JOT: a Variational Signal Decomposition into Jump, Oscillation and Trend

Antonio Cicone and Martin Huska and Sung-Ha Kang and Serena Morigi

***Abstract*—We propose a two stages signal decomposition method which efficiently separates a given signal into Jump, Oscillation and Trend. While there have been numerous advances in signal processing in past few decades, they mainly aim to analyze the signal in terms of oscillating (underlying frequencies) or non-oscillating (underlying trend) features. Both traditional Time-Frequency analysis methods, like Short Time Fourier Transform, wavelet, and advanced ones, like Synchrosqueezing wavelet, Hilbert Huang Transform or IMFogram, can fail when abrupt changes and jump discontinuities appear in the signal.**

We present a variational framework separating piece-wise constant jump features as well as smooth trends and oscillating features of a given signal. In the first stage, a three component signal decomposition is applied, using sparsity promoting regularization, and Sobolev spaces of negative differentiability to model oscillations. In the second stage, components are refined using residuals of other components. The proposed method finds big and small jumps, is stable against high level of noise, is independent from the choice of basis functions, and does not have different level of decompositions which can be affected by large discontinuities. This variational framework is free from training in network-based approaches, and can be used for generating training data. The optimization problem is efficiently solved by an alternating minimization strategy. Applied as pre-processing for time-frequency analysis and Synchrosqueezing, it allows for improvements in results showing much clearer separation without artifacts. The proposed method is tested against synthetic data, where the ground truth is known, and real world data.

I. Introduction

The goal of signal decomposition is extraction and separation of signal components from composite signals, which should preferably be related to semantic features. The underlying features which characterize the components to be separated are typically defined by criteria that indicate (mathematical) homogeneity. Examples for these are a smooth trend, high oscillating components and piece-wise constant behaviour.

Over the past decades, a wide variety of signal processing methods have been proposed to perform a time-frequency analysis of a given signal and to extract its features. Among them we mention here the most popular methods, like short-time Fourier transform (STFT), wavelet transform, and recent ones, like the Synchrosqueezing transform (SST) based on continuous wavelet transform (CWT) [1], Hilbert-Huang Transform (HHT) based on Empirical Mode Decomposition (EMD) [2], Fourier–Singh analytic signal (FSAS) representations based on Fourier quadrature transforms (FQTs) [3], spectral kurtosis (SK) [4], or IMFogram based on Fast Iterative Filtering (FIF) [5].

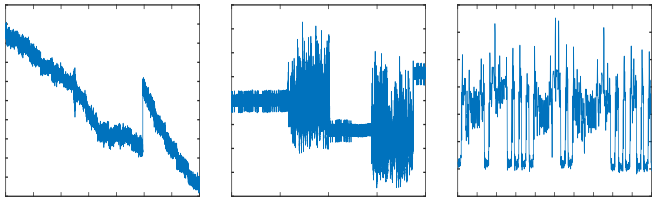
These conventional methods suffer in dealing with interference elements, such as noise which exists in the whole frequency range, the presence of oscillatory and non-oscillatory behaviors, and anomalies such as impulse features like jumps, spikes and sudden drops. This kind of anomalies affects the detection of the underlying trends in time series data or signals, since these abrupt changes corrupt the mean level of a signal. Often, the jump step is small and the time series is damaged by some type of noise, and this makes the problem even more challenging because the jump step may be hidden by the noise. In [6] the authors show how anomalies in signals can badly effect a decomposition produced using EMD– or FIF–based algorithms. In particular, they highlight how anomalies, that are commonly present in a real life signal, can have negative influence in a decomposition and in the derived time-frequency representations if not properly handled. This is true, for low frequency components, even far away from the locations of the anomalies. Such jumps, spikes and sudden drops can impact the physical meaningfulness of a signal decomposition and the associated time-frequency representation produced with all the aforementioned techniques.

In frequency domain, a jump can be present in all frequencies, which become active by the time localization of the jump nearby [6]. This is the reason why any filtering approach, which relies on a certain cutoff frequency to separate jumps from the meaningful signal, has problems in removing these artifacts in a neat way.

There is the need for a new accurate method able to identify and remove these anomalies, and possibly separate data trend from noise if present in the measurements, such that the state of the art time-frequency analysis methods can produce crisper plots which are less prone to artifacts.

