of 10^{-2} to the input signal ensemble $\mathbf{X}$. Then, at each iteration, $k = 1, \ldots, K$, the algorithm calculates a residual vector $\mathbf{r}$, as the difference between the current denoised signals $\mathbf{S}^{(t)}$ and the recorded noisy signals $\mathbf{X}$. Then, the residual $\mathbf{r}$ is used to create a diagonal matrix $\mathbf{D} \in \mathbb{R}^{N \times N}$ with the normalized weights $\mathbf{w}_n$ on its main diagonal, which also depends on the noise impulsiveness (via the exponent $(p-2)/2$).

By setting the derivative of the objective function Q_2, with respect to $\mathbf{S}$, equal to zero, an estimate of the denoised signals is given by

$$\begin{aligned} \hat{\mathbf{S}} &= \left(\mathbf{D}^T\mathbf{D} + b(\mathbf{I} - \mathbf{A})^H(\mathbf{I} - \mathbf{A})\right)^{-1} (\mathbf{D}^T\mathbf{D})\mathbf{X} \\ \overset{\mathbf{D^T = D}}{\Longrightarrow} \hat{\mathbf{S}} &= \left(\mathbf{D}^2 + b(\mathbf{I} - \mathbf{A})^H(\mathbf{I} - \mathbf{A})\right)^{-1} \mathbf{D}^2\mathbf{X} \ , \end{aligned} \tag{10}$$

where $\mathbf{I}$ denotes the identity matrix, and H is the Hermitian transpose.

Having obtained a first estimate of the denoised signals, $\hat{\mathbf{S}}$, a further improvement is achieved through the use of a gradient descent scheme applied on the initial objective function Q_1. Specifically, after some algebraic manipulation, the gradient-based updated solution is given by line 12 of Algorithm 1, with

$$\nabla Q_1(\mathbf{S}^{(t)}) = p\mathbf{W}_1^{(t)}(\mathbf{S}^{(t)} - \mathbf{X}) + p\mathbf{W}_2^{(t)}(\mathbf{S}^{(t)} - \mathbf{A}\mathbf{S}^{(t)}) \ , \tag{11}$$

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

$$\mathbf{W}_1^{(t)}(n,n) = \sum_{n=1}^{N} \left(|\mathbf{I}(n,:)\mathbf{S}^{(t)^T} - \mathbf{X}(:,\mathbf{n})^T|^{(p-2)} \right) \ , \tag{12}$$

$$\mathbf{W}_2^{(t)}(n,n) = \sum_{n=1}^{N} \left(|(\mathbf{I}(n,:) - \mathbf{A}(n,:))\mathbf{S}^{(t)^T}|^{(p-2)} \right) \ . \tag{13}$$

The elements of the diagonal matrices $\mathbf{W}_1^{(t)}$ and $\mathbf{W}_2^{(t)}$ correspond to each one of the N electrodes (channels).

Finally, the algorithm terminates when a predetermined number of iterations I_{max} has been reached. For computational stability, in order to avoid zero values in the denominator of the normalized weights $\mathbf{w}_n$, a small constant ϵ is added in the residual $\mathbf{r}$.

III. Adjacency Matrix Design

This section analyzes how the distinct adjacency matrices are designed. Specifically, the adjacency matrices considered herein are inspired by statistical and topological information extraction. In the former case, statistical approaches transfer

Algorithm 1 ℓ_p-regularized graph filter

Inputs: $\mathbf{X}$, $\mathbf{A}$, I_{max}, b, ϵ
Outputs: $\hat{\mathbf{S}}$
Initialization: $\mathbf{S}^{(0)} = \mathbf{X} + \epsilon$

for t = 1:I_{max} **do**
  for k = 1:K **do**
    $\mathbf{r} = \mathbf{S}(k,:)^{(t)} - \mathbf{X}(k,:)$
    $\mathbf{w} = |\mathbf{r} + \epsilon|^{(p-2)/2}$
    $\mathbf{w}_n = \frac{\mathbf{w}}{\max(\mathbf{w})}$
    $\mathbf{D} = \mathrm{diag}(\mathbf{w}_n)$
    $\mathbf{S}(k,:)^{(t)} = \left((\mathbf{D}^2 + b(\mathbf{I}-\mathbf{A})^H(\mathbf{I}-\mathbf{A}))^{-1}\mathbf{D}^2\right)\mathbf{X}(k,:)^T$
  end for
  $\mathbf{S}^{(t+1)} = \mathbf{S}^{(t)} + q\nabla Q_1(\mathbf{S}^{(t)})$
end for
$\hat{\mathbf{S}} = \mathbf{S}^{(Imax)}$

the functional brain information, while in the latter case, the structural brain information is described by the topological adjacency matrices.

