Sparse Reconstruction for Enhancement of the Empirical Mode Decomposition-Based Signal Denoising

KRZYSZTOF BRZOSTOWSKI[iD]

Faculty of Computer Science and Management, Wrocław University of Science and Technology, 50-370 Wrocław, Poland

e-mail: krzysztof.brzostowski@pwr.edu.pl

ABSTRACT Effective signal denoising methods are essential for science and engineering. In general, denoising algorithms may be either linear or non-linear. Most of the linear ones are unable to remove the noise from the real-world measurements. More suitable methods are usually based on non-linear approaches. One of the possible algorithms to signal denoising is based on empirical mode decomposition. The typical approach to the empirical mode decomposition-based signal denoising is the partial reconstruction. More recently, a new concept inspired by the wavelet thresholding principle was proposed. The method is named the interval thresholding. In this article, we further extend the concept by the application of the sparse reconstruction to the empirical mode decomposition-based signal denoising algorithm. To this end, we state and then solve the problem of signal denoising as a regularization problem. In the article, we consider three cases, that is, three types of penalty functions. The first algorithm is combining total variation denoising with empirical mode decomposition approach. In the second one, we applied the fused LASSO Signal Approximator to design the empirical mode decomposition–based signal denoising algorithm. The third approach solves the denoising problem by applying a non-convex sparse regularization. The proposed algorithms were validated on synthetic and real-world signals. We found that the proposed methods have the ability to improve the accuracy of the signal denoising in comparison to the reference methods. Significant improvements from both the synthetic and the real-world signals were obtained for the algorithm based on non-convex sparse regularization. The presented results show that the proposed approach to signal denoising based on empirical mode decomposition algorithm and sparse regularization gives a great improvement of accuracy, and it is the promising direction of future research.

INDEX TERMS Non-linear signal processing, non-convex optimization, gyroscopes.

I. INTRODUCTION

The signal denoising is one of the fundamental and the most challenging tasks in science and engineering. Because of the imperfections present in the measurement system, the obtained signals can be deteriorated during the processes of acquisition and transmission. Thus, the main challenge of signal denoising is to preserve and enhance the desirable features of the collected signals.

In general, filters may be either linear or nonlinear. Removing noises from the acquired signals by applying linear filters usually leads to unsatisfactory results. This is due to the fact that these filters are suitable when the unknown signal is restricted to a known frequency band. Therefore, when the real-world signals are considered, nonlinear filters have to be taken into account. The reason is that these measurements are more complex and not restricted to a specific frequency band. A further complication is a non-stationarity of the acquired signals. Non-stationarity of signals means that signals have time-varying frequency contents [1]. Hence, most of the denoising algorithms, used in practice, are nonlinear.

Linear filters are out of the scope of this work. However, we refer interested readers to [2].

Nonlinear filters are a viable alternative to linear filters. These methods offer an advantage in applications in which the underlying process is non-stationary and nonlinear. As we mentioned, non-stationarity can produce signals with time-varying frequency contents. Another disadvantage of the

The associate editor coordinating the review of this manuscript and approving it for publication was Naveed Ur Rehman[iD].

discussed process is high-amplitude noise which is challenging problems for linear filters [3].

There are many papers proving that significant improvements in performance can be achieved by applying nonlinear signal processing methods [4]–[6], [76]. One of the examples is wavelet transform-based approach [4], [5]. The typical technique widely used in the approach is thresholding [6], [76]. In these papers, the authors introduce two basic approaches for wavelet transform-based denoising, that is, hard and soft thresholding. One of the major problems is that methods based on this technique might remove wavelet coefficients containing useful information. Moreover, the challenging task in hard and soft thresholding is the choice of the suitable threshold value.

Non-local means (NLM) is another signal denoising technique. Originally, this approach was applied to attenuate noise from images [7]. Recently [8], a non-local means approach for ECG signals denoising was proposed. One of the distinctive features of the method is that it has a small effect on the high amplitude of the signals. Hence, denoising is focused on low amplitude regions in the processed signals [9], [10].

In work [11], an approach combining the advantage of non-local means and wavelet transform is proposed. The proposed method differs from the existing algorithm. In this case, the block of samples is estimated in a collaborative manner. Then, these estimates are averaged to find their final values. The authors of the paper suggest that a sparsely transformed domain will improve the overall performance of the denoising algorithm. In the paper, it was also highlighted that the methods combining non-local means with wavelet transform-based approaches are able to preserve crucial components of the signals much better than the existing algorithms [11].

Empirical mode decomposition (EMD) was firstly proposed in [12]. The algorithm decomposes the input signal into a set of modes called Intrinsic Mode Functions (IMFs) and a residual signal. Advantage of empirical mode decomposition is that this method that does not rely on the predefined basis function nor other parameters. However, empirical mode decomposition has some disadvantages as well. One of their drawbacks is that the method is prone to mode mixing. Another is that the approach still has no sound mathematical theory and it is described only by algorithm.

The EMD method has been successfully applied to the signal denoising in various research fields, i.e., biomedical signals [13]–[15], [72] acoustic signals [16]–[18] and signals acquired from inertial sensors [19]–[21], [109]. For example, in [15] the authors verified the proposed EMD-based denoising methods on electrocardiography data, whereas in [20] and [21] the noise attenuating method was tested on gyroscopic data.

The effectiveness of EMD methods was also combined with wavelet transform and presented in [22], [23]. Recently, empirical mode decomposition was incorporated with non-local means as well [10].

Despite the fact that the EMD method was successfully applied to various real-world problems, it is still an approach that suffers from some drawbacks such as sensitivity to noise or sampling [24].

In [24] an alternative to the EMD approach was proposed. The method is entirely non-recursive and the modes are extracted concurrently. The method is named variational mode decomposition and, contrary to EMD, has strong mathematical foundations and a better noise robustness. Moreover, the computational effectiveness of the algorithm is higher than the empirical mode decomposition method. Variational mode decomposition algorithm was used to denoise, for example, biomedical images [25], [26], biomedical signals [27], and inertial data [28].

In the paper, we proposed an approach to design EMD-based denoising algorithms exploiting the idea of sparse reconstruction. In general, the proposed approach is an extension of the thresholding technique originally proposed in [29]. The algorithm to signal denoising suggested in that work is based on traditional ℓ_0 and ℓ_1 norms which are used as functions to promote sparsity. In this paper, we consider other functions that can be used to promote sparsity more strongly. It leads to new formulations of optimization tasks which solutions are applied to the empirical mode decomposition–based signal denoising methods.

A. CONTRIBUTION

The main contributions of this paper are new algorithms for signal denoising based on empirical mode decomposition and sparse reconstruction:

a) EEMD-TVD; a denoising algorithm applies the total variation denoising (TVD) method to modify IMFs extracted via EEMD;

b) EEMD-FLS; a denoising method based on the fused LASSO signal approximator (FLS) to modify EEMD-related modes;

c) EEMD-NSR; a non-convex sparse regularization-based approach to modify IMFs.

We also compared EEMD-TVD, EEMD-FLS and EEMD-NSR with the EEMD-based denoising algorithm (the method based on the procedure proposed in [29]), the non-local means method [8] and wavelet-TV denoising (WATV) [34]. Verification of the proposed methods was performed based on synthetic signals with the fractional Gaussian noise and different noise level as well as on the real-world signals acquired from the three-axis gyroscope.

The paper is organized as follows. Section II presents the state-of-the-art methods for signal denoising based on sparse reconstruction. In section III we review basic methods applied in the proposed approach. The main results are presented in section IV, i.e., we formulate the denoising problem, then we propose the model and the solution to the problem. In section V comprehensive numerical studies for the synthetic and the real-world signals are performed. We finally conclude this paper in sections VI and VII.

II. RELATED WORKS

Sparsity-based signal processing is one of the new branches relevant to the denoising algorithms [30]. These algorithms are most suitable when the signal of interest is itself sparse or when it permits a sparse representation. The idea of sparsity has been successfully applied in signal processing [30]–[33], [34], [41], [42].

Total variation (TV) is one of the well-known sparsity-based approaches to signal denoising [33]. Unfortunately, in the process of TV-based denoising, stair-case artifacts are often generated. Moreover, the high-amplitude components of the signal are underestimated. To improve the traditional TV algorithm some variants of the original method are proposed. For example, in [33] the authors suggest total variation denoising approach based on majorization-minimization and altering direction method of multipliers (ADMM), whereas in [43], the method of suppression of the low- and high-frequency components was proposed. It is worth highlighting that the authors of [43] introduced a method based on an advanced penalty function.

Combining the sparse regularization with TV denoising we can deduce the problem of the fused Least Absolute Shrinkage and Selection Operator (fused LASSO) [44]. In general, the fused LASSO is closely related to the standard LASSO. This method utilizes the sparsity of piecewise constant signals in the discrete derivative domain [45].

The fused LASSO formulation may be further extended to a non-convex penalty function such as the regularization [46]. Most of the algorithms for solving the fused LASSO problem are based on some iterative algorithms such as Majorization-Minimization or ADMM.

The idea of sparsity can be applied in the wavelet domain as well. One of the successful examples of the sparsity-based regularization is presented in [34]. The approach proposed in that work is based on the ADMM framework. In the method, the sparse regularization in the wavelet domain is expressed by a non-convex penalty function since it induces sparsity more strongly. The proposed method is relatively resistant to pseudo-Gibbs oscillations and spurious noisy spikes [34]

Another example of the application of the sparsity concept in the wavelet domain is discussed in [31]. In work, to induce the wavelet-domain sparsity, the authors employ non-convex penalty functions.

In [31] the wavelet regularization is combined with the total variation for solving the problem of denoising. To this end, the authors reformulate the unconstrained optimization task to a constrained optimization problem. The results reported in the paper show that the proposed solution outperforms the reference methods presented in the paper.

Another line of works concentrates on the decomposition of the signal by determining the sparsest representation of a signal over a dictionary consisting of all IMFs generated by the empirical mode decomposition method. For example, the main idea of [47] is to determine the sparsest representation of the measured signal within the largest possible dictionary of components. In work, the dictionary can be seen as a

collection of all the possible IMFs. Thus, the aim is to find the best modes based on the elements of the dictionary. It leads to the problem of dictionary learning. A subsequent work of the authors [47] is the paper [48], in which they have proposed an algorithm to find sparse representations of the signal over a parametrized dictionary. In work, the components of the dictionary are produced via the EMD method.

It is worth mentioning that the dictionary learning can be combined with other methods to compute dictionaries. For example, K-singular value decomposition is one of the most representative among others [49]–[51]. Apart from K-singular value decomposition and EMD algorithms to determine dictionary, it is possible to use, for example, wavelet transform, short-time Fourier transform, and Gabor transform [52].

A different example of the combination empirical mode decomposition and sparsity is considered in [53]. The algorithm generates a robust and effective representation of the data at a relatively low sparsifying level. The presented in the paper method is effective and computationally efficient. Results of experiments have confirmed the competitive performance of the proposed algorithm with other representative methods. In the paper, the authors have shown the promising results in the denoising problem.

In the paper [54], sparse optimization is used for extracting the underlying trends of signals via a combination of empirical mode decomposition and sparse optimization. In turn, work [65] suggests an empirical approach for dictionary learning where the dictionary is learned from the acquired signal. The proposed method is applied to signal classification.

Empirical mode decomposition is also combined with other methods to solve the denoising problem. For example, work [55] considers the problem of lidar signal denoising based on EMD algorithm supported by soft thresholding and roughness penalty. The proposed approach efficiently improved the lidar signal and extend its detection range. On the contrary, [56] proposes a new signal decomposition method which is inspired by EMD. Pattern aliasing and the lack of signal spectrum and reconstruction signals are the main drawbacks of the EMD algorithm. In that work, the authors propose a novel adaptive ensemble empirical mode decomposition in which only the first-order mode is decomposed after adding white noise to the original signal, then adding the noise ceaselessly to the residual signal. The proposed algorithm was applied to solve the problem of denoising of switchgear partial discharge signal.

EMD algorithm can be also used to analyze multivariate signals. In the work [57] multivariate EMD was combined with canonical correlation analysis to remove muscle artifacts from EEG recording. In the first step, the proposed method jointly decomposes the few-channel EEG recordings into multivariate intrinsic mode functions with the use of multivariate EMD. Although the canonical correlation analysis is used to decompose the reorganized multivariate EMD-based IMFs into the underlying sources. Artifact-free EEG signals

are obtained by reconstructing the signals using only artifact-free sources.

Multivariate signals are also under investigation in the field of geophysics. Seismic data are multi-channel thus Multivariate EMD or its modifications, for instance, fast multivariate empirical mode decomposition [58]. Empirical mode decomposition and its modifications are applied to denoise seismic signals [59], [60], [62], [64] and enhancing seismic reflections [61] or ground roll attenuation [63].

III. RESEARCH METHODOLOGY

A. EMPIRICAL MODE DECOMPOSITION

Empirical mode decomposition is a method for the adaptive splitting of a given signal into a finite set of oscillatory components. The acquired components are termed intrinsic mode functions. Although the process of their determination is named *sifting process*. In this section, the EMD procedure and its extensions are briefly reviewed. The details of the EMD method can be found, for example, in [12].

Consider a noisy signal $y(n)$ given by

$$y(n) = s(n) + \xi(n), \quad n = 1, 2, \ldots, N, \tag{1}$$

where $s(n)$ is the noise-free signal and $\xi(n)$ denotes the additive noise, and N stands for total number of data samples.