Figure 1 shows some examples of real life signals containing spikes and jumps which impose challenges to classical and modern techniques for nonstationary signal processing. Figure 1(a) is a sample of the Earth's Electric Field measured using the CSES-01 satellite [7]. This signal shows small discontinuities which are created by measurement interruptions in the period of data acquisition; Figure 1(b) is an example of clean Electrocardiogram (ECG) data perturbed due to simulated electrode motion artifacts on certain segments of

DISIM, University of L'Aquila, L'Aquila, Italy
Istituto di Astrofisica e Planetologia Spaziali, INAF, Roma, Italy
Istituto Nazionale di Geofisica e Vulcanologia, Roma, Italy (antonio.cicone@univaq.it)
Department of Mathematics, University of Bologna, Bologna, Italy (martin.huska@unibo.it)
School of Mathematics, Georgia Institute of Technology, Atlanta, GA, USA (kang@math.gatech.edu)
Department of Mathematics, University of Bologna, Bologna, Italy (serena.morigi@unibo.it)



(a) (b) (c)

Fig. 1. Anomalies present in signals: (a) Earth's Magnetic Field, (b) ECG Data with electrode motion artifact, (c) CPU Magnetic Field.

the data set [8]; Figure 1(c) is an example of CPU emanated electromagnetic signal, where discontinuities are introduced by some operations performed by the device [9].

We propose an innovative approach to address the problem of separating signals like those shown in Figure 1, where the underlying data comprises jump-sparse discontinuities (anomalies), a smoothly varying component (underlying trend), and a highly oscillating component (noise/features). The main idea is to model a given finite-length discrete-time signal $f \in \mathbb{R}^N$ as

$$f = v^* + w^* + n^* \ , \tag{1}$$

where $v^* \in \mathbb{R}^N$ represents a piecewise-constant (sparse-derivative) component, $w^* \in \mathbb{R}^N$ is a smooth low-oscillating function, and $n^* \in \mathbb{R}^N$ contains zero-mean high-oscillations of the signal, which can include additive white noise, as well as a meaningful oscillatory components.

In Figure 2, we illustrate two examples of decomposition of an original signal f (on the left) and their extracted components v^*, w^*, n^*, respectively (on the right) separated using the proposed method. In these two examples v^* represents the abrupt changes occurring in the signal. The goal for the first row example is to recover the jumps v^* and the trend w^* from a noisy signal f_1, while for the second row signal f_2 we want to eliminate the jumps v^* from the measured signal f_2 to obtain jump-free high-frequency data with a low frequency trend $n^* + w^*$. The proposed method gives such flexibility. Figure 3 shows a time-frequency plot of the raw data f_2 illustrated in Figure 2 obtained by the continuous wavelet transform (CWT). The results are affected by artifacts coming from the sudden steps which mask the frequency patterns of interest, see Figure 3 (left). Whereas, if we remove the steps first, the new time-frequency plot becomes clearer, Figure 3 (right).

We propose a two-stage variational approach. The first stage separates the given signal f according to the decomposition model (1), then the second stage applies a refined restoration, using the residuals from other components.

The main contributions of this paper are

1) the development of a stable separation method for jump discontinuity, underlying trend and oscillatory components estimation, based on a three-component variational approach, which does not require any training set;
2) efficient decomposition algorithm which works well for real data without any a priori knowledge.
3) automatic recovery of jump locations without any a prior knowledge of the type or number of the jumps that occur on the signal.

We refer to the proposed method as Jump, Oscillation and Trend (JOT) decomposition framework.

Following the literature review in this section, in Section II, we present the details of the proposed JOT decomposition framework. Section III contains numerical implementations details. Numerical results are reported in Section IV, and, in Section V, concluding remarks are presented.

A. Related Work

The problem of estimating underlying trends in time series data or signals arises in a variety of disciplines including macroeconomics, geophysics, financial time series analysis, social sciences, biological and medical science. Many trend filtering methods have been proposed, and most of them are linear filtering; see [10] for a survey of linear filtering methods in trend estimation. The most widely used methods are bandpass filtering [10], moving average filtering, exponential smoothing, and Hodrick–Prescott (H-P) filtering, [11]. In case the signals possess discontinuities, classical first-order Mumford–Shah (MS) models take this explicitly into account by incorporating the jump sets. In [12], the higher order MS model introduced is capable, for given order of polynomial, to simultaneously estimate a discontinuity set, i.e. a domain partitioning, and a corresponding piecewise smooth signal.