A. Statistical adjacency matrices

The functional connectivity provides information about the interrelations between distinct brain areas. In the literature, the most commonly used statistical adjacency matrix is based on Pearson's correlations (hereafter denoted by "Corr"). In this work, we investigate an alternative way for calculating the interrelations between the nodes (electrodes). Specifically, a FLOM-based covariation matrix (hereafter denoted by "Cov") is employed as a more efficient adjacency matrix that better adapts to the underlying SαS distributed heavy-tailed noise.

As mentioned in Section II-B, the covariation matrix is appropriately normalized to become a valid analogue of the conventional correlation matrix for further use in the graph filter.

B. Topological adjacency matrices

The intuition behind the use of the visibility graph theory stems from the methodology proposed by [6]. Therein, the automatic segmentation of the epileptic part of EEG signals is obtained via visibility graphs. In this work, we extend this analysis towards using visibility graphs to better represent the topological information of the brain.

Specifically, the procedure followed for the construction of the associated adjacency matrix consists of two phases, namely, a) at the samples level, and b) at the nodes level. At the first phase, consider that for each pair of nodes, (m, n), with $m, n = 1, \ldots, N$, their corresponding signals $\mathbf{x}_m$, $\mathbf{x}_n$ are processed as follows:

- m-th node:
  1) Sort the elements of $\mathbf{x}_n$ in ascending amplitude.
  2) Take the corresponding ordering for the $\mathbf{x}_m$.
  3) Generate a visibility graph $\mathbf{A}_m$.
- n-th node:
  1) Sort the elements of $\mathbf{x}_m$ in ascending amplitude.
  2) Take the corresponding ordering for the $\mathbf{x}_n$.
  3) Generate a visibility graph $\mathbf{A}_n$.

As a simple illustration, let the samples of the m-th and n-th nodes be of amplitudes $[2, 4, 8, 3]$ and $[3, 5, 1, 8]$, respectively. First, we sort the n-th node's amplitudes in ascending order, i.e., $[1, 3, 5, 8]$ and we take the corresponding positions $\{3, 1, 2, 4\}$ as the output (step 1). These positions are finally used in the m-th node and its corresponding signal is converted to $[8, 2, 4, 3]$ (step 2).

Regarding step 3, it concerns the main procedure for the construction of the visibility graph, as shown in Fig. 1. In particular, now we compare the amplitudes of the new sorted $\mathbf{x}_m$ signal. This procedure is based on the following assumption: for each element k out of the K signal's length, we assume the range $[k-s, k+s]$, where s is a shift parameter defining how many neighbors we will take into account from the left and the right side of the examined element k. In this range of selected samples, if the examined k element is greater than each one of the neighboring samples, i.e., $k-s, \ldots, k-1, k+1, \ldots, k+s$, then an "1" fills the adjacency matrix corresponding position. However, if the k-th element is less than an element which belongs in the range $[k-s, k+s]$, then the adjacency matrix corresponding index is equal to "0".

The conventional way to construct the visibility graph-based adjacency matrix is by putting "1" if two objects (samples) can see each other (i.e., one has greater amplitude than the other), otherwise we put "0". However, motivated by [6], a weight can be assigned to each edge, making the visibility graph-based adjacency matrix more robust. To calculate the strength of these connections, let's assume, for instance, that we are located at the k-th position out of the K ones in the sorted $\mathbf{x}_m$ signal. Then, we have to examine s samples on the right and s on the left side of the k-th sample (accounting properly for the bounding conditions whenever s samples do not exist either on the left or the right side). As a consequence, the corresponding weights for the range $l \in [k-s, k+s]$ are given by