The IMFs, determined by applying the EMD procedure to the signal $y(n)$, are satisfying two conditions. The first one assumes that the number of zero-crossings and extrema are equal or the difference is no greater than 1. The second one assumes that the local mean value of the upper and lower envelopes is zero. The obtained IMF represents a certain frequency band of $y(n)$, that is, from the higher-frequency bands (the first couple of IMFs) to lower-frequency bands (the last few IMFs) [66].

When the decomposition is completed, the signal $y(n)$ is represented by

$$y(n) = \sum_{k=1}^{K} h_k(n) + r(n), \quad n = 1, 2, \ldots, N, \tag{2}$$

where $h_k(n)$ is a k-th IMF, $r(n)$ is the final residual, and K denotes the total number of IMFs. The component $r(n)$ can be considered as an IMF and it is denoted by $h_{K+1}(n)$. Hence we have

$$y(n) = \sum_{k=1}^{K+1} h_k(n), \quad n = 1, 2, \ldots, N. \tag{3}$$

The determined components are nearly orthogonal to each other [12], [67], and the set of modes is complete. Let us denote these sets of IMFs as

$$\mathcal{H} = \{h_k(n)\}_{k=1}^{K+1}. \tag{4}$$

Thus, we denote the EMD-based decomposition procedure of the signal $y(n)$ as

$$\mathcal{H} = \mathcal{E}(y(n)), \tag{5}$$

whereas the procedure of signal estimation based on determined IMFs (see eq. 3) is represented as

$$y(n) = \mathcal{E}^{-1}(\mathcal{H}). \tag{6}$$

The main advantage of the EMD approach relates to the fully data-driven nature of this method. The data-driven decomposition of the given signal means that the discovery of IMFs does not require rigid assumptions of the basic functions. This feature of the EMD method makes it suitable to analyze non-linear and non-stationary signals [12].

One of the weaknesses of the EMD method is the *mode mixing*. The phenomenon is defined in two ways. In the first one, the mode mixing means that a single intrinsic mode function contains signals of different scales. In the second one, a signal of similar scale exists in different IMF component [68], [69].

The EMD method also suffers from *aliasing*, that is, the overlapping of IMF spectra caused by a sub-Nyquist nature of extrema sampling and *end effect artifacts*. The letter one effect is related to an insufficient number of extrema detected at the beginning and at the end of the data segment. Lack of a sufficient number of extremes does not allow to determine the envelope successfully.

To overcome the mentioned drawbacks of the EMD method, the ensemble empirical mode decomposition (EEMD) is introduced in [68]. In the method, the white noise is added to the initial data. Thus, the determined IMFs are obtained by averaging the ensemble of the original signal and the different realization of white noise. The described modification of the original EMD algorithm makes the determined IMFs less prone to mode mixing.

The EEMD algorithm also has some disadvantages. One of them is that the signal decomposition cannot be complete. Additionally, applying the averaging techniques, by adding white noise, can lead to a different number of determined IMFs. It leads to the so-called reconstruction problem.

The complementary EEMD (CEEMD) is mainly proposed with the aim to deal with the reconstruction problem [70]. In general, the basic structure of CEEMD is similar to EEMD. The difference lies in the method of generation of white noise. In the CEEMD algorithm, the white noise is added in pairs to the original signal. Thus, we obtain two sets of ensemble IMFs.

Overcoming the problems of EEMD, the CEEMD algorithm introduces new ones. First, the ensemble process is time-consuming, whereas the number of trials is large [71]. It is related to the increasing number of sifting processes. The second problem is that the approach to determine IMFs in CEEMD can result in some false and meaningless IMF components. Moreover, some of them may not meet the definition of IMF's conditions [71].

The mentioned disadvantageous effects of the CEEMD method are minimized in CEEMDAN [72]. The main feature of the algorithm is the ability to control the noise level at every decomposition stage. Contrary to the previous method,

the reconstruction is complete and noise-free. The computational costs are less because the number of trials is decreased.

B. FILTER BANK PROPERTY OF EMD

An essential feature of the empirical mode decomposition method is that the determined set of IMF exhibits a quasi-dyadic filter bank property. Filter banks represent a set of bandpass filters used to isolate different frequency bands in the original signal. It was shown that IMFs determined by applying EMD algorithm provides frequency responses similar to the dyadic filter bank [66], [73].

The two well-known problems that can be solved utilizing the filter bank property of EMD methods are *denoising* and *detrending* of the signal. In general, the technique used to solve these problems is the partial reconstruction. For denoising, the new signal is constructed without modes that are identified as modes having high-frequency noise terms. In turn, for detrending, the modes with low-frequency noise terms are removed.

C. THE EMD-BASED DENOISING

The problem of signal denoising relies on estimating the signal $s(n)$ based on a noisy signal (1) for $n = 1, 2, \ldots, N$.

In the conventional EMD-based denoising approach reported in the literature, the first step is to separate relevant and irrelevant IMFs. The approach termed partial reconstruction was firstly proposed in [66]. In this approach, the estimated signal can be expressed as

$$\hat{s}(n) = \sum_{k=k_{th}}^{K+1} h_k(n), \quad n = 1, 2, \ldots, N, \tag{7}$$

where $1 < k_{th} \leq K + 1$ denotes the index discriminating relevant and irrelevant modes. There are various methods to find k_{th}. In the paper [15] an approach based on probabilistic similarity measure between the probability density function of the signal and that of each determined IMF. A similar approach to EMD-based signal denoising is introduced in the paper [20]. The work [74] considers the energy-based method to discriminate relevant and irrelevant modes. In turn, to find relevant modes the detrended fluctuation analysis is applied in the paper [75].

An alternative approach for the EMD-based signal denoising is a technique known from the wavelet domain. In this case, the reconstruction of the unknown signal bases on the whole set of the determined IMFs. However, before the reconstruction is performed, the IMFs are modified by applying wavelet-inspired thresholding procedure. The denoising procedure by thresholding method can be expressed as

$$\hat{s}(n) = \sum_{k=1}^{K+1} \tilde{h}_k(n), \quad n = 1, 2, \ldots, N, \tag{8}$$

where

$$\tilde{h}_k(n) = \begin{cases} h_k(n) & |h_k(n)| > T_k \\ 0 & |h_k(n)| \leq T_k, \end{cases} \tag{9}$$

for the hard-thresholding, and

$$\tilde{h}_k(n) = \begin{cases} \text{sgn}\,(h_k(n))\,(|h_k(n)| - T_k) & |h_k(n)| > T_k \\ 0 & |h_k(n)| \leq T_k, \end{cases} \tag{10}$$

for the soft-thresholding. T_k is the threshold of the k-th IMF.

In some cases, the thresholding procedure can be combined with the partial reconstruction. A generalized algorithm for reconstruction of the denoised signal is given by [20], [29]

$$\hat{s}(n) = \sum_{k=k_{th_1}}^{k_{th_2}} \tilde{h}_k(n) + \sum_{k=k_{th_2}+1}^{K+1} h_k(n), \quad n = 1, 2, \ldots, N. \tag{11}$$

The two additional parameters, k_{th_1} and k_{th_2} give flexibility on the exclusion of the noisy low-order IMFs and on the optional thresholding of the high-order ones. The paper [29] proposes an approach to determine these parameters.

The wavelet-inspired direct thresholding applied to the EMD-based denoising approach can introduce the discontinuity in the reconstructed signal [20], [29]. The remedy for the problem is the interval thresholding.

Let us consider an interval of zero-crossings $\mathbf{z}_k(n_j) = [z_k(n_j) \quad z_k(n_{j+1})]$ in the k-th IMF. Replacing the single sample by the interval $\mathbf{z}_k(n_j)$, the formulas (9) and (10) can be adopted to the following form [20]

$$\tilde{h}_k(\mathbf{z}_k(n_j)) = \begin{cases} h_k(\mathbf{z}_k(n_j)) & |h(r_k(n_j))| > T_k \\ 0 & |h(r_k(n_j))| \leq T_k, \end{cases} \tag{12}$$

for the hard-thresholding, and

$$\tilde{h}_k(\mathbf{z}_k(n_j))$$
$$= \begin{cases} \text{sgn}\,\big(h_k(\mathbf{z}_k(n_j))\big)\,\big(\big|h_k(\mathbf{z}_k(n_j))\big| - T_k\big) & |h(r_k(n_j))| > T_k \\ 0 & |h(r_k(n_j))| \leq T_k, \end{cases}$$
$$\tag{13}$$

for the soft-thresholding. The term $h(r_k(n_j))$ is the single extremum of the corresponding zero-crossing interval $\mathbf{z}_k(n_j)$ whereas $h_k(\mathbf{z}_k(n_j))$ represents all samples from $z_k(n_j)$ to $z_k(n_{j+1})$.

The problem of threshold (T_k) selection is considered for example in [20], [29].

D. SPARSITY IN SIGNAL DENOISING

The success of the sparsity in the field of signal processing has been proven. This idea has been successfully applied in various areas, including signal [76]–[78] and image [79] denoising.

It is well known that the problem of signal denoising is an ill-posed problem [80]. The ill-posedness nature of the denoising problem requires a regularization technique to be used for making a meaningful solution [81]. In other words, the solution that should have some of the required properties can be formulated as a constrained optimization problem.

Consider the estimation of signal $s(n)$ observed in additive white Gaussian noise (1). The components of (1) are represented as vectors in $\mathbb{R}^N$, that is, $y \in \mathbb{R}^N$, $s \in \mathbb{R}^N$, and $\boldsymbol{\xi} \in \mathbb{R}^N$.

The general form of the constrained optimization problem has the form

$$\hat{s} = \arg\min_{s}\{F(s) = \frac{1}{2}||y - s||_2^2 + \lambda\phi(s)\}, \qquad (14)$$

where $\phi(s) : \mathbb{R}^N \to \mathbb{R}$ is the regularization (or penalty) term and $\lambda > 0$ is the regularization parameter.

The optimization problem (14) can be formulated as convex or non-convex. The advantage of the convex formulation is that the wealth of convex optimization theory can be leveraged, and robust algorithms with guaranteed convergence are available [82]. In turn, the non-convex approaches are advantageous in terms of yielding sparser solutions. However, non-convex formulations are generally more difficult to solve. Generally, the convex approaches are based on sparsity-promoting convex penalty functions (e.g., ℓ_2 norm), whereas non-convex approaches are based on non-convex regularizers (e.g., ℓ_p pseudo norm with $p < 1$).

The choice of selecting a good penalty function $\phi(s)$ is an active area of the research. The typical form of regularizer $\phi(s)$ is ℓ_2 norm of forward difference vector which is defined as

$$\phi(s) = ||Ds||_2^2, \qquad (15)$$

where

$$D = \begin{bmatrix} -1 & 1 & & \\ & -1 & 1 & \\ & & \ddots & \ddots \\ & & & -1 & 1 \end{bmatrix}. \qquad (16)$$

is the first-order difference matrix.

In this case, the signal s is modeled as sparse and having a sparse derivative.

The norm ℓ_2 can be replaced, for example, by ℓ_0 in the optimization task (14) [54]. Taking into account new form of the norm, the optimization task (14) may be reformulated to the following ℓ_0-norm-based optimization task

$$\hat{s} = \arg\min_{s}\{F(s) = \frac{1}{2}||y - s||_2^2 + \lambda||s||_0\}. \qquad (17)$$

The approximate solution of the optimization problem (17) can be found by applying the following operator [83]

$$\hat{s} = \text{hard}(s, \lambda) = \begin{cases} s & |s| > \lambda \\ 0 & |s| \le \lambda. \end{cases} \qquad (18)$$

The mathematical formula (18) is termed the *hard-thresholding operator*. The operator on the real line is illustrated in Fig. 1.

It should be highlighted that when we apply hard-thresholding to vector s, we apply it component-wise, that is,

$$[\text{hard}(s, \lambda)]_n = \text{hard}(s(n), \lambda). \qquad (19)$$

Instead of ℓ_0 norm applied for the signal s we can use ℓ_1. It means that penalty function has the form $\phi(s) = ||s||_1$,

FIGURE 1. The hard-threshold operator.

whereas reformulated optimization problem may be written as

$$\hat{s} = \arg\min_{s}\{F(s) = \frac{1}{2}||y - s||_2^2 + \lambda||s||_1\}. \qquad (20)$$

The problem (20) is known in the literature as the *basis pursuit denoising*. The solution of the problem (20) can be found with [83]

$$\hat{s} = \text{soft}(s, \lambda) = \begin{cases} \text{sgn}(s)(|s| - \lambda) & |s| > \lambda \\ 0 & |s| \le \lambda, \end{cases} \qquad (21)$$

where $|| \cdot ||_1$ represents ℓ_0 norm of s.

The function soft(s, λ) is called the *soft-thresholding operator*. The operator is illustrated in Fig. 2.

FIGURE 2. The soft-threshold operator.

To improve the denoising performance of the algorithm, a variety of non-convex penalty terms have been developed to replace ℓ_0 and ℓ_1. For example, in [84] the authors applied $\ell_2 - \ell_0$ penalty function instead of ℓ_1-norm functional.

In the next section, we propose a framework to design denoising algorithms based on penalty functions that promote sparsity more strongly.

IV. SIGNAL DENOISING BASED ON SPARSIFYING IMFs
A. PROBLEM STATEMENT
In the paper, we consider the model of the measured discrete time series in the form of (1). The components of the model are represented as vectors (section III-D).

The problem is to design an algorithm that attenuates the noise $\boldsymbol{\xi}$ based on the given measurements y and keeps the signal s as intact as possible, that is, minimize difference $\boldsymbol{\epsilon} = \hat{s} - s$ ($\hat{s}$ stands for estimate of s).