Recently, sparsity-enhancing representation has attracted a great deal of attention in the field of feature extraction, and extensive applications have been considered in signal and image processing. Signal decomposition methods via sparse optimization have achieved good performance exploiting ℓ_1 norm-based regularizers [13]. In [14] the authors proposed the ℓ_1 trend filtering, a variation on H-P filtering which substitutes a sum of absolute values (i.e., an ℓ_1 norm) for the ℓ_2 norm used in H-P filtering to penalize variations in the estimated trend. The convex ℓ_1 norm may reduce the amplitudes of the signal, thus affecting the accuracy of the decomposition. To overcome this problem, non-convex sparse regularizers have been used. The non-convex ℓ_0 pseudonorm-based penalty is the main ingredient for the recovery of jump-sparse and sparse signals from noisy data, as used in the proposal of the inverse Potts energy functionals [15] for signal decomposition. This approach reveals great ability to well recover pure piecewise-constant signals. A wide class of non-convex penalties have been successfully proposed following the convex non-convex strategy via variational optimization. Their use in sparse optimization allows to significantly improve signal denoising [16] and decomposition [17]. The minimax convex penalty considered in our sparse optimization falls in this non-convex sparsity-inducing penalty class [18].

For the scientific signal processing community, the study of techniques for the time-frequency analysis and decomposition of signals is a long lasting line of research which has led over the decades to the development of many important algorithms and approaches, which are nowadays commonly used in many

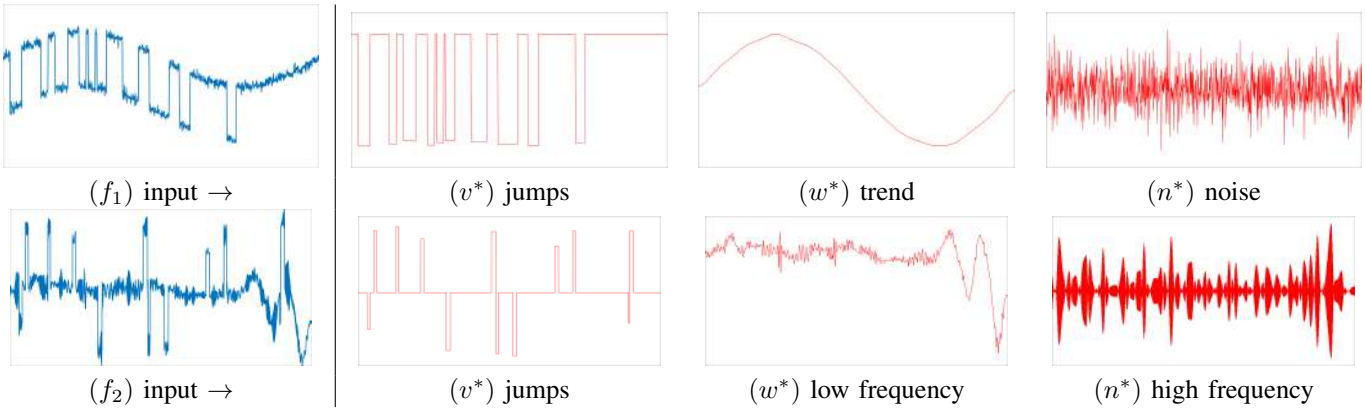


Fig. 2. Decomposition results of the proposed method: (first column) the given signals f_1 and f_2; (first row) the jumps v^*, the trend w^* and noise n^* are separated from the given signal f_1; (second row) eliminate the steps v^* from f_2 to obtain jump-free high-frequency data with a low frequency signal $w^* + n^*$.