$$w(k,l) = |\mathbf{x}_m(l) - \mathbf{x}_m(k)| \ , \ l > k \ , \tag{14}$$

for each one of the s right neighbors, and

$$w(k,l) = |\mathbf{x}_m(k) - \mathbf{x}_m(l)| \ , \ k > l \ , \tag{15}$$

for each one of the s left neighbors, and for $k = 1, \ldots, K$. These calculations are performed for each node, resulting in the final $\mathbf{A}_m \in \mathbb{R}^{K\times K}$ and $\mathbf{A}_n \in \mathbb{R}^{K\times K}$ adjacency matrices, respectively. For each (m, n) pair, the above two matrices are combined to obtain the (m, n)-th element of the joint adjacency matrix, as follows,

$$\mathbf{A}_{VG}(m,n) = \sum_{k=1}^{K}\sum_{l=1}^{K}(\mathbf{A}_m(k,l) + \mathbf{A}_n(k,l)) \ . \tag{16}$$

Having taken into account the interrelations among the samples of the EEG signals, and given the degrees of connectivity of each examined (m, n) combination, we move to

the second phase, i.e., the node-level analysis. Specifically, suppose that each row n out of the N of the $\mathbf{A}_{VG}$ matrix corresponds to a "signal" of length N. Doing so, a $1 \times N$ vector is exported as presented in Fig. 2. Since the chosen row consists of N elements, each one of them is examined through the VG procedure, i.e., whether it has a larger or smaller degree of connectivity (value) compared with its s right and s left neighbors. Notice that, in this iterative process, overlapping/re-examined elements may occur. In such a case, their extracted weights for the corresponding element's position (the previous, the current and the next ones) will be added. Finally, each exported vector corresponds to the n-th row of the new $\mathbf{H} \in \mathbb{R}^{N \times N}$ matrix, which encodes the topological information of the brain. Notice that $\mathbf{H}$ must be rescaled in the range $[0, 1]$, and then be symmetrized, in accordance to (6) and (7).

The visibility graph theory enables the computation of interrelations between graph nodes in a more accurate way, since it does not account only for the current nodes, but also for their neighbors. Through the resulting visibility graphs we aim to compute a weighted adjacency matrix which represents the interrelations between signal samples, as well as the inter-dependencies between the nodes of the graph. This remark motivates the construction of our proposed visibility graph-based adjacency matrix, hereafter denoted by "VG".

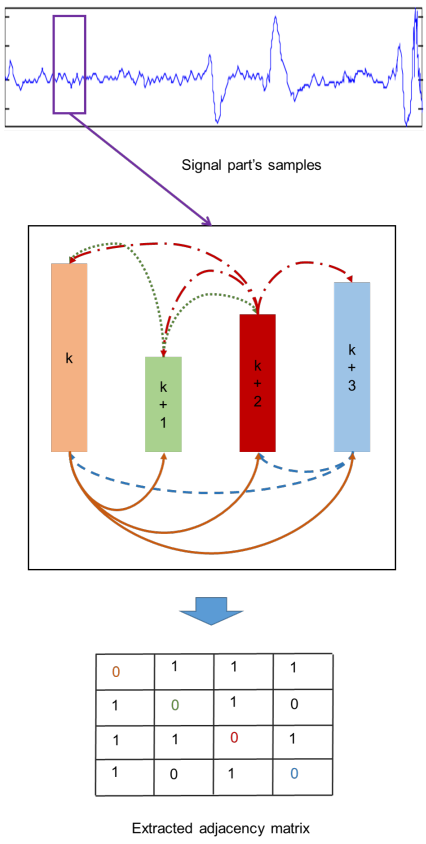

Fig. 1. VG procedure (at the signal level).