B. TOTAL VARIATION DENOISING
An introduced in the previous section operators for hard and soft thresholding are effective and computationally efficient approaches for noise reduction. Unfortunately, these operators suffer from some drawbacks, that is, spurious noise

spikes and pseudo-Gibbs oscillations [85], [86]. An alternative is to use the total variation Denoising that attracted great attention in 1-D signal and 2-D image processing [87], [88].

In this section, we propose the EMD-based denoising method applying the total variation regularization as an alternative to interval hard-, and soft-thresholdings discussed in section III-C. To estimate IMFs of the noisy signal y we form following optimization problem [89]

$$\tilde{h}_k = \arg\min_{h_k} \left\{ F(\tilde{h}_k) = \frac{1}{2} \|\mathcal{E}(\mathbf{y}) - h_k\|_2^2 \right.$$
$$\left. + \lambda_0 \left\| D\mathcal{E}^{-1}(\mathcal{H}) \right\|_1 \right\}, \quad (22)$$

where $k = 1, 2, \ldots, K + 1$, λ_0 is the penalty parameter $\lambda_0 > 0$, $\mathcal{H}$ stands for sets of IMFs represented as vectors, that is, $\mathcal{H} = \{h_k\}_{k=1}^{K+1}$, D represents the first-order difference matrix (16).

The regularizer in (22), that is, $\left\| D\mathcal{E}^{-1}(\mathcal{H}) \right\|_1$, is the total variation of the estimate of signal s [34].

The solution of the problem (22) is [33]

$$\tilde{h}_k = \text{tvd}(h_k, \lambda_0), \quad k = 1, 2, \ldots, K + 1, \quad (23)$$

where tvd() is the total variation denoising algorithm. It should be mentioned that there is no explicit solution for TVD, but recently, the direct algorithm to determine the exact solution (5) by the *taut string algorithm* is proposed in [90].

It is worth to emphasize that one of the main drawbacks of the denoising algorithms based on TV regularization is undesirable staircase artifacts [34].

Having determined $\tilde{\mathbf{h}}_k$, we can estimate the signal s from formula (6). Because the algorithm proposed in the section is based on the EEMD decomposition method and total variation denoising technique, we named it as EEMD-TVD (Algorithm 1).

Algorithm 1 The EEMD-TVD Algorithm

Require: noisy signal $\mathbf{y}$, parameter λ_0
Ensure: estimated signal $\hat{\mathbf{s}}$
$\quad \mathcal{H} = \mathcal{E}(\mathbf{y})$ {EEMD-based signal decomposition}
$\quad$ **for** $k = 1$ to $K + 1$ **do**
$\quad\quad \tilde{\mathbf{h}}_k = \text{tvd}(\mathbf{h}_k, \lambda_0)$ {Total Variation Denoising}
$\quad$ **end for**
$\quad \tilde{\mathcal{H}} = \{\tilde{\mathbf{h}}_k\}_{k=1}^{K+1}$
$\quad \hat{\mathbf{s}} = \mathcal{E}^{-1}\left(\tilde{\mathcal{H}}\right)$ {EEMD-based signal recomposition}

C. FUSED LASSO SIGNAL APPROXIMATOR

The TVD approach can be extended by adding an extra penalty term. Let us state the new optimization problem as an extension of the problem (22)

$$\tilde{h}_k = \arg\min_{h_k} \left\{ F(\tilde{h}_k) = \frac{1}{2} \|\mathcal{E}(\mathbf{y}) - h_k\|_2^2 + \lambda_1 \|h_k\|_1 \right.$$
$$\left. + \lambda_0 \left\| D\mathcal{E}^{-1}(\mathcal{H}) \right\|_1 \right\}, \quad (24)$$

where λ_0 and λ_1 are the penalty parameters, $k = 1, 2, \ldots, K + 1$. The additional regularizer in (24), that is, $\|h_k\|_1$ controls the sparsity (number of zeros) [91].

The solution to (24) [92] (*Lemma A.1*) is readily obtained by applying the Fused LASSO method [44]

$$\tilde{h}_k = \text{soft}(\text{tvd}(h_k, \lambda_0), \lambda_1), \quad (25)$$

where soft() is the soft-thresholding operator. For the EMD-based method, the formula (13) should be applied for the soft-thresholding.

The algorithm to signal denoising based on (25) is named EEMD-FLS (Algorithm 2) since the signal decomposition is based on the EEMD method.

Algorithm 2 The EEMD-FLS Algorithm

Require: noisy signal $\mathbf{y}$, parameters: λ_0, λ_1
Ensure: estimated signal $\hat{\mathbf{s}}$
$\quad \mathcal{H} = \mathcal{E}(\mathbf{y})$ {EEMD-based signal decomposition}
$\quad$ **for** $k = 1$ to $K + 1$ **do**
$\quad\quad \tilde{\mathbf{h}}_k = \text{soft}(\text{tvd}(\mathbf{h}_k, \lambda_0), \lambda_1)$ {Fused LASSO Signal Approximator with soft-thresholding (eq. 13)}
$\quad$ **end for**
$\quad \tilde{\mathcal{H}} = \{\tilde{\mathbf{h}}_k\}_{k=1}^{K+1}$
$\quad \hat{\mathbf{s}} = \mathcal{E}^{-1}\left(\tilde{\mathcal{H}}\right)$ {EEMD-based signal recomposition}

D. NON-CONVEX SPARSE REGULARIZATION

Instead of the ℓ_1 norm, it is possible to formulate the optimization problem with another regularization function. Let us generalize problems (22) and (24)

$$\tilde{h}_k = \arg\min_h \left\{ F(\tilde{h}_k) = \frac{1}{2} \|\mathcal{E}(\mathbf{y}) - h_k\|_2^2 \right.$$
$$\left. + \sum_{k=1}^{K+1} \lambda_k \phi(h_k; a_k) + \lambda_0 \left\| D\mathcal{E}^{-1}(\mathcal{H}) \right\|_1 \right\}, \quad (26)$$

where λ_k are Lagrangian multipliers and ϕ is a non-convex sparsity inducing penalty function. In the previous formulation of the optimization problem, the penalty term ϕ is based on ℓ_1 norm. However, it is possible to use several other functions instead of considered ℓ_1 norm [93].

1) PENALTY FUNCTION

In general, the penalty function ϕ should satisfy the following conditions [34]:

1) ϕ is continuous on $\mathbb{R}$;
2) ϕ is twice continuously differentiable, increasing, and concave on $\mathbb{R}_+$;
3) $\phi(h; 0) = |h|$;
4) $\phi(0; a) = 0$;
5) $\phi(-h; a) = \phi(h; a)$;
6) $\phi'(0^+; a) = 1$;
7) $\phi''(h; a) \geq a$ for all $h \neq 0$;

In this study, we use the SCAD function (smoothly clipped absolute deviation) as it induces sparsity more strongly.

The SCAD function has the following general form [94], [95]

$$\phi(h; T, a) = \begin{cases} T|h| & |h| \leq T \\ -\left(\dfrac{|h|^2 - 2aT|h| + T^2}{2(a-1)}\right) & \lambda \leq |h| \leq aT \\ \dfrac{(a+1)T^2}{2} & |h| > aT, \end{cases} \quad (27)$$

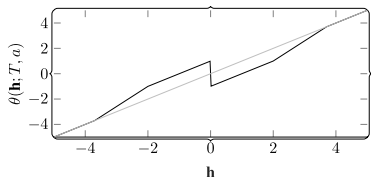where T denotes threshold, a is parameter $a > 2$. The SCAD-thresholding operator is illustrated in Fig. 3.

FIGURE 3. The SCAD-thresholding operator.

2) VARIABLE SPLITTING

To solve the formulated optimization problem (26) we applied the technique known as the variable splitting. The technique can be applied when the objective function of the optimization problem is the sum of two functions, one of which is written as the composition of two functions

$$\tilde{h} = \arg\min_h g_1(h) + g_2(f(h)). \quad (28)$$

Based on the variable splitting, we created a new variable, for example u, to serve the argument of function g_2, under the constraint that $f(h) = u$ [96]. This leads to the following constrained optimization problem

$$\{\tilde{h}, \tilde{u}\} = \arg\min_{h,u} g_1(h) + g_2(u)$$
$$\text{s.t.} \quad f(h) = u, \quad (29)$$

which is equivalent to the unconstrained problem (28). The premise for the application of the variable splitting is that it could be easier to solve the constrained problem (29) than its unconstrained version (28). It is possible by introducing the auxiliary variable u. The variable allow the decoupling of the two functions g_1 and g_2. Each of them applies to one specific optimization variable, that is, h and u.

3) ALTERNATING DIRECTION METHOD OF MULTIPLIERS

Based on the variable splitting technique, we divide (26) into two functions

$$g_1(h_k) = \frac{1}{2}\|\mathcal{E}(\mathbf{y}) - h_k\|_2^2 + \sum_{k=1}^{K+1} \lambda_k \phi(h_k; a_k), \quad (30)$$

$$g_2(u_k) = \lambda_0 \left\| D\mathcal{E}^{-1}(\mathcal{U}) \right\|_1, \quad (31)$$

in which $u_k = h_k$, and $\mathcal{U} = \{u_k\}_{k=1}^{K+1}$.

Taking into account (30) and (31), the formula (26) can be rewritten in the form [98]

$$\{\tilde{h}_k, \tilde{u}_k\} = \arg\min_{h_k, u_k} \{g_1(h_k) + g_2(u_k)\}$$
$$\text{s.t.} \quad u_k = h_k. \quad (32)$$

A typical way for solving the introduced constrained optimization problem (32) is to deploy the augmented Lagrangian approach (see section III) [99]. Let us rewrite the augmented Lagrangian function of the problem (32) in the form [98]

$$L(h_k, u_k, \mu) = g_1(h_k) + g_2(u_k) + \frac{\mu}{2}\|u_k - h_k - d_k\|_2^2, \quad (33)$$

where μ is a positive parameter that balances the penalization. In general, the value of μ is varied. Typically, the value of μ is initially low, but as the solution converges, its value is progressively increased.

For convex function L, ADMM algorithm is guaranteed to converge for any positive value of μ [97].

To solve the formulated optimization problem, an iterative scheme should be applied [89]

$$\tilde{h}_k^{(m+1)} = \arg\min_{h_k} \left\{ g_1\left(h_k^{(m)}\right) + \frac{\mu}{2}\left(u_k^{(m)} - h_k^{(m)} - d_k^{(m)}\right)\right\}, \quad (34a)$$

$$\tilde{u}_k^{(m+1)} = \arg\min_{u_k} \left\{ g_2\left(u_k^{(m)}\right) + \frac{\mu}{2}\left(u_k^{(m)} - h_k^{(m)} - d_k^{(m)}\right)\right\}, \quad (34b)$$

$$\tilde{d}_k^{(m+1)} = d_k^{(m)} - \left(u_k^{(m)} - h_k^{(m)}\right). \quad (34c)$$

The algorithm (34a – 34c) is named alternating direction method of multipliers. Instead of the ADMM method, the SALSA algorithm (split augmented Lagrangian shrinkage algorithm) can also be applied [96].

The problem (34a) can be solved exactly [89]. Based on [34], the formula to estimate h_k has the form

$$\tilde{h}_k = \theta\left(p_k; \frac{\lambda_k}{\mu + 1}, a_k\right), \quad (35)$$

where θ is the thresholding operator based on SCAD penalty function (27) and

$$p_k = (\mathcal{E}(y) + \mu(u_k - d_k)/(\mu + 1)). \quad (36)$$

The equation (34b) can be solved exactly as well. The formulas were derived in [33]

$$v_k = d_k + \tilde{h}_k, \quad (37a)$$

$$u_k = v_k + \mathcal{E}(tvd(\mathcal{E}^{-1}(\mathcal{V}), \lambda_0/\mu) - \mathcal{E}^{-1}(\mathcal{V})), \quad (37b)$$

where $\mathcal{V} = \{v_k\}_{k=1}^{K+1}$.

Having estimated $\tilde{h}_k$ we can calculate the signal from (6). The algorithm to signal denoising, proposed in this section is named EEMD-NSR (Algorithm 3) since the signal decomposition is based on EEMD approach and non-convex sparse regularization.

4) REMARK

In the formula (35) we apply the thresholding operator based on SCAD function (38), as shown at the bottom of the next page. The direct application of SCAD function to IMFs is not possible, because it introduces the discontinuity in the estimated signal. Instead of (38), we propose the interval SCAD thresholding operator (39), as shown at the bottom of the next page. The SCAD-based operator is adopted for IMFs.

Algorithm 3 The EEMD-NSR Algorithm

Require: noisy signal $\mathbf{y}$, parameters: $\lambda_0, \lambda_1, \ldots, \lambda_{K+1}, \mu$
Ensure: estimated signal $\hat{\mathbf{s}}$

  for $k = 1$ to $K + 1$ **do**
    $a_k = 1/\lambda_k$
  end for
  $\mathcal{H} = \mathcal{E}(\mathbf{y})$ {EEMD-based signal decomposition}
  repeat
    $\mathbf{p}_k^{(m+1)} = \left(\mathcal{E}(\mathbf{y}) + \mu \left(\mathbf{u}_k^{(m)} - \mathbf{d}_k^{(m)} \right) / (\mu + 1) \right)$
    for $k = 1$ to $K + 1$ **do**
      $\tilde{\mathbf{h}}_k^{(m)} = \theta \left(p_k^{(m)}(\mathbf{z}_k(n_j)); \frac{T_k}{\mu+1}, a_k \right)$ {SCAD-thresholding operator (eq. 39)}
    end for
    $\mathbf{v}^{(m+1)} = \mathbf{d}^{(m)} + \tilde{\mathbf{h}}_k^{(m)}$
    $\mathbf{u}^{(m+1)} = \mathbf{v}^{(m)} + \mathcal{E}(tvd(\mathcal{E}^{-1}(\mathbf{v}^{(m)}), \lambda_0/\mu) - \mathcal{E}^{-1}(\mathbf{v}^{(m)}))$
    $\mathbf{d}^{(m+1)} = \mathbf{d}^{(m)} - (\mathbf{u}^{(m)} - \tilde{\mathbf{h}}_k^{(m)})$
    $m = m + 1$
  until convergence
  $\tilde{\mathcal{H}} = \{\tilde{\mathbf{h}}_k\}_{k=1}^{K+1}$
  $\hat{\mathbf{s}} = \mathcal{E}^{-1}\left(\tilde{\mathcal{H}} \right)$ {EEMD-based signal recomposition}

---

5) PARAMETERS SELECTION

The proposed algorithm requires the parameters a_k, $\lambda_0, \lambda_1, \ldots, \lambda_{K+1}$. To assure convexity of (34a), the value of the parameter a_k should be taken from the range $0 \leq a_k \leq \frac{1}{\lambda_k}$ for $k = 1, 2, \ldots, K+1$. For example in work [34] the authors suggest using $a_k = \frac{1}{\lambda_k}$ (or slightly less, e.g. $a_k = \frac{0.95}{\lambda_k}$) to maximally induced sparsity.