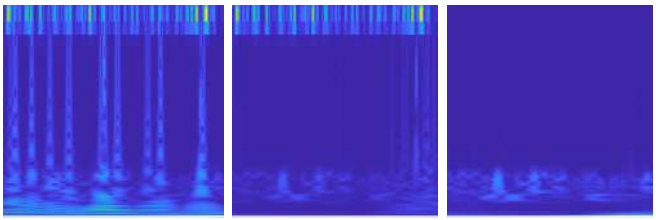


Fig. 3. (left) CWT time-frequency plot of the raw signal f_2 shown in Figure 2; (middle) CWT of the f_2 signal after Potts model jump removal; (right) CWT after removing the jumps and high frequency noise using the proposed work.

research fields [19]. When the signal under analysis is non-stationary, standard methods, like (Short Time) Fourier Transform and Wavelet Transform, were proved to be inadequate to provide detailed time-frequency information, due to their inherent linearity. Two decades ago Huang and collaborators introduced a game changer method called Hilbert Huang Transform (HHT) [2]. This is an iterative, local and adaptive data-driven method based on a "divide et impera" approach. The key idea is simple, but powerful: the signal f is first subdivided into several simple components, called Intrinsic Mode Functions (IMFs), plus a trend via what is called the Empirical Mode Decomposition (EMD) method; then, each IMF is analyzed separately in the time-frequency domain via the computation of each component instantaneous frequency obtained using the Hilbert Transform. In [2] the authors define informally the IMFs as oscillatory functions that fulfill two properties: the number of zero crossing equals the number of its extrema, plus or minus one; the envelops connecting its maxima and minima have to be symmetric with respect to the horizontal axis. The decompositions produced using the EMD algorithm proved to be successful for a wide range of applications, see for instance [6] and references therein. An alternative iterative method for a signal decomposition, is the so called Iterative Filtering (IF) algorithm [20], which, as for the EMD, does not require any "a priori" assumption on the signal under analysis. The structure of IF resemble the EMD, yet, the key difference is in the signal moving average computation, which is obtained in the IF method as the convolution of the signal f with an a priori chosen filter function. This apparently small difference between the IF and EMD moving average computation opened the doors to the mathematical analysis of IF [21]. The Fast Iterative Filtering (FIF) method allows to speed up the IF computations via FFT [21].

Regarding time-frequency analysis of the IMFs, besides the Hilbert Transform, recently it has been published a new representation method called IMFogram, which allows to produce crisp and detailed time-frequency plots of nonstationary signals [5] and converges, in the limit, to the spectrogram based on the STFT [22].

We propose a nonconvex variational model for signal decomposition where a sparsity-inducing regularizer is introduced to allow for effectively extracting the impulse features from harmonic and highly oscillating parts of the original signal. This approach is motivated by image decomposition stated from the seminal work of Y. Meyer [23], who proposed separating the image into geometric part, e.g. using TV denoising [24], and oscillatory texture or noisy part. Various work has followed [25], [26], [27], [28] just to mention a few. In particular, in [29], the authors proposed a model which decomposes the given image into a piecewise-constant part, a harmonic part and a noisy part. This is based on the work in [30], which is proposed for the decomposition of functions on triangulated surfaces. Compared to other decomposition models, the model proposed in [29] captures the structure part more clearly with the sparsity enforcing regularization term, which is well suited for jump separation in the signal processing.

II. The Proposed JOT Signal Decomposition Framework

The Jump, Oscillation and Trend (JOT) decomposition framework consists of two stages: (i) three component signal decomposition using sparsity promoting regularization and Sobolev spaces of negative differentiability to model oscillations, and (ii) component refinement using residuals. Fig. 5 presents an outline of the proposed method.

A. Stage 1: Three-Component Signal Decomposition

The first step of our sparsity-enhanced signal decomposition is performed via the following non-convex minimization problem

$$\{\bar{v},\bar{w},\bar{n}\} \leftarrow \arg\min_{v,w,n\in\mathbb{R}^N} \mathcal{J}(v,w,n)\,, \tag{2}$$
$$\mathcal{J}(v,w,n) := \frac{1}{2}\,\|v+w+n-f\|_2^2$$
$$+\bar{\gamma_1}\sum_{j=1}^{N}\phi\left(\|(Dv)_j\|;a\right)+\bar{\gamma_2}\,\|Hw\|_2^2+\,\bar{\gamma_3}\|n\|_{\mathcal{H}^{-1}}^4$$

where the scalar value $\bar{\gamma_i}$ represent the regularization parameters balancing the regularization terms and the fidelity term. The function $\phi(\cdot;a) : [0,+\infty) \to [0,1]$ is a non-convex parametrized sparsity promoting penalty function, the operators D and H represent first and second-order derivatives, respectively. We denote by $\|\cdot\|_{\mathcal{H}^{-1}}$ the norm of the negative Sobolev space $\mathcal{H}^{-1}$ which is proved to be bounded for oscillatory functions [28], being therefore useful for modeling the highly oscillating component of a signal. The minimax-concave (MC) penalty, introduced in [18], and used in the proposed signal decomposition model (2), is a piecewise-quadratic function defined by