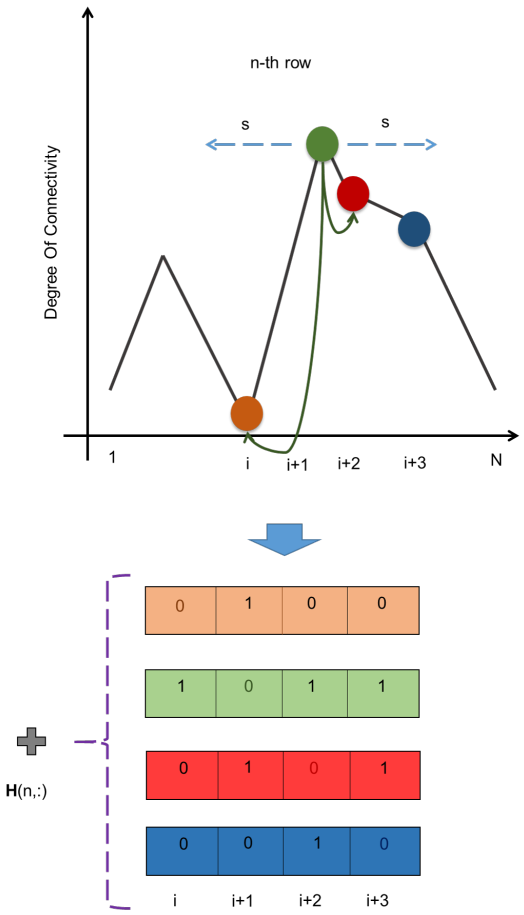

Fig. 2. VG procedure (at the node level).

C. Combined adjacency matrix

In this section, we propose a new adjacency matrix which exploits both the functional and topological brain connectivity. Analogously to the formula presented in [16], herein we compute the topological connectivity through the visibility graph, whereas the functional connectivity is expressed by the covariation matrix. As a result, our proposed joint adjacency matrix, hereafter denoted by "Comb", is given by

$$\mathbf{A} = \mathbf{F}^{(\mathbf{1}+\mathbf{H})} \ . \tag{17}$$

Notice that, in the above equation, the power corresponds to an element-wise calculation.

Before its incorporation in (8), the adjacency matrix $\mathbf{A}$ is normalized by calculating the eigen-decomposition of each examined adjacency matrix, i.e., each one of the {Corr, Cov, VG, Comb} matrices compared in this work, and dividing with the maximum absolute eigenvalue, as follows,

$$\mathbf{A} = \frac{\mathbf{A}_{examined}}{\max(|\lambda|)} \ . \tag{18}$$

IV. Experimental Evaluation

In this section, the denoising performance of the four adjacency matrices presented above is evaluated on real data.

The publicly available dataset [17], [18] we utilize consists of 23-channel EEG signals with 921600 samples per channel, and for four distinct patients. Moreover, the denoising performance is evaluated for distinct characteristic exponent values, α, ranging from very impulsive (e.g. $\alpha = 1.1$) to light-tailed noise (e.g. $\alpha = 1.9$), while the dispersion is set to $\gamma = 1$. For each patient, 4 parts are randomly selected, i.e., 16 signals of 256 samples are examined, including both epileptic and non-epileptic behaviours. The learning rate q is set empirically equal to 10^{-8}, the tuning parameter b, which controls the smoothness of the denoised signals is fixed to 0.001, the constant ϵ is set to 0.1, and the number of iterations I_{max} varies in $[12 : -2 : 4]$, depending on the α value. In the first phase (signal level), the shift value is set equal to 5, whilst in the second phase (nodes level), the shift is fixed at 3. Finally, the experimental results are averaged over 50 Monte-Carlo runs, as well as over all channels, examined parts and patients.

Regarding the filtering process, three performance metrics are employed, namely, (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 the denoised signals, respectively, (ii) the Spearman correlation, motivated by [19], [20], and (iii) the structural similarity index (SSIM) [21].

As a first illustration, in Fig. 3 the 16 parts, 4 per patient, are depicted for the purposes of our analysis. Specifically, this figure concerns the 3rd, 6th, 9th, 12th channels of each examined patient, respectively, and the noise case is for $\alpha = 1.1$, i.e., the most impulsive. As it can be seen, the ℓ_p-regularized filter coupled with the proposed combined adjacency matrix gives well-denoised signals for all the patients and channels, while the alternative adjacency matrices tend to retain more spikes due to the impulsive noise. Notice that the choice of 4 patients is motivated by the fact that, as each examined person has a different brain geometric anatomy, we are interested in examining how the different adjacency matrices adapt.

Fig. 4 depicts the average SER values for the four adjacency matrices. Clearly, our proposed matrix achieves a higher performance compared to its counterparts, whilst the correlation-based adjacency matrix is the one closer to ours, which is obvious in the most impulsive noise case.