In turn, according to [20], the parameters $\lambda_1, \lambda_2, \ldots, \lambda_{K+1}$ can be determined from

$$\lambda_k = \sqrt{2E_k \ln N}, \tag{40}$$

where N is the number of signal data points, and E_k is obtained from formulas

$$E_1 = \frac{\text{median}(|h_1(n)|)}{0.6745}, \tag{41}$$

$$E_k = \frac{E_1^2}{\gamma} \rho^{-k}, \quad k = 2, 3, \ldots, K+1, \tag{42}$$

where γ and ρ are parameters to be estimated by large number of independent noise realizations and their IMFs [20] and n stands for n-th data point. In [100] the authors proposed the values 0.719 for γ and 2.01 for ρ.

To set λ_0, based on the papers [34], [101] we suggest the following formula

$$\lambda_0 = \frac{1}{4}(1 - \eta)\sqrt{N}E_k, \tag{43}$$

in which $0 < \eta < 1$. In accordance with [34] we suggest setting $\eta = 0.95$.

In (43) we used E_k instead the standard deviation of noise σ_k based on the remark in [20].

V. NUMERICAL RESULTS

The performance of the proposed methods is evaluated based on the signal-to-noise ratio:

$$SNR = 10 \log_{10} \left(\frac{\sum_{n=1}^{N} s(n)}{\sum_{n=1}^{N} \left(s(n) - \hat{s}(n)\right)^2} \right), \tag{44}$$

where $\hat{s}(n)$ is the denoised signal, $s(n)$ is the original signal and N is the length of the data. The SNR ratio is applied for tests in which synthetic signals are used. To test the performance of the proposed methods, based on the real-world signal, we applied the Allan variance (AV). The Allan variance is a method of representing the root means square of noise as a function of averaging time. The method can be used to determine the characteristics of the underlying random processes in the measurements. In the work AV is denoted as $\sigma(\tau)$ where τ represents average (correlation) time.

In this work, we test the performance of the proposed algorithms for the signal denoising problem. The first of the proposed algorithm is EEMD-TVD (please see section IV-B), the second one is EEMD-FLS (section IV-C), and the third one is EEMD-NSR (section IV-D). The presented methods are based on the following implementation of the EEMD method [72]. The proposed methods are compared with the EEMD interval thresholding algorithm (EEMD-CIIT). The EEMD-CIIT method based on the EMD-CIIT algorithm introduced in [29]. Additionally, for comparison purposes,

$$\theta\left(p_k(n); \frac{T_k}{\mu+1}, a_k\right) = \begin{cases} \text{sgn}(p_k(n)) \max \left(|p_k(n)| - \dfrac{T_k}{\mu+1} \right) & |p_k(n)| \leq 2\dfrac{T_k}{\mu+1} \\ \dfrac{(a_k - 1)p_k(n) - a_k \dfrac{T_k}{\mu+1} \text{sgn}(p_k(n))}{a_k - 2} & 2\dfrac{T_k}{\mu+1} < |p_k(n)| \leq a_k \dfrac{T_k}{\mu+1} \\ p_k(n) & |p_k(n)| > a_k \dfrac{T_k}{\mu+1}m, \end{cases} \tag{38}$$

$$\theta\left(p_k(\mathbf{z}_k(n_j)); \frac{T_k}{\mu+1}, a_k\right) = \begin{cases} p_k(\mathbf{z}_k(n_j)) \dfrac{\max\left(0, |p_k(r_k(n_j))| - T_k\right)}{|p_k(r_k(n_j))|} & |p_k(r_k(n_j))| \leq 2\dfrac{T_k}{\mu+1} \\ p_k(\mathbf{z}_k(n_j)) \dfrac{(a - 1)|p_k(r_k(n_j))| - aT}{(a - 2)|p_k(r_k(n_j))|} & 2\dfrac{T_k}{\mu+1} < |p_k(r_k(n_j))| \leq a\dfrac{T_k}{\mu+1} \\ 1 & |p_k(r_k(n_j))| > a\dfrac{T_k}{\mu+1}. \end{cases} \tag{39}$$

we applied a reference approach not based on the empirical mode decomposition algorithm, that is, the non-local means filter [7], [102]. NLM algorithm's parameters, that is, the *bandwidth*, the *patch half-width*, and the *neighborhood half-width* are determined empirically [8], [103].

We also used the Wavelet-TV denoising algorithm as a reference. In this approach, the sparse theory was applied to induce wavelet-domain sparsity [34].

At the same time, we test the robustness of designed algorithms to conduct a sequence of simulations under different signals and SNR.

This section includes the test on simulated (noise-corrupted synthetic signals) and real-world noisy signals (i.e., measurements data from gyroscope).

A. RESULTS ON SIMULATED SIGNALS

1) SIGNALS DETAILS

In this section, we compare the performance of the proposed methods with other denoising algorithms. Coupled with performance analysis we study the robustness of suggested methods under various signal forms and SNR.

In our research, we consider *Blocks*, *Bumps* and *Doppler* (Fig. 4) signals of length 4096 samples. The signals are generated by the *Wavelab* [104] function *MakeSignal*.
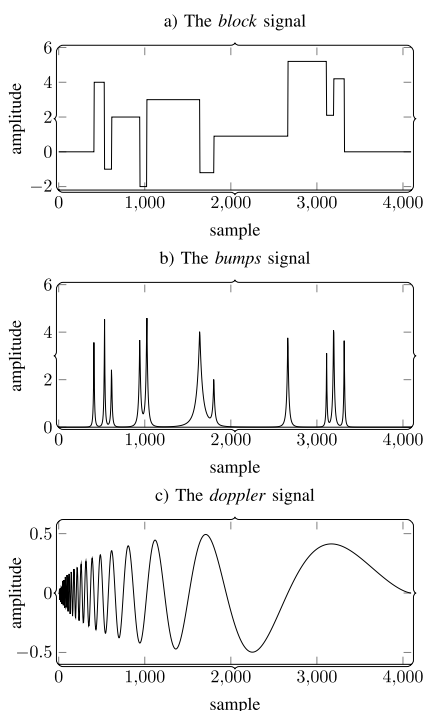
FIGURE 4. The synthetic signals used in studies.

We add the noise to the generated signals to assess the denoising performance of the proposed methods. In our research, we study the signals with different SNR (0, 5, 10 and 15 [dB]) and different short and long dependencies. Thus, the signals are corrupted by fractional Gaussian noise (fGn) with the Hurst exponent equal to H = 0.2,

H = 0.5 and H = 0.8. The value 0.5 of Hurst indicates the absence of long-range dependence. On the other hand, when $0.5 < H < 1$ the signal has a long-range dependency property, and when $0 < H < 0.5$ the signal indicates strong negative correlation (the opposite of the long-range dependency) [105].

To verify the effectiveness and robustness of the proposed methods, we carried out 30 trials for noisy *Block*, *Bumps*, and *Doppler* signals with various SNR ratios (0, 5, 10 and 15 [dB]) under various levels of the Hurst exponent (0.2, 0.5, 0.8). As an example, the synthetic signals (*Block*, *Bumps*, and *Doppler*) contaminated by fractal Gaussian noise with parameter H = 0.8 and SNR = 15 [dB] is shown in Fig. 5.

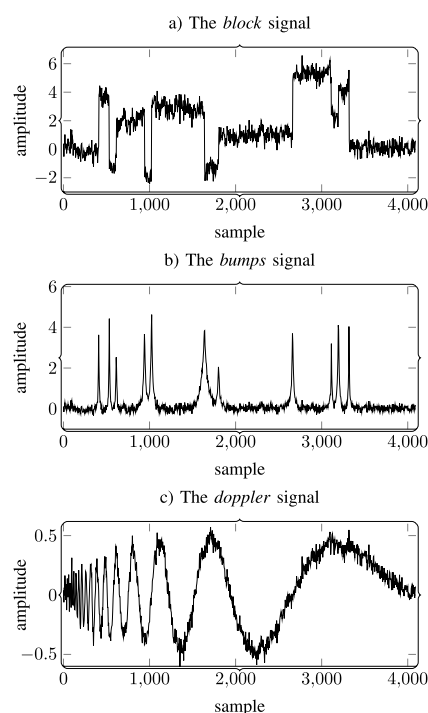FIGURE 5. The synthetic signals in noise (fractional Gaussian noise H = 0.8; SNR = 15 [dB]).

2) EXPERIMENTAL RESULTS

The results for *Blocks*, *Bumps* and *Doppler* with SNR ratio equal to 15 [dB] and H = 0.8 are displayed in Fig. 6. In Tables 1 to 9 more results for SNR varying from 0 to 15 [dB] and the Hurst exponent equals 0.2, 0.5 and 0.8 are presented.

Let us take the *Blocks* signal as an example. The Tables 1, 4, and 7 indicate that, in almost all cases, one of the proposed methods (EEMD-NSR) gives the best results when the Hurst exponent equal to H = 0.2 and H = 0.5. For H = 0.2 and SNR = 0 and SNR = 5 [dB] the best performance was achieved by Wavelet-TV. For Hurst H = 0.8 the best performance is obtained from non-local means. The methods EEMD-TVD and EEMD-FLS give better results
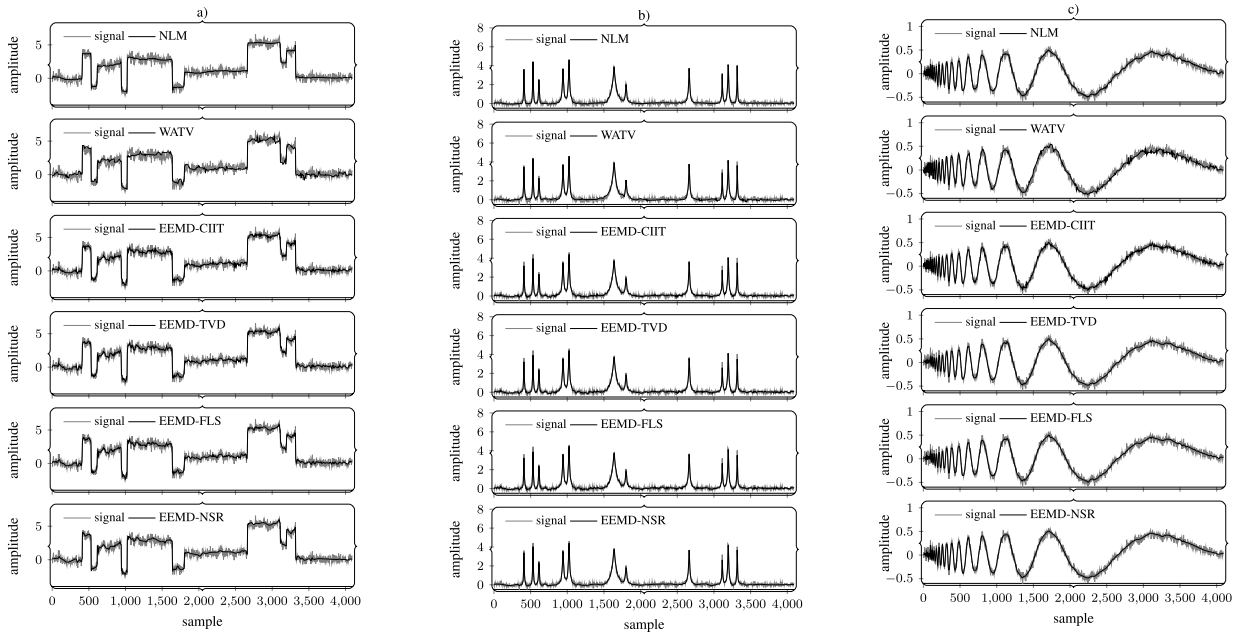
FIGURE 6. The results of signal denoising (*Blocks* (a), *Bumps* (b), *Doppler* (c) signals, fractional Gaussian noise H = 0.8, SNR = 15 [dB]).

TABLE 1. The denoising results for the *Blocks* signal, fractional Gaussian noise H = 0.2.

Methods		SNR [dB]			
		0	5	10	15
NLM	mean	13,94	20,75	25,72	29,57
	std	0,15	0,18	0,32	0,49
WATV	mean	**18,60**	**22,34**	27,03	29,47
	std	0,66	0,64	0,53	0,93
EEMD-CIIT	mean	15,68	18,37	22,05	25,02
	std	0,51	0,32	0,22	0,52
EEMD-TVD	mean	16,97	20,10	23,70	27,91
	std	0,25	0,22	0,30	0,31
EEMD-FLS	mean	18,14	21,43	24,74	27,76
	std	0,29	0,32	0,37	0,30
EEMD-NSR	mean	18,48	21,88	**27,28**	**30,48**
	std	0,45	0,44	1,57	2,05

TABLE 2. The denoising results for the *Bumps* signal, fractional Gaussian noise H = 0.2.