$$\phi(t;a) \;=\; \begin{cases} -\dfrac{a}{2}\,t^2+\sqrt{2a}\,t & \text{for } t\in\left[\,0,\sqrt{2/a}\,\right), \\ 1 & \text{for } t\in\left[\,\sqrt{2/a},+\infty\,\right). \end{cases} \tag{3}$$

The parameter a in $\phi(\cdot;a)$ affects the degree of non-convexity, such that $\phi(\cdot;a)$ tends to ℓ_0 pseudonorm for $a \to \infty$. For $a = 0$, the MC penalty is defined as $\phi(t;a) = |t|$. Unlike the classical ℓ_1 norm regularizer the proposed MC penalty can accurately preserve the amplitude of the piecewise-constant signal component v while better representing sharp discontinuities, thus to improve the decomposition accuracy, [29], [31], [16]. The smooth varying component (underlying trend) w is obtained via penalizing the second derivative in the signal to capture non-oscillating features.

For the oscillatory components n, we consider the negative Sobolev space $\mathcal{H}^{-1}(\Omega)$, on the interval Ω, which is the dual space of $\mathcal{H}_0^1(\Omega)$, and is endowed with the following norm

$$\|n\|_{\mathcal{H}^{-1}} = \inf\left\{\sqrt{\sum_i |g_i|^2} \;\;\Big|\; n = D^T g\right\} \approx \|g\|_2\,, \tag{4}$$

where D denotes the first-order linear differential operator, and D^T the adjoint operator. Therefore, given the inner product $<\cdot,\cdot>$, D and D^T satisfy the fundamental relation $<x,Dy>=<D^Tx,y>$. A visual insight on the motivation for using Sobolev spaces of negative differentiability to model oscillations is given in Fig. 4 where the function g (red colored) is over-imposed on the function n (black colored) which consists of a piecewise-constant signal of six different frequencies with the same amplitude. The associated function g, computed as $g = (DD^T)^{-1}Dn$, has ℓ_2-norm values which decrease with increasing frequency of n. Therefore, by penalizing $\|g\|_2$ we aim to capture the higher-frequency oscillatory parts of the signal. We remark that since we do not assume any a priori knowledge, Sobolev spaces of negative differentiability may capture both oscillating features as well as noise. If a priori more information about the type of oscillation is known, different G-norms can be explored [32].

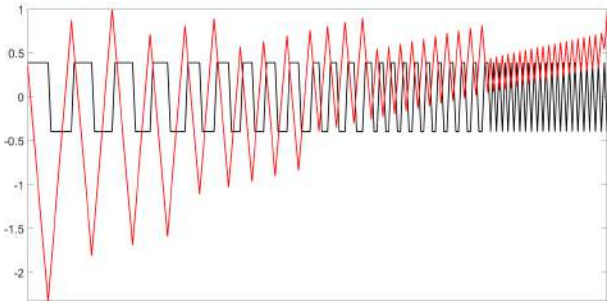


Fig. 4. Signal n (black) with increasing frequency and same magnitude oscillations; associated function g (red). As the frequency of n increases, $\|g\|_2$ decreases.

By adopting the norm in (4) the considered variational problem reads as

$$\{\bar{v},\bar{w},\bar{g}\} \leftarrow \arg\min_{v,w,g\in\mathbb{R}^N} \mathcal{J}(v,w,g)\,, \tag{5}$$
$$\mathcal{J}(v,w,g) := \frac{1}{2}\,\|v+w+D^Tg-f\|_2^2$$
$$+\gamma_1\sum_{j=1}^{N}\phi\left(\|(Dv)_j\|;a\right)+\gamma_2\,\|Hw\|_2^2+\,\gamma_3\|g\|_2^4,$$

where the non-negative parameters γ_1,γ_2, and γ_3 are appropriately selected to balance the energies in the minimizing function $\mathcal{J}$. The oscillatory component $\bar{n}$ is finally obtained as $\bar{n} = D^T\bar{g}$.