On the other hand, the SSIM is a performance metric that better resembles human visual perception. The average SSIM values shown in Fig. 5 reveal that the proposed combined matrix results in a higher SSIM compared to its counterparts (Corr, Cov and VG) and this performance difference is more intense in the high-level noise case, i.e., for $\alpha = 1.1$. However, for less impulsive noise, i.e., for $\alpha = 1.3$, it is comparable when compared with the well-established Corr matrix. Most importantly, as the noise becomes less impulsive (close to the Gaussian) the SSIM of all the examined cases increases. Finally, in Fig. 6, the average Spearman correlations are presented, proving that our proposed adjacency matrix results in not only well-denoised signals, but also in high correlated ones with the original recordings.

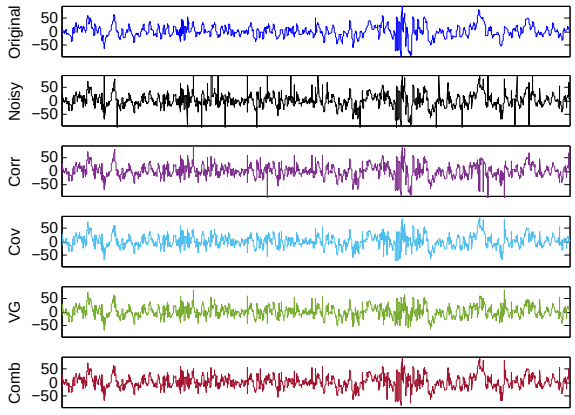

Fig. 3. Sample signals of the random channels from all the patients: original, noisy and their denoising versions from Corr, Cov, VG and comb ($\alpha = 1.1, \gamma = 1$).

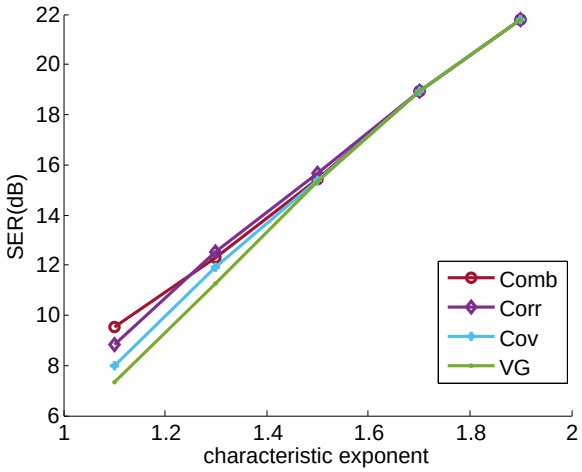

Fig. 4. SER between the original and the denoised data.

V. Conclusions and Future Work

In this work, we couple a recently introduced graph-based denoising filter with distinct adjacency matrices, which capture either functional or topological, or both, information in the brain. Specifically, our proposed combined adjacency matrix, which combines topological and functional information, is compared with well-established adjacency matrices, namely, the correlation, the covariation, and the visibility graphs. Comparison of several performance metrics reveals that our proposed adjacency matrix results in a more accurate reconstruction of the original (noiseless) recordings for a broad range of noise behaviours, as well as in higher Spearman correlation and SSIM values.

As a future extension of this work, we are interested in investigating the performance of the proposed framework towards addressing more challenging problems, such as the automatic detection of seizure events, by jointly accounting

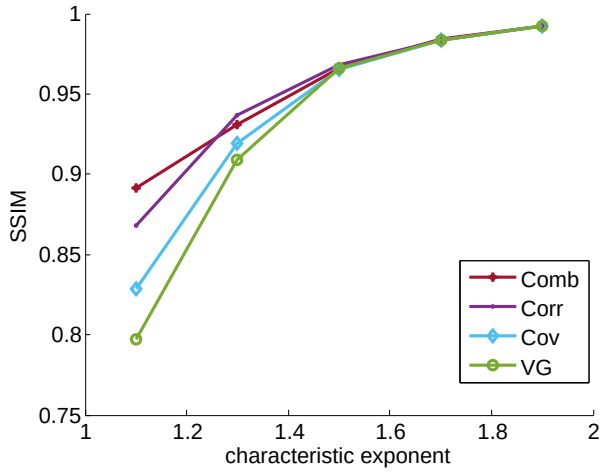

Fig. 5. SSIM between the original and the denoised data.