Methods		SNR [dB]			
		0	5	10	15
NLM	mean	12,73	17,83	22,07	25,68
	std	0,29	0,37	0,35	0,32
WATV	mean	**16,14**	**21,60**	22,36	**26,10**
	std	0,22	0,46	0,29	0,30
EEMD-CIIT	mean	14,08	18,04	21,92	25,16
	std	0,32	0,34	0,26	0,32
EEMD-TVD	mean	13,72	17,79	21,38	25,02
	std	0,22	0,28	0,29	0,29
EEMD-FLS	mean	15,22	19,07	22,77	25,88
	std	0,31	0,32	0,36	0,30
EEMD-NSR	mean	16,00	20,29	**22,86**	25,36
	std	1,10	0,40	0,46	0,33

than EEMD-CIIT in all cases. However, their performance is weaker than both EEMD-NSR and the reference methods NLM and Wavelet-TV. Therefore, these results demonstrate that one of the proposed methods, that is, EEMD-NSR is more suitable for denoising piecewise constant signals with jumps (like the *Blocks* signal) for the white Gaussian noise (H = 0.5) and the noise with strong negative correlation (when 0 < H < 0.5). The results of *Blocks* signal denoising is presented in Fig. 6.

Another example is the *Bumps* signal. The best results for H = 0.2 and SNR equals 10 [dB] is obtained from the EEMD-NSR method. While WATV is the most effective for SNR equals 0, 5 and 15 [dB]. For Hurst H = 0.5 and 0.8 the best performance is achieved in almost all cases from the reference method, that is, NLM. Only for H = 0.5

and SNR = 0 the best performance was obtained from EEMD-NSR. The method EEMD-TVD gives better results than EEMD-CIIT but weaker than EEMD-NSR, EEMD-FLS, non-local means and Wavelet-TV. The results show that for signals like the *Bumps* signal one of the proposed methods, that is, EEMD-NSR can be used to remove noises with strong negative correlation (when 0 < H < 0.5). In other cases, for H = 0.5 and 0.5 < H < 1, the reference methods NLM and WATV outperforms the other algorithms. In turn, the results for denoising the *Bumps* signal is presented in Fig. 6.

For *Doppler* signals, the EEMD-NSR method is more successful than other methods. Only for H = 0.2 and SNR = 10 and 15 [dB], the EEMD-FLS method gives better results. Moreover, the proposed approaches, EEMD-FLS and EEMD-TVD, are better than the EEMD-CIIT algorithm.

TABLE 3. The denoising results for the *Doppler* signal, fractional Gaussian noise H = 0.2.

Methods		SNR [dB]			
		0	5	10	15
NLM	mean	14,07	18,58	23,75	29,09
	std	0,13	0,20	0,28	0,31
WATV	mean	18,86	23,21	27,34	31,79
	std	0,23	0,41	0,32	0,53
EEMD-CIIT	mean	17,24	21,00	24,69	28,99
	std	0,42	0,43	0,47	0,40
EEMD-TVD	mean	19,02	22,87	26,92	31,06
	std	0,39	0,41	0,41	0,35
EEMD-FLS	mean	19,18	22,92	**27,40**	**32,12**
	std	0,32	0,37	0,42	0,36
EEMD-NSR	mean	**20,60**	**23,86**	26,39	29,12
	std	0,53	0,52	0,45	0,45

TABLE 4. The denoising results for the *Blocks* signal, fractional Gaussian noise H = 0.5.

Methods		SNR [dB]			
		0	5	10	15
NLM	mean	14,51	19,72	**24,77**	28,64
	std	0,36	0,41	0,46	0,49
WATV	mean	15,08	20,65	24,23	29,46
	std	0,60	0,67	0,54	0,56
EEMD-CIIT	mean	13,38	16,56	19,92	23,73
	std	0,41	0,37	0,37	0,30
EEMD-TVD	mean	14,64	17,68	21,46	25,56
	std	0,40	0,34	0,31	0,31
EEMD-FLS	mean	14,84	17,86	21,50	25,71
	std	0,41	0,37	0,35	0,32
EEMD-NSR	mean	**15,31**	**20,86**	24,54	**30,24**
	std	0,54	0,67	0,73	0,73

TABLE 5. The denoising results for the *Bumps* signal, fractional Gaussian noise H = 0.5.

Methods		SNR [dB]			
		0	5	10	15
NLM	mean	12,37	**17,38**	**21,73**	**25,81**
	std	0,46	0,39	0,30	0,27
WATV	mean	10,73	15,62	16,03	22,11
	std	0,30	0,31	0,10	0,32
EEMD-CIIT	mean	11,20	15,26	19,34	23,41
	std	0,40	0,43	0,37	0,33
EEMD-TVD	mean	11,71	15,51	19,49	23,57
	std	0,34	0,32	0,28	0,24
EEMD-FLS	mean	11,78	15,60	19,62	23,48
	std	0,37	0,35	0,29	0,24
EEMD-NSR	mean	**12,76**	16,97	20,54	23,94
	std	0,49	0,44	0,37	0,84

TABLE 6. The denoising results for the *Doppler* signal, fractional Gaussian noise H = 0.5.

Methods		SNR [dB]			
		0	5	10	15
NLM	mean	13,89	17,23	21,54	26,61
	std	0,41	0,42	0,37	0,41
WATV	mean	13,96	18,21	23,22	27,29
	std	0,45	0,39	0,55	0,45
EEMD-CIIT	mean	14,11	17,86	21,71	25,79
	std	0,74	0,67	0,61	0,62
EEMD-TVD	mean	15,89	17,81	23,50	27,61
	std	0,63	0,49	0,54	0,44
EEMD-FLS	mean	15,80	18,83	23,51	27,66
	std	0,66	0,52	0,48	0,43
EEMD-NSR	mean	**16,19**	**20,01**	**23,89**	**27,72**
	std	0,77	0,69	0,66	0,45

TABLE 7. The denoising results for the *Blocks* signal, fractional Gaussian noise H = 0.8.

Methods		SNR [dB]			
		0	5	10	15
NLM	mean	**7,46**	**12,29**	**17,82**	**22,55**
	std	1,23	1,13	1,36	1,24
WATV	mean	6,72	10,48	16,41	21,27
	std	0,93	0,63	0,89	0,88
EEMD-CIIT	mean	6,73	10,84	15,19	19,71
	std	1,09	0,83	0,72	0,63
EEMD-TVD	mean	7,42	11,18	15,27	19,75
	std	1,20	0,88	0,68	0,58
EEMD-FLS	mean	7,43	11,19	15,25	19,73
	std	1,20	0,87	0,67	0,56
EEMD-NSR	mean	6,84	11,28	16,16	21,07
	std	1,07	0,88	0,82	0,90

TABLE 8. The denoising results for the *Bumps* signal, fractional Gaussian noise H = 0.8.

Methods		SNR [dB]			
		0	5	10	15
NLM	mean	**7,19**	**12,07**	**16,59**	**21,11**
	std	1,24	1,24	1,02	0,97
WATV	mean	6,27	10,92	15,80	20,41
	std	0,92	0,80	0,85	0,68
EEMD-CIIT	mean	6,20	10,44	14,91	19,39
	std	0,85	0,72	0,62	0,56
EEMD-TVD	mean	5,70	9,96	14,40	18,93
	std	0,75	0,61	0,52	0,42
EEMD-FLS	mean	5,70	9,97	14,40	19,00
	std	0,73	0,63	0,52	0,47
EEMD-NSR	mean	6,47	10,87	15,30	20,08
	std	0,95	0,88	0,79	0,74

We observe that the EEMD-NSR method is well suited for denoising high-frequency oscillating signals like *Doppler*. The results of the *Doppler* signal denoising is presented in Fig. 6.

For illustrative purposes, in Fig. 7 (on the left) we present the results of *Doppler* signal decomposition, whereas in Fig. 7 (on the right), the results of sparsification of the first seven IMFs of *Doppler* signal by the EEMD-NSR approach is presented.

We can also observe from the results (Tables 1 to 9) that the performance of EEMD-CIIT, the proposed methods (EEMD-TVD, EEMD-FLS, and EEMD-NSR), and reference non-local means and wavelet-TV decreases with the value of the Hurst exponent close to 1. When the Hurst exponent

TABLE 9. The denoising results for the *Doppler* signal, fractional Gaussian noise H = 0.8.

Methods		SNR [dB]			
		0	5	10	15
NLM	mean	7,09	11,37	15,67	20,25
	std	1,06	0,92	0,72	0,66
WATV	mean	7,00	11,54	15,80	20,76
	std	1,07	0,96	0,75	0,78
EEMD-CIIT	mean	6,72	11,11	15,53	19,94
	std	1,20	0,99	0,81	0,74
EEMD-TVD	mean	7,81	11,93	16,20	20,62
	std	1,31	1,08	0,87	0,70
EEMD-FLS	mean	7,81	11,93	16,17	20,61
	std	1,31	1,07	0,86	0,71
EEMD-NSR	mean	**8,11**	**12,39**	**16,82**	**21,21**
	std	1,38	1,26	1,01	0,90
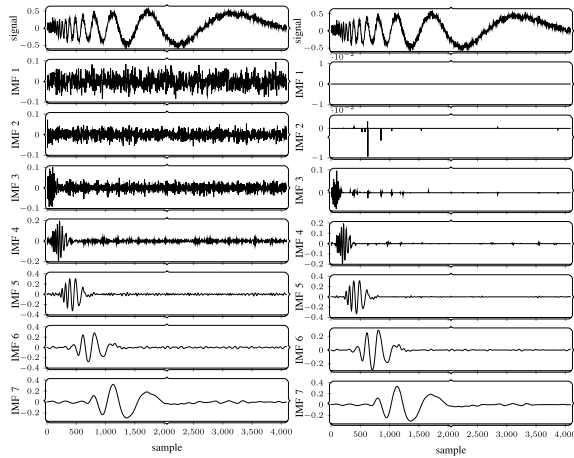
FIGURE 7. The results of the *Doppler* signal decomposition by the EEMD method (the first 7 IMFs) for the fractional Gaussian noise H = 0.8, SNR = 15 [dB] (left), and the results of sparsification of the first 7 IMFs (right).

is close to 1, we can also observe that standard deviation increases. We should emphasize the fact that EEMD-NSR is a slightly higher standard deviation than other EMD-based approaches (EEMD-CIIT, EEMD-TVD, and EEMD-FLS).

The conducted investigations also have shown that the proposed methods can obtain satisfying robustness under various signal forms and SNR.

B. RESULTS ON REAL-WORLD SIGNALS
1) SIGNALS DETAILS
The micro-electromechanical system can be used to manufacture, for example, gyroscopic sensors. The gyroscope is a device for measuring the angular velocity of a moving object. It is kind of sensor widely used in, for instance, human motion tracking and detection due to its small size, low cost, long lifespan, and no moving parts. Gyroscope sensor has some disadvantages related to its vulnerability to inferences such as temperature, vibration and pressure [20]. These phenomena result in different noise effects that degrade the accuracy of

the data and limit its applications. One of the examples of such noise is drift.

The drift is a weak nonlinear, non-stationary, and sensitive to environmental conditions example of the noise, and a critical research question is how to suppress it from gyroscopic measurements. It is crucial since the enhancement of the accuracy of these data will significantly improve the precision of the systems based on the signals acquired from gyroscopes.

The issues of the gyroscopic signal denoising have been presented in many research papers. For example, in work [106], the wavelet transform was applied to suppress noises in gyroscopic data. In turn, Neural Networks and Kalman filter were used, for example, in [107] and [108], respectively. Empirical mode decomposition was also applied to enhance the gyroscopic measurements [20], [21], [29], [109], [111].

In [21], [111] hybrid methods combining EMD and FLP (forward linear prediction) are used to enhance the gyroscopic data. In turn, in the papers [20], [29] empirical mode decomposition based on interval thresholding was suggested to gyroscope signal denoising. Although work [109] introduces a method based on partial reconstruction technique for EMD algorithm.

To verify feasibility and effectiveness of the proposed signal denoising algorithms, practical experimental data were collected from the three-axis gyroscope. The experiments were performed in static conditions. It means that during the tests the gyroscope is kept stationary. The measurements were collected by the *Shimmer3* unit with a temperature of 22°C and with the sample frequency equal to 50 [Hz] (Fig. 8 a,b,c, on the left). The angular rate was recorded for x-, y-, and z-axis.
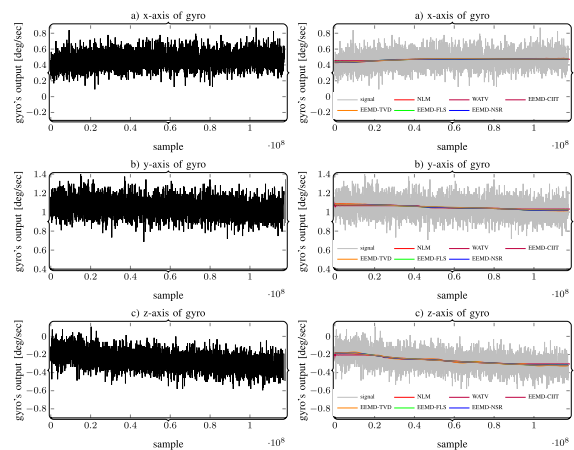

FIGURE 8. The original signals from gyroscope (left), and the results of their denoising (right).