Among various image decomposition models which can be applied to signal decomposition, the proposed model of separation to three components, Jumps, Oscillations and Trend, is particularly well-suited for real data, such as the ones illustrated in Fig. 1. The main idea behind the model (5) is that the noise (and high frequency component) is captured by $\mathcal{H}^{-1}$ norm, which gives small value when it is highly oscillatory, while it is balanced with the sparsity promoting function $\phi(\cdot)$, that well filters the sparse piecewise-constant jumps, and the change in the trend is captured by the non-oscillatory component w. Notice all three norms are emphasizing different aspects of the signal: ϕ uses the sparsity of first derivatives, w is found by minimizing second derivatives for smoothness, and n is found from minimizing L_2 norm of g where $D^Tg = n$.

The proposed model (5) identifies jumps, represented by v, without any a priori knowledge on their localization; where a jump is characterized by sparse intensity discontinuities which is not part of noisy oscillations nor a trend. In addition, as side result from the numerical procedure applied to minimize the functional (5), the model determines the signal $\bar{z}$ in (20) which allows for an automatic detection of the jump location.

In many practical applications the amplitude h of the constant step jumps contained in v is a priori known, see e.g., Fig. 1(b). In such cases, we can impose a constraint in the minimization problem (5) to force $\bar{v}$ to be a binary vector. This

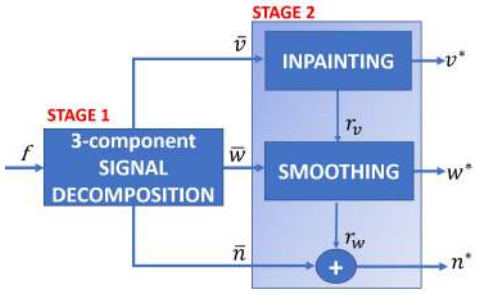


Fig. 5. Diagram scheme of the proposed JOT framework.

leads to the following constrained minimization problem:

$$\{\bar{v},\bar{w},\bar{g}\} \leftarrow \arg\min_{v,w,g\in\mathbb{R}^N} \mathcal{J}(v,w,g) \;\text{ s.t. }\; v\in\{0,h\}^N \quad (6)$$

where the functional $\mathcal{J}(v,w,g)$ is the same as in (5).

B. Stage 2: Residual Aided Refinement

Due to the ill-posedness of the inverse decomposition problem, in practice the three resulting components from stage 1 may be slightly mixed, regardless of the choice of the regularization parameters γ_1,γ_2,γ_3, which anyway strongly depend on the signal morphology. Some trend can be find in $\bar{v}$, and some noise (highly oscillating behaviour) in the smoothed component $\bar{w}$. The trend contribution in $\bar{v}$ should be extracted and added to $\bar{w}$, as well as the noise, mistakenly introduced into $\bar{w}$, should then be added to $\bar{n}$. This is illustrated in the diagram scheme in Fig. 5.

The goal of this second stage is thus to refine the results $\bar{v}$, $\bar{w}$, $\bar{n}$ obtained from stage 1 to satisfy the specific application target when needed. We note here that in real experiments results from stage 1 can be good enough. In the numerical section, we report some final results obtained directly from stage 1, e.g. Figure 11, 12 and 15.

Refinement of $\bar{v}$. The refined v^* component is obtained from $\bar{v}$, solution of problem (5), by subtracting the intrinsic oscillating parts:

$$v^* = \bar{v} - r_v \ ,$$

where r_v is carried out by filling in $\bar{v}$ the intervals corresponding to the jumps with piecewise-constant samples (hence named *inpainting*).

To this aim, by exploiting the automatic jump-detection result, provided in stage 1 and discussed in Section III, we build a binary vector $d \in \mathbb{R}^N$ with components one only in correspondence of the indices of $\bar{v}$ that belong to the K localized jumps. We set the matrix $M \in \mathbb{R}^{(N-K)\times N}$ to be the sampling matrix obtained from the identity matrix I_N by eliminating its rows corresponding to the non-zero indices in d. We denote by $M_c \in \mathbb{R}^{K\times N}$ its complement matrix, consisting of the rows of I_N not appearing in M.