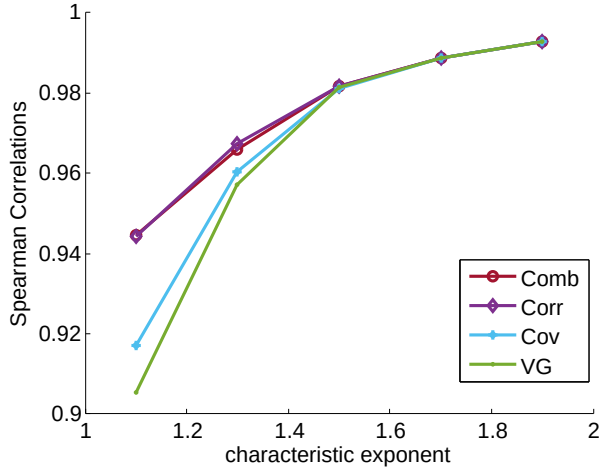

Fig. 6. Spearman correlation between the original and the denoised data.

for functional and topological connectivity patterns between the EEG signals.

ACKNOWLEDGMENT

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

REFERENCES

[1] Z. Lasefr S. Shiva, V. Ayyalasomayajula, K. Elleithy, "An Efficient Automated Technique for Epilepsy Seizure Detection Using EEG signals", 2017.

[2] A. Garcés Correa, L.L. Orosco, P. Diez, "Adaptive Filtering for Epileptic Event Detection in the EEG", *J. Med. Biol. Eng. 39*, 2019.

[3] Y. Wang,Y. Qi, Y. Wang, L. Zhen, X. Zheng, G. Pan, "Delving into - stable distribution in noise suppression for seizure detection from scalp EEG", *Journal of neural engineering*, 2016.

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

[5] A. Pentari, G. Tzagkarakis, K. Marias, P. Tsakalides, "Graph-based denoising of EEG signals in impulsive environments", *EUSIPCO*, 2020.

[6] S. Supriya, S. Siuly, H. Wang, G. Zhuo Y. Zhang, "Analyzing EEG signal data for detection of epileptic seizure: Introducing Weight on Visibility Graph with Complex Network Feature", *Springer International Publishing*, 2016.

[7] S. Sannino, S. Stramaglia, L. Lacasa, D. Marinazzo, "Visibility Graphs for fMRI Data: Multiplex Temporal Graphs and Their Modulations Across Resting-State Networks", *Network Neuroscience*, 2017.

[8] 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.

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

[10] 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.

[11] L. Zie, Q. Li, "Planar Visibility Graph Network Algorithm For Two Dimensional Timeseries", *29th Chinese Control And Decision Conference (CCDC)*, pp. 1352-1357, 2017.

[12] N. Ahmadi, R. M. H. Besseling, M. Pechenizkiy, "Assessment of visibility graph similarity as a synchronization measure for chaotic, noisy and stochastic time series", *Social Network Analysis and Mining*, 8(1),[47], 2018.

[13] P. Djuric, C. Richard, "Cooperative and Graph Signal Processing", *Elsevier*, 2018.

[14] S. Chen, A. Sandryhaila, J. M. F. Moura, J. Kovacevic, "Signal denoising on graphs via graph filtering", *in Proc. IEEE GlobalSIP*, 2014.

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

[16] J. Budd, H. Cuntz, S. Eglen, P. Krieger, "Editorial: Quantitative Analysis of Neuroanatomy", *Frontiers in neuroanatomy*, vol. 9 143. 2015.

[17] A. Goldberger, L. Amaral, L. Glass, JM. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C. K. Peng, H. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals", *Circulation*, 2000.

[18] A. Shoeb, J. Guttag, "Application of Machine Learning To Epileptic Seizure Detection", *27th International Conference on Machine Learning (ICML)*, pp. 975-982, 2010.

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

[20] M. Lai, M. Demuru, A. Hillebrand and M. Fraschini, "A comparison between scalp- and source-reconstructed EEG networks," *Scientific Reports*, Aug. 2018.

[21] Z. Wang, AC. Bovik, HR. Sheikh, EP. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600-612, 2004.