The results of the gyroscopic signal denoising are analyzed quantitatively with the use of the Allan Variance.

In general, the gyroscopic signal includes quantization error (QE), angle random walk (ARW), bias instability (BI), rate random walk (RRW), and rate ramp (RR) [110].

TABLE 10. Allan Variance results of x-axis gyroscope denoising specified by five noise terms.

	QE [arcsec]	ARW [$deg/h^{1/2}$]	BI [deg/h]	RRW [$deg/h^{3/2}$]	RR [deg/h^2]
Signal	$0,0053 \cdot 10^{-3}$	$0,1754 \cdot 10^{-3}$	$0,0652 \cdot 10^{-3}$	$-0,0632 \cdot 10^{-3}$	$0,0170 \cdot 10^{-3}$
NL-means	$-0,0102 \cdot 10^{-3}$	$0,1077 \cdot 10^{-3}$	$-0,0398 \cdot 10^{-3}$	$0,0169 \cdot 10^{-3}$	$-0,0021 \cdot 10^{-3}$
WATV	$0,0081 \cdot 10^{-6}$	$-0,0641 \cdot 10^{-6}$	$0,1481 \cdot 10^{-6}$	$0,0259 \cdot 10^{-6}$	$0,2096 \cdot 10^{-6}$
EEMD-CIIT	$0,0011 \cdot 10^{-6}$	$-0,0118 \cdot 10^{-6}$	$0,0409 \cdot 10^{-6}$	$-0,0544 \cdot 10^{-6}$	$0,3259 \cdot 10^{-6}$
EEMD-TVD	$0,0002 \cdot 10^{-6}$	$-0,0022 \cdot 10^{-6}$	$0,0059 \cdot 10^{-6}$	$-0,0004 \cdot 10^{-6}$	$0,3710 \cdot 10^{-6}$
EEMD-FLS	$0,0004 \cdot 10^{-6}$	$-0,0044 \cdot 10^{-6}$	$0,0130 \cdot 10^{-6}$	$-0,0087 \cdot 10^{-6}$	$0,4173 \cdot 10^{-6}$
EEMD-NSR	$0,0012 \cdot 10^{-6}$	$-0,0124 \cdot 10^{-6}$	$0,0429 \cdot 10^{-6}$	$-0,0574 \cdot 10^{-6}$	$0,3005 \cdot 10^{-6}$

2) EXPERIMENTAL RESULTS

Fig. 8 a,b,c (on the right) show the results of removing the drift from the angular velocity (acquired from gyroscope). Gray signal is the original noisy signal. The results of signal denoising are obtained by NLM, Wavelet-TV, EEMD-CIIT, EEMD-TVD, EEMD-FLS, and EEMD-NSR.

A quantitative comparison of the results obtained by Allan Variance is pictured in Fig. 9 a,b,c and tabulated in Tables 10 – 12. From Allan Variance plots of the gyroscopic signal, we can observe three different scales $-1/2$, 0, and $+1/2$. It indicates, among others, the presence of angle random walk, bias instability, and rate random walk noises [112]. The gyroscopic signal can also include the quantization and rate ramp noises. In our analysis, we take into account all of these noise terms and their influence on the precision of gyroscope. All components of the gyroscope slope signal are determined based on the code accompanying the paper [110].

Allan Variance plots show that each of the EEMD-based methods and Wavelet-TV algorithm can remove the angle random walk noise. The non-local means algorithm is not able to attenuate this component. The observation is also verified in Tables 10 – 12 (see column ARW). It meets the remark in [8]. The NLM algorithm is designed for the white Gaussian noise, and its ability to attenuate the non-white Gaussian noise is limited (please see also section V-A).

The second reference method (Wavelet-TV) and the typical EEMD-CIIT algorithm can effectively improve the quality of the gyroscope signal. The values of QE, ARW, BI, RRW, and RR differ by three orders of magnitude (Tables 10 and 11) and two orders of magnitude (Table 12) compared with the original signal.

Comparing EEMD-CIIT with the proposed methods, we can see that, in almost all cases, the latter ones produce better results. From Tables 10 – 12, it is observed that the EEMD-TVD and EEMD-FLS predominate in removing QE, ARW, BI, and RRW, whereas for RR noise, the best results are obtained from the EEMD-NSR method.

In turn, we can observe the best results of RR attenuation we obtain for the wavelet-TV algorithm for x, y, and z axes of the gyroscope.

C. TIME COMPLEXITY

In this section, we analyze the time complexity of the proposed algorithms, that is, EEMD-TVD, EEMD-FLS, and EEMD-NSR. The obtained results are compared with the results of the similar analysis performed for EEMD-CIIT and the reference method, that is, non-local means.

In each of the proposed EEMD-based denoising algorithms, the process of signal decomposition is performed. It is the most demanding part of the proposed methods. However, in EEMD-TVD and EEMD-FLS, the decomposition is performed only once, whereas in EEMD-NSR, the process of IMFs extraction is repeated and the number of repetitions depends on the number of the executed iterations. Analyzing the algorithm (3), we can see that EEMD-NSR has higher time complexity.

It is worth to mention that in both EEMD-TVD and EEMD-FLS, an additional operation related to total variation denoising is performed. However, in these cases, we used fast implementation of TVD algorithm proposed in [90].

FIGURE 9. The Allan Variance plot of original and denoised gyroscopic signal.

TABLE 11. Allan Variance results of y-axis gyroscope denoising specified by five noise terms.

	QE [arcsec]	ARW $[deg/h^{1/2}]$	BI $[deg/h]$	RRW $[deg/h^{3/2}]$	RR $[deg/h^2]$
Signal	$-0,0249 \cdot 10^{-3}$	$0,3503 \cdot 10^{-3}$	$-0,3490 \cdot 10^{-3}$	$0,2059 \cdot 10^{-3}$	$-0,0397 \cdot 10^{-3}$
NL-means	$-0,0694 \cdot 10^{-3}$	$0,6978 \cdot 10^{-3}$	$-0,3080 \cdot 10^{-3}$	$0,1276 \cdot 10^{-3}$	$-0,0206 \cdot 10^{-3}$
WATV	$-0,0010 \cdot 10^{-6}$	$0,0378 \cdot 10^{-6}$	$-0,2917 \cdot 10^{-6}$	$0,9166 \cdot 10^{-6}$	$0,2273 \cdot 10^{-6}$
EEMD-CIIT	$0,0019 \cdot 10^{-6}$	$-0,0199 \cdot 10^{-6}$	$0,0686 \cdot 10^{-6}$	$-0,0915 \cdot 10^{-6}$	$0,5519 \cdot 10^{-6}$
EEMD-TVD	$0,0006 \cdot 10^{-6}$	$-0,0060 \cdot 10^{-6}$	$0,0198 \cdot 10^{-6}$	$-0,0222 \cdot 10^{-6}$	$0,4433 \cdot 10^{-6}$
EEMD-FLS	$0,0008 \cdot 10^{-6}$	$-0,0079 \cdot 10^{-6}$	$0,0265 \cdot 10^{-6}$	$-0,0315 \cdot 10^{-6}$	$0,4409 \cdot 10^{-6}$
EEMD-NSR	$0,0017 \cdot 10^{-6}$	$-0,0179 \cdot 10^{-6}$	$0,0617 \cdot 10^{-6}$	$-0,0824 \cdot 10^{-6}$	$0,4463 \cdot 10^{-6}$

TABLE 12. Allan Variance results of z-axis gyroscope denoising specified by five noise terms.

	QE [arcsec]	ARW $[deg/h^{1/2}]$	BI $[deg/h]$	RRW $[deg/h^{3/2}]$	RR $[deg/h^2]$
Signal	$0,0104 \cdot 10^{-3}$	$0,1029 \cdot 10^{-3}$	$0,0882 \cdot 10^{-3}$	$-0,0431 \cdot 10^{-3}$	$0,0091 \cdot 10^{-3}$
NL-means	$-0,0025 \cdot 10^{-4}$	$0,1118 \cdot 10^{-4}$	$0,0943 \cdot 10^{-4}$	$-0,0419 \cdot 10^{-4}$	$0,0250 \cdot 10^{-4}$
WATV	$0,0038 \cdot 10^{-5}$	$-0,0243 \cdot 10^{-5}$	$-0,0064 \cdot 10^{-5}$	$0,2440 \cdot 10^{-5}$	$0,1541 \cdot 10^{-5}$
EEMD-CIIT	$0,0009 \cdot 10^{-5}$	$-0,0097 \cdot 10^{-5}$	$0,0335 \cdot 10^{-5}$	$-0,0445 \cdot 10^{-5}$	$0,2255 \cdot 10^{-5}$
EEMD-TVD	$0,0007 \cdot 10^{-5}$	$-0,0074 \cdot 10^{-5}$	$0,0255 \cdot 10^{-5}$	$-0,0331 \cdot 10^{-5}$	$0,2210 \cdot 10^{-5}$
EEMD-FLS	$0,0007 \cdot 10^{-5}$	$-0,0076 \cdot 10^{-5}$	$0,0260 \cdot 10^{-5}$	$-0,0322 \cdot 10^{-5}$	$0,2323 \cdot 10^{-5}$
EEMD-NSR	$0,0008 \cdot 10^{-5}$	$-0,0088 \cdot 10^{-5}$	$0,0302 \cdot 10^{-5}$	$-0,0401 \cdot 10^{-5}$	$0,1890 \cdot 10^{-5}$

We have performed a series of experiments to verify the time complexity of the specified signal denoising methods. The test signal was the *Doppler* signal with the length ranging from 128 to 1024, where H = 0.8 and SNR = 15 [dB]. The maximal number of iterations in the EEMD algorithm was set to 50 for the tested methods. The experiments were performed with the following computer: Intel® Core™ i7 @ 2.20 [GHz] and 16.00 GB RAM memory running Windows 10. The execution time is shown in Tab. 13. The NLM algorithm has the lowest time complexity (i.e. $0,08 \pm 0,01$ seconds for the signal of length 1024) in comparison to the other tested methods.

Compared with non-local means and wavelet-TV, the remaining methods are characterized by a higher time complexity. The reference method EEMD-CIIT has a very long running time equal to $177,44 \pm 33,42$ seconds (for the signal of 1024 samples). Similarly, one of the proposed method, that is, EEMD-NSR has a very high time complexity of $241,75 \pm 40,75$ seconds for the signal of 1024 samples. It is converging to our previous remark that in EEMD-NSR, the process of IMFs extraction is repeated and increases the overall amount of time taken by the method.

In turn, unlike EEMD-CIIT and EEMD-NSR, EEMD-TVD and EEMD-FLS methods give a much shorter running time ranging between $18,10 \pm 4,50$ to $18,42 \pm 4,47$ seconds for the signal of length 1024. EEMD-TVD and EEMD-FLS are a good alternative for the EEMD-CIIT algorithm, particularly, for signals like *Blocks* and *Doppler*. A similar analysis can be performed for the signals of length 128 and 512.

TABLE 13. Relationship between the sample number and the execution time (The *Doppler* signal, fractional Gaussian Noise H = 0.8, SNR = 15 [dB]).

Methods	Time (s)	Number of samples		
		128	512	1024
NL-means	mean	0,01	0,04	0,08
	std	0,0010	0,0025	0,01
WATV	mean	0,15	0,15	0,19
	std	0,01	0,01	0,02
EEMD-CIIT	mean	67,16	118,35	177,44
	std	1,91	5,72	33,42
EEMD-TVD	mean	6,62	11,76	18,10
	std	0,35	0,49	4,50
EEMD-FLS	mean	6,67	11,81	18,42
	std	0,41	0,75	4,47
EEMD-NSR	mean	89,36	155,54	241,75
	std	3,30	6,45	40,71

The obtained results indicate that one of the future directions of our works should be focused on the reduction of the time complexity of the proposed algorithms.

D. PARAMETERS ANALYSIS

The EEMD-NSR algorithm has two parameters to be set before running the denoising algorithm. One of the parameters is η (eq. 43). In the article [34], the value of the

parameter η is suggested to be $\eta = 0.95$. Based on our numerical analysis for the *Doppler* signal (see Fig. 10) we confirm that parameter η should have the value 0.95. In Fig. 10 we can see that SNR is maximal for this value. A similar analysis can be performed for other forms of signals.
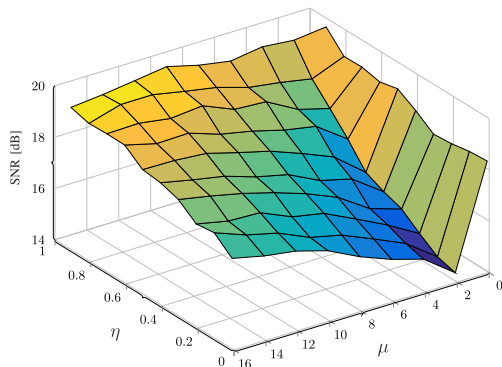

FIGURE 10. Surface plot of the SNR ratio's relationship with η and μ parameters for EEMD-NSR method (The *Doppler* signal, fractional Gaussian Noise H = 0.8, SNR = 15 [dB]).

The second parameter to be set is μ (please see eq. 34a and 34b) which is used to balance the penalization (33), and its value must be positive. Typically, the value of μ is initially low, but as the solution converges, its value is progressively increased [35].

In the work [36] a simple scheme that can be used to iteratively determine values of parameter μ is suggested.

E. PROPOSED APPROACH IN ANALYSIS OF MULTIVARIATE SIGNALS

Empirical mode decomposition and its variations are also useful to extract the components in multivariate signals [38]. Among the modern approaches to processing the multivariate signals, the low-rank approximation-based approach is one of the most popular. Its successful performance has been proven, for example, in the processing of EEG [37] and seismic signals [38].