The inpainted residual $r_v \in \mathbb{R}^N$ is then expressed as a sum of two terms

$$r_v = M^T y + M_c^T x \ , \quad (7)$$

where $y = M\bar{v}$ represents the samples from $\bar{v}$ to be kept and $x \in \mathbb{R}^K$ is the remaining unknown samples. This is obtained by solving the following small dimensional linear least squares problem

$$x^* \leftarrow \arg\min_{x\in\mathbb{R}^K} \left\| D(M^T y + M_c^T x)\right\|_2^2 . \quad (8)$$

Refinement of $\bar{w}$. The spurious trend extracted from $\bar{v}$, computed as r_v in (7), can now be added to $\bar{w}$. The final smooth component w^* is then computed by solving the quadratic optimization problem

$$w^* \leftarrow \arg\min_{w\in\mathbb{R}^N} \frac{1}{2}\|w-(\bar{w}+r_v)\|_2^2 + \alpha\|Dw\|_2^2 \quad (9)$$

which smoothes the $\bar{w}+r_v$ signal. The convex optimization problem (9) is a Tikhonov-like variational formulation that enforces smooth solutions. We note that problem (9) generalizes to the H-P trend estimation method (see [11]) when D is replaced by second-order difference matrix H. The H-P filtering is supported in several standard software packages for statistical data analysis, e.g., SAS, R, and Stata.

The component $\bar{w}$ may contain residues of highly oscillating components, that we collect in the residual signal given by

$$r_w = (\bar{w}+r_v) - w^* . \quad (10)$$

Refinement of $\bar{n}$. The $\bar{n}$ component from Stage 1 is updated by adding to it the highly oscillating residual from w^*, computed as r_w in (10),

$$n^* = \bar{n} + r_w . \quad (11)$$

In Fig. 6, we present the workflow for the signal f_1 shown in Fig. 2, following the proposed JOT decomposition framework illustrated in Fig. 5. Each graph of $\bar{v},\bar{w},\bar{n},v^*,w^*$ and n^* is presented in red and superposed against the ground truth signals in blue to show the effects of each stage. We observe the important role of the residuals which refine the decomposition results obtained in Stage 1.

In real applications, one may combine three components v^*, w^*, and n^* differently: (i) If the focus is on the jumps with an expected amplitude, e.g. in noisy barcode reading, the clean component v^* should provide this information; (ii) If the focus is on further analysis of the signal without jumps, one can analyze either w^* or n^* separately, or provide the reconstructed signal without jumps w^*+n^*.

III. Numerical Implementations

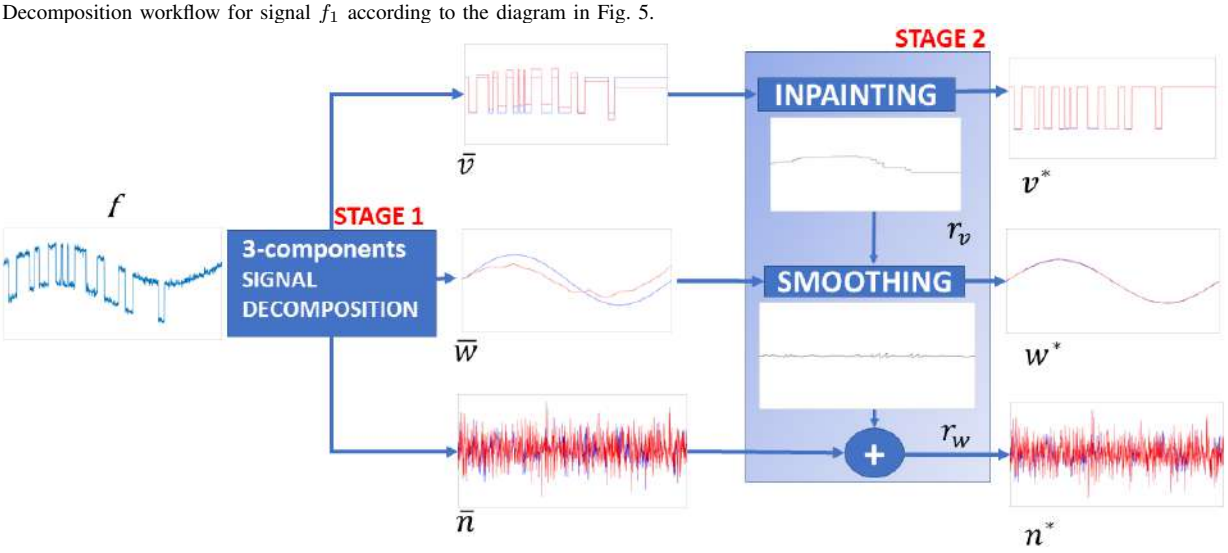
The first-order difference of an N-point signal x is approximated by forward finite difference scheme and represented in matrix vector form as Dx where D is the matrix

$$D = \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{(N-1)\times N}. \quad (12)$$