Recently, a simultaneous cosparsity and a low-rank optimization problem [39] have shown usefulness in the processing of EEG signals [40]. The idea of cosparsity has several advantages in the processing of multivariate signals. For example, this approach allows us to estimate the signal directly, unlike a typical sparse synthesis in which firstly the sparse vector is estimated, and then, based on obtained results, the signal is determined.

In the paper [39], to solve the optimization problem of simultaneous cosparsity and low-rank, the authors recommend using the ADMM method. In this work, the ADMM algorithm has been compared with other methods: block sparse Bayesian learning, simultaneous orthogonal matching pursuit, and simultaneous greedy analysis pursuit. Conducted tests showed that ADMM outperforms the other ones in accuracy. Some of them are faster; however, their accuracy is much worse.

As proposed in this paper, to apply the approach to process multivariate signals, the considered optimization models (22, 24 and 26) have to be reformulated. In some cases for multivariate signals analysis, instead of norms ℓ_1 and ℓ_2, we should use *Schatten-p* norm [39], [40].

Based on the literature review, it is clear that the proposed approach in this work is based on empirical mode decomposition, and sparse reconstruction algorithms can be applied in multivariate signal processing. It is an interesting direction of future research.

VI. DISCUSSION

The experimental results of the fGn noise removal from synthetic signals (*Blocks*, *Bumps* and *Doppler*) indicate that the proposed approaches provided an improvement over the EEMD-CIIT method for all signals, a significant advantage we observed for *Blocks* and *Doppler* signals. This part of the experiments shows that the proposed approaches remarkably improve the denoising performance in comparison with EEMD-CIIT. It is worth stressing that the EEMD-NSR method always gives better results in comparison to EEMD-CIIT and, in almost all cases, to other proposed algorithms (EEMD-TVD and EEMD-FLS).

In the second experiment, we test the methods with the signals obtained from the gyroscope. In this case, the proposed method EEMD-NSR is always better at removing the noise from gyroscope measurements. EEMD-TVD and EEMD-FLS give better results for z axis and similar results for x and y axes in comparison to the EEMD-CIIT approach.

VII. CONCLUSION

We proposed new algorithms for signal denoising which are improving the performance of the existing EEMD-based approach. The main novelty of our methods is the application of sparse modeling in the EEMD-based signal algorithms.

The proposed approach exhibits an enhanced performance compared with the wavelet inspired algorithm EEMD-CIIT in almost all cases. The proposed algorithms outperform the EEMD-CIIT method for the synthetic signals with various SNRs and in the presence of short and long dependencies in noise. Similarly, in the case of gyroscopic signals, the proposed algorithm gives better results.

We also compare EEMD-TVD, EEMD-FLS, and EEMD-NSR algorithms to the non-local means and wavelet-TV methods. We can see that for the *Bumps* signal, the non-local means outperform the proposed algorithms in almost all cases. On the other hand, in the presence of short dependencies in noise, EEMD-NSR, EEMD-FLS, and reference wavelet-TV algorithm give better results in comparison to NLM. However, comparing this method with EEMD-NSR and EEMD-FLS for *Doppler* signal, we reported that these algorithms provide better results in all cases.

We have also shown the robustness of proposed methods by conducting a number of simulations with various signal forms and SNR.

The proposed algorithms also have some drawbacks. The main disadvantage of the methods is the high computational complexity. The computational complexity is remarkably higher compared, for example, with non-local means. The reason is that the EEMD algorithm which we used in the current implementation to signal decomposition has a high computational complexity [113]. On the one hand, the computational complexity of the proposed algorithm can be improved by applying fast implementation of the EEMD algorithm. On the other hand, it is possible to improve the running time of the proposed algorithm by replacing the ADMM algorithm with non-iterative methods. Further works will be focused, inter alia, on improving the time complexity of the proposed methods.

The major advantages of the proposed methods are summarized below:

1) remarkable improvement of denoising performance for synthetic signals (*Blocks* and *Doppler* signals);
2) high SNR in the presence of short dependencies noise in signals;
3) significant improvement of the signal acquired from the gyroscope sensor.

REFERENCES

[1] S. Chen, X. Dong, G. Xing, Z. Peng, W. Zhang, and G. Meng, "Separation of overlapped non-stationary signals by ridge path regrouping and intrinsic chirp component decomposition," *IEEE Sensors J.*, vol. 17, no. 18, pp. 5994–6005, Sep. 2017.

[2] S. W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*. San Diego, CA, USA: California Tech. Pub., 1997.

[3] G. R. Arce, *Nonlinear Signal Processing: A Statistical Approach*. Hoboken, NJ, USA: Wiley, 2005.

[4] J. Gao, H. Sultan, J. Hu, and W.-W. Tung, "Denoising nonlinear time series by adaptive filtering and wavelet shrinkage: A comparison," *IEEE Signal Process. Lett.*, vol. 17, no. 3, pp. 237–240, Mar. 2010.

[5] D. Wei and Y.-M. Li, "Generalized wavelet transform based on the convolution operator in the linear canonical transform domain," *Optik*, vol. 125, no. 16, pp. 4491–4496, Aug. 2014.

[6] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[7] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2005, pp. 60–65.

[8] B. H. Tracey and E. L. Miller, "Nonlocal means denoising of ECG signals," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2383–2386, Sep. 2012.

[9] C. Chen, J. Shi, and P. Huang, "Unsupervised discriminant analysis based on the local and non-local mean," *Phys. Procedia*, vol. 24, pp. 1967–1973, Jan. 2012.

[10] S. Kumar, D. Panigrahy, and P. K. Sahu, "Denoising of electrocardiogram (ECG) signal by using empirical mode decomposition (EMD) with non-local mean (NLM) technique," *Biocybern. Biomed. Eng.*, vol. 38, no. 2, pp. 297–312, 2018.

[11] S. K. Yadav, P. K. Bora, and R. Sinha, "Electrocardiogram signal denoising using non-local wavelet transform domain filtering," *IET Signal Process.*, vol. 9, no. 1, pp. 88–96, Feb. 2015.

[12] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London. Ser., Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[13] B. Weng, M. Blanco-Velasco, and K. E. Barner, "ECG denoising based on the empirical mode decomposition," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2006, pp. 1–4.

[14] M. Blanco-Velasco, B. Weng, and K. E. Barner, "ECG signal denoising and baseline wander correction based on the empirical mode decomposition," *Comput. Biol. Med.*, vol. 38, no. 1, pp. 1–13, Jan. 2008.

[15] A. Komaty, A.-O. Boudraa, B. Augier, and D. Dare-Emzivat, "EMD-based filtering using similarity measure between probability density functions of IMFs," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 1, pp. 27–34, Jan. 2014.

[16] K. Khaldi, M. Turki-Hadj Alouane, and A.-O. Boudraa, "A new EMD denoising approach dedicated to voiced speech signals," in *Proc. 2nd Int. Conf. Signals, Circuits Syst.*, Nov. 2008, pp. 1–5.

[17] N. Chatlani and J. J. Soraghan, "EMD-based filtering (EMDF) of low-frequency noise for speech enhancement," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 4, pp. 1158–1166, May 2012.

[18] K. Khaldi, A.-O. Boudraa, A. Bouchikhi, and M. T.-H. Alouane, "Speech enhancement via EMD," *EURASIP J. Adv. Signal Process.*, vol. 2008, no. 1, Dec. 2008, Art. no. 873204.

[19] X. Guo, C. Sun, P. Wang, and L. Huang, "Hybrid methods for MEMS gyro signal noise reduction with fast convergence rate and small steady-state error," *Sens. Actuators A, Phys.*, vol. 269, pp. 145–159, Jan. 2018.

[20] G. Yang, Y. Liu, Y. Wang, and Z. Zhu, "EMD interval thresholding denoising based on similarity measure to select relevant modes," *Signal Process.*, vol. 109, pp. 95–109, Apr. 2015.

[21] B. Cui and X. Chen, "Improved hybrid filter for fiber optic gyroscope signal denoising based on EMD and forward linear prediction," *Sens. Actuators A, Phys.*, vol. 230, pp. 150–155, Jul. 2015.

[22] N. Li and P. Li, "An improved algorithm based on EMD-wavelet for ECG signal de-noising," in *Proc. IEEE Int. Joint Conf. Comput. Sci. Optim.*, Apr. 2009, pp. 825–827.

[23] M. A. Kabir and C. Shahnaz, "Denoising of ECG signals based on noise reduction algorithms in EMD and wavelet domains," *Biomed. Signal Process. Control*, vol. 7, no. 5, pp. 481–489, Sep. 2012.

[24] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.

[25] S. Lahmiri and M. Boukadoum, "Biomedical image denoising using variational mode decomposition," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Oct. 2014, pp. 340–343.

[26] W. Liu, S. Y. Cao, and Y. He, "Ground roll attenuation using variational mode decomposition," in *Proc. 77th EAGE Conf. Exhib.*, Jun. 2015, pp. 1–5.

[27] S. Lahmiri, "Comparative study of ECG signal denoising by wavelet thresholding in empirical and variational mode decomposition domains," *Healthcare Technol. Lett.*, vol. 1, no. 3, pp. 104–109, Sep. 2014.

[28] Y. Liu, G. Yang, M. Li, and H. Yin, "Variational mode decomposition denoising combined the detrended fluctuation analysis," *Signal Process.*, vol. 125, pp. 349–364, Aug. 2016.

[29] Y. Kopsinis and S. McLaughlin, "Development of EMD-based denoising methods inspired by wavelet thresholding," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1351–1362, Apr. 2009.

[30] A. Gholami and S. M. Hosseini, "A general framework for sparsity-based denoising and inversion," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5202–5211, Nov. 2011.

[31] U. Kamilov, E. Bostan, and M. Unser, "Wavelet shrinkage with consistent cycle spinning generalizes total variation denoising," *IEEE Signal Process. Lett.*, vol. 19, no. 4, pp. 187–190, Apr. 2012.

[32] I. W. Selesnick, A. Parekh, and I. Bayram, "Convex 1-D total variation denoising with non-convex regularization," *IEEE Signal Process. Lett.*, vol. 22, no. 2, pp. 141–144, Feb. 2015.

[33] I. W. Selesnick, H. L. Graber, D. S. Pfeil, and R. L. Barbour, "Simultaneous low-pass filtering and total variation denoising," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1109–1124, Mar. 2014.

[34] Y. Ding and I. W. Selesnick, "Artifact-free wavelet denoising: Non-convex sparse regularization, convex optimization," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1364–1368, Sep. 2015.

[35] A. Majumdar, *Compressed Sensing for Engineers*. Boca Raton, FL, USA: CRC Press, 2018.

[36] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2010.

[37] A. Majumdar, A. Gogna, and R. Ward, "A low-rank matrix recovery approach for energy efficient EEG acquisition for a wireless body area network," *Sensors*, vol. 14, no. 9, pp. 15729–15748, Aug. 2014.

[38] Y. Chen, Y. Zhou, W. Chen, S. Zu, W. Huang, and D. Zhang, "Empirical low-rank approximation for seismic noise attenuation," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4696–4711, Aug. 2017.

[39] Y. Liu, M. De Vos, and S. Van Huffel, "Compressed sensing of multi-channel EEG signals: The simultaneous cosparsity and low-rank optimization," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 8, pp. 2055–2061, Aug. 2015.

[40] J. Zhu, C. Chen, S. Su, and Z. Chang, "Compressive sensing of multi-channel EEG signals via lq norm and Schatten-p norm regularization," *Math. Problems Eng.*, vol. 2016, pp. 1–7, Nov. 2016.

[41] Q. Wang, X. Zhang, Y. Wu, L. Tang, and Z. Zha, "Nonconvex weighted ℓ_p minimization based group sparse representation framework for image denoising," *IEEE Signal Process. Lett.*, vol. 24, no. 11, pp. 1686–1690, Nov. 2017.

[42] X. Wang, Y. Zhou, M. Shu, Y. Wang, and A. Dong, "ECG baseline wander correction and denoising based on sparsity," *IEEE Access*, vol. 7, pp. 31573–31585, 2019.

[43] X. Ning, I. W. Selesnick, and L. Duval, "Chromatogram baseline estimation and denoising using sparsity (BEADS)," *Chemometric Intell. Lab. Syst.*, vol. 139, pp. 156–167, Dec. 2014.

[44] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, and K. Knight, "Sparsity and smoothness via the fused LASSO," *J. Roy. Stat. Soc., Ser. B, Stat. Methodol.*, vol. 67, no. 1, pp. 91–108, 2005.

[45] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc., Ser. B, Stat. Methodol.*, vol. 58, no. 1, pp. 267–288, Jan. 1996.

[46] I. Bayram, P.-Y. Chen, and I. W. Selesnick, "Fused lasso with a non-convex sparsity inducing penalty," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2014, pp. 4156–4160.

[47] T. Y. Hou and Z. Shi, "Adaptive data analysis via sparse time-frequency representation," *Adv. Adapt. Data Anal.*, vol. 3, no. 1, pp. 1–28, Apr. 2011.

[48] T. Y. Hou and Z. Shi, "Sparse time-frequency decomposition based on dictionary adaptation," *Phil. Trans. Roy. Soc., Math., Phys. Eng. Sci.*, vol. 374, no. 2065, pp. 1–16, 2016.

[49] Z. Zhang, J. Ren, W. Jiang, Z. Zhang, R. Hong, S. Yan, and M. Wang, "Joint subspace recovery and enhanced locality driven robust flexible discriminative dictionary learning," *IEEE Trans. Circuits Syst. Video Technol.*, early access, Jun. 14, 2019, doi: 10.1109/TCSVT.2019.2923007.

[50] Z. Zhang, W. Jiang, J. Qin, L. Zhang, F. Li, M. Zhang, and S. Yan, "Jointly learning structured analysis discriminative dictionary and analysis multi-class classifier," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 8, pp. 3798–3814, Aug. 2018.

[51] Z. Zhang, W. Jiang, Z. Zhang, S. Li, G. Liu, and J. Qin, "Scalable block-diagonal locality-constrained projective dictionary learning," 2019, *arXiv:1905.10568*. [Online]. Available: http://arxiv.org/abs/1905.10568

[52] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proc. IEEE*, vol. 98, no. 6, pp. 1045–1057, Jun. 2010.

[53] R. Chen, H. Jia, X. Xie, and W. Gao, "Learning a collaborative multiscale dictionary based on robust empirical mode decomposition," *Neurocomputing*, vol. 287, pp. 196–207, Apr. 2018.

[54] Z. Yang, B. W.-K. Ling, and C. Bingham, "Joint empirical mode decomposition and sparse binary programming for underlying trend extraction," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 10, pp. 2673–2682, Oct. 2013.

[55] J. Chang, L. Zhu, H. Li, F. Xu, B. Liu, and Z. Yang, "Noise reduction in lidar signal using correlation-based EMD combined with soft thresholding and roughness penalty," *Opt. Commun.*, vol. 407, pp. 290–295, Jan. 2018.

[56] T. Jin, Q. Li, and M. A. Mohamed, "A novel adaptive EEMD method for switchgear partial discharge signal denoising," *IEEE Access*, vol. 7, pp. 58139–58147, 2019.

[57] X. Chen, X. Xu, A. Liu, M. J. McKeown, and Z. J. Wang, "The use of multivariate EMD and CCA for denoising muscle artifacts from few-channel EEG recordings," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 2, pp. 359–370, Feb. 2018.

[58] X. Lang, Q. Zheng, Z. Zhang, S. Lu, L. Xie, A. Horch, and H. Su, "Fast multivariate empirical mode decomposition," *IEEE Access*, vol. 6, pp. 65521–65538, 2018.

[59] Y. Chen and J. Ma, "Random noise attenuation by $f − x$ empirical-mode decomposition predictive filtering," *Geophysics*, vol. 79, no. 3, pp. V81–V91, May 2014.

[60] Y. Chen, C. Zhou, J. Yuan, and Z. Jin, "Applications of empirical mode decomposition in random noise attenuation of seismic data," *J. Seismic Explor.*, vol. 23, no. 5, pp. 481–495, 2014.

[61] Y. Chen, G. Zhang, S. Gan, and C. Zhang, "Enhancing seismic reflections using empirical mode decomposition in the flattened domain," *J. Appl. Geophysics*, vol. 119, pp. 99–105, Aug. 2015.

[62] Y. Chen, "Dip-separated structural filtering using seislet transform and adaptive empirical mode decomposition based dip filter," *Geophys. J. Int.*, vol. 206, no. 1, pp. 457–469, Jul. 2016.

[63] W. Chen, Y. Chen and W. Liu, "Ground roll attenuation using improved complete ensemble empirical mode decomposition," *J. Seismic Explor.*, vol. 25, no. 5, pp. 485–495, 2016.

[64] W. Chen, J. Xie, S. Zu, S. Gan, and Y. Chen, "Multiple-reflection noise attenuation using adaptive randomized-order empirical mode decomposition," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 1, pp. 18–22, Jan. 2017.

[65] M. Kaleem, A. Guergachi, and S. Krishnan, "Empirical mode decomposition based sparse dictionary learning with application to signal classification," in *Proc. IEEE Digit. Signal Process. Signal Process. Edu. Meeting (DSP/SPE)*, Aug. 2013, pp. 18–23.

[66] P. Flandrin, G. Rilling, and P. Goncalves, "Empirical mode decomposition as a filter bank," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 112–114, Feb. 2004.

[67] T.-L. Huang, M.-L. Lou, H.-P. Chen, and N.-B. Wang, "An orthogonal Hilbert-huang transform and its application in the spectral representation of earthquake accelerograms," *Soil Dyn. Earthq. Eng.*, vol. 104, pp. 378–389, Jan. 2018.

[68] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, Jan. 2009.

[69] D. P. Mandic, N. U. Rehman, Z. Wu, and N. E. Huang, "Empirical mode decomposition-based time-frequency analysis of multivariate signals: The power of adaptive data analysis," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 74–86, Nov. 2013.

[70] J.-R. Yeh, J.-S. Shieh, and N. E. Huang, "Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method," *Adv. Adapt. Data Anal.*, vol. 2, no. 2, pp. 135–156, Apr. 2010.

[71] J. Zheng, J. Cheng, and Y. Yang, "Partly ensemble empirical mode decomposition: An improved noise-assisted method for eliminating mode mixing," *Signal Process.*, vol. 96, pp. 362–374, Mar. 2014.

[72] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 4144–4147.

[73] Z. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proc. Roy. Soc. London. Ser., Math., Phys. Eng. Sci.*, vol. 460, no. 2046, pp. 1597–1611, Jun. 2004.

[74] A. Moghtaderi, P. Flandrin, and P. Borgnat, "Trend filtering via empirical mode decompositions," *Comput. Statist. Data Anal.*, vol. 58, pp. 114–126, Feb. 2013.

[75] A. Mert and A. Akan, "Detrended fluctuation thresholding for empirical mode decomposition based denoising," *Digit. Signal Process.*, vol. 32, pp. 48–56, Sep. 2014.

[76] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, Sep. 1994.

[77] Z. Jin, A. Dong, M. Shu, and Y. Wang, "Sparse ECG denoising with generalized minimax concave penalty," *Sensors*, vol. 19, no. 7, p. 1718, Apr. 2019.

[78] P.-Y. Chen and I. W. Selesnick, "Group-sparse signal denoising: Nonconvex regularization, convex optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 13, pp. 3464–3478, Jul. 2014.

[79] L. Yang, C. Li, J. Han, C. Chen, Q. Ye, B. Zhang, X. Cao, and W. Liu, "Image reconstruction via manifold constrained convolutional sparse coding for image sets," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 7, pp. 1072–1081, Oct. 2017.

[80] M. Gong, X. Jiang, and H. Li, "Optimization methods for regularization-based ill-posed problems: A survey and a multi-objective framework," *Frontiers Comput. Sci.*, vol. 11, no. 3, pp. 362–391, Jun. 2017.

[81] A. N. Tikhonov and V. I. Arsenin, *Solutions ill-Posed Problems*. Washington, DC, USA: Winston, 1977.

[82] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[83] I. Rish and G. Grabarnik, *Sparse Modeling: Theory, Algorithms, and Applications*. Boca Raton, FL, USA: CRC Press, 2014.

[84] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, nos. 5–6, pp. 877–905, 2008.

[85] Y. Lin and J. Cai, "A new threshold function for signal denoising based on wavelet transform," in *Proc. Int. Conf. Measuring Technol. Mechatronics Autom.*, Mar. 2010, pp. 200–203.

[86] S.-W. Deng and J.-Q. Han, "Efficient general sparse denoising with non-convex sparse constraint and total variation regularization," *Digit. Signal Process.*, vol. 78, pp. 259–264, Jul. 2018.

[87] M. A. Little and N. S. Jones, "Generalized methods and solvers for noise removal from piecewise constant signals. I. Background theory," *Proc. Roy. Soc., Math., Phys. Eng. Sci.*, vol. 467, no. 2135, pp. 3088–3114, Nov. 2011.

[88] M. Zhu, S. J. Wright, and T. F. Chan, "Duality-based algorithms for total-variation-regularized image restoration," *Comput. Optim. Appl.*, vol. 47, no. 3, pp. 377–400, Nov. 2010.

[89] T. Hastie, R. Tibshirani, and M. Wainwright, *Stat. Learn. with sparsity: LASSO and generalizations*. Boca Raton, FL, USA: CRC Press, 2015.

[90] L. Condat, "A direct algorithm for 1-D total variation denoising," *IEEE Signal Process. Lett.*, vol. 20, no. 11, pp. 1054–1057, Nov. 2013.

[91] J. Qian and J. Jia, "On stepwise pattern recovery of the fused lasso," *Comput. Statist. Data Anal.*, vol. 94, pp. 221–237, Feb. 2016.

[92] J. Friedman, T. Hastie, H. Höfling, and R. Tibshirani, "Pathwise coordinate optimization," *Ann. Appl. Statist.*, vol. 1, no. 2, pp. 302–332, Dec. 2007.

[93] I. W. Selesnick, H. L. Graber, Y. Ding, T. Zhang, and R. L. Barbour, "Transient artifact reduction algorithm (TARA) based on sparse optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6596–6611, Dec. 2014.

[94] J. Fan and R. Li, "Variable selection via nonconcave penalized likelihood and its oracle properties," *J. Amer. Stat. Assoc.*, vol. 96, no. 456, pp. 1348–1360, Dec. 2001.

[95] A. Mehranian, H. S. Rad, A. Rahmim, M. R. Ay, and H. Zaidi, "Smoothly clipped absolute deviation (SCAD) regularization for compressed sensing MRI using an augmented lagrangian scheme," *Magn. Reson. Imag.*, vol. 31, no. 8, pp. 1399–1411, Oct. 2013.

[96] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345–2356, Sep. 2010.

[97] J. Eckstein and D. P. Bertsekas, "On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Program.*, vol. 55, nos. 1–3, pp. 293–318, Apr. 1992.

[98] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, Mar. 2011.

[99] J. Nocedal and S. J. Wright, *Sequential Quadratic Programming*. New York, NY, USA: Springer, 2006.

[100] P. Flandrin, P. Gonçalves, and G. Rilling, "EMD equivalent filter banks, from interpretation to applications," in *Hilbert–Huang Transform and its Applications*. Singapore: World Scientific, 2014, pp. 99–116.

[101] L. Dümbgen and A. Kovac, "Extensions of smoothing via taut strings," *Electron. J. Statist.*, vol. 3, pp. 41–75, 2009. [Online]. Available: https://projecteuclid.org/download/pdfview_1/euclid.ejs/1233176790, doi: 10.1214/08-EJS216.

[102] *Non-Local Means (NLM) Denoising for Time Series, Applied to ECG*. Accessed: Sep. 5, 2019. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/41762-non-local-means–nlm–denoising-for-time-series–applied-to-ecg

[103] M. Van, H.-J. Kang, and K.-S. Shin, "Rolling element bearing fault diagnosis based on non-local means de-noising and empirical mode decomposition," *IET Sci., Meas. Technol.*, vol. 8, no. 6, pp. 571–578, Nov. 2014.

[104] *Wavelab 850*. Accessed: Sep. 5, 20019. [Online]. Available: http://statweb.stanford.edu/~wavelab/

[105] S. Bregni and L. Jmoda, "Accurate estimation of the hurst parameter of long-range dependent traffic using modified allan and Hadamard variances," *IEEE Trans. Commun.*, vol. 56, no. 11, pp. 1900–1906, Nov. 2008.

[106] Q. Huaming and M. Jichen, "Research on fiber optic gyro signal denoising based on wavelet packet soft-threshold," *J. Syst. Eng. Electron.*, vol. 20, no. 3, pp. 607–612, 2009.

[107] A. El-Rabbany and M. El-Diasty, "An efficient neural network model for de-noising of MEMS-based inertial data," *J. Navigat.*, vol. 57, no. 3, pp. 407–415, Sep. 2004.

[108] M. Narasimhappa, S. L. Sabat, and J. Nayak, "Fiber-optic gyroscope signal denoising using an adaptive robust Kalman filter," *IEEE Sensors J.*, vol. 16, no. 10, pp. 3711–3718, May 2016.

[109] L. Qian, G. Xu, W. Tian, and J. Wang, "A novel hybrid EMD-based drift denoising method for a dynamically tuned gyroscope (DTG)," *Measurement*, vol. 42, no. 6, pp. 927–932, Jul. 2009.

[110] K. Jerath, S. Brennan, and C. Lagoa, "Bridging the gap between sensor noise modeling and sensor characterization," *Measurement*, vol. 116, pp. 350–366, Feb. 2018.

[111] C. Shen, H. Cao, J. Li, J. Tang, X. Zhang, Y. Shi, W. Yang, and J. Liu, "Hybrid de-noising approach for fiber optic gyroscopes combining improved empirical mode decomposition and forward linear prediction algorithms," *Rev. Sci. strum.*, vol. 87, no. 3, Mar. 2016, Art. no. 033305.

[112] N. El-Sheimy, H. Hou, and X. Niu, "Analysis and modeling of inertial sensors using allan variance," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 1, pp. 140–149, Jan. 2008.

[113] Y.-H. Wang, C.-H. Yeh, H.-W.-V. Young, K. Hu, and M.-T. Lo, "On the computational complexity of the empirical mode decomposition algorithm," *Phys. A, Stat. Mech. Appl.*, vol. 400, pp. 159–167, Apr. 2014.

KRZYSZTOF BRZOSTOWSKI received the M.Sc. degree in electronics from the Wrocław University of Science and Technology (WUST), in 2003, and the Ph.D. degree from the Faculty of Computer Science and Management, WUST, in 2009. He joined the Faculty of Computer Science and Management, WUST, where he is currently an Assistant Professor. His current research interests include the areas of nonlinear signal processing, time-frequency signal representation, sparsity techniques, system identification, multi-sensor fusion, biomedical signal processing, and inertial navigation systems.

• • •