Analogously, the second-order difference operator H is approximated by the finite central difference scheme and represented in matrix vector form as Hx where H is the matrix

$$H = \begin{bmatrix} -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \end{bmatrix} \in \mathbb{R}^{(N-2)\times N}. \quad (13)$$

Fig. 6. Decomposition workflow for signal f_1 according to the diagram in Fig. 5.



The original signal is padded to obtain zero Neumann boundary conditions for first and second-order differential operators.

A. Stage 1: Three-Component Decomposition

The numerical solution of the optimization problem (5) is obtained via alternating strategy which deals with the variables separability and makes full use of the separable structure of the objective function. The iterative scheme consists in minimizing successively with respect to each variable, that we chose to be $v \in \mathbb{R}^N$ and $x := (w; g)^T \in \mathbb{R}^{2N}$, that is to solve iteratively, for initialized $v^{(0)}$ and $x^{(0)}$, the two following optimization problems:

$$v^{(k+1)} = \underset{v \in \mathbb{R}^N}{\operatorname{argmin}} \mathcal{J}(v, x^{(k)}) \,, \tag{14}$$

$$x^{(k+1)} = \underset{x \in \mathbb{R}^{2N}}{\operatorname{argmin}} \mathcal{J}(v^{(k+1)}, x) \,. \tag{15}$$

For the x**-subproblem** (15), by imposing the first-order optimality conditions, and replacing the nonlinear term $\left\|g^{(k+1)}\right\|_2^2$ with the value at the previous iteration k, we obtain the following linear system of equations

$$L x^{(k+1)} = y \,, \tag{16}$$

where

$$L = \begin{bmatrix} I_N + \gamma_2 H^T H & D^T \\ D & DD^T + 2\gamma_3 \left\|g^{(k)}\right\|_2^2 I_N \end{bmatrix},$$

$$y = \begin{bmatrix} f - v^{(k+1)} \\ D(f - v^{(k+1)}) \end{bmatrix}$$

which is solved for $x^{(k+1)} = (w^{(k+1)}; g^{(k+1)})^T$.

The system is symmetric, positive definite, but can be slightly ill-conditioned, which leads to add the regularization term $\frac{\kappa}{2}\|x\|_2$ to the x-subproblem (15), with $\kappa > 0$ a very small scalar parameter. Hence, a suitable regularized solution is determined by replacing (16), with the following regularized system

$$(L + \kappa I_{2N})\, x^{(k+1)} = y, \tag{17}$$

where I_{2N} denotes the identity matrix of order $2N$. The system is efficiently solved using iterative preconditioned conjugate gradient linear solver.

The v**-subproblem** (14) can be written as

$$v^{(k+1)} = \underset{v \in \mathbb{R}^N}{\operatorname{argmin}} \mathcal{J}_1(v), \tag{18}$$

$$\mathcal{J}_1(v) := \frac{1}{2}\|v - q\|_2^2 + \gamma_1 \sum_j \phi(|(Dv)_j|; a),$$

where $q := f - w^{(k)} - D^T g^{(k)}$. Necessary conditions for strong convexity of the cost function $\mathcal{J}_1(v)$ in (18) are reported in Proposition 1 and demonstrated in [16]. The minimizer of (18) is then carried out by a forward-backward splitting (FBS) iterative scheme which is proved to converge to the unique solution of (18).

Proposition 1. *Let $f \in \mathbb{R}^N$, $\gamma_1 > 0$, and $\phi(\cdot; a)$ be the penalty function defined in* (3)*. If*

$$0 \le a \le \frac{1}{4\gamma_1} \tag{19}$$

then the cost function $\mathcal{J}_1(v)$ in (18) *is strongly convex and the the FBS iterative algorithm*

$$z^{(i)} = D^T\big(Dv^{(i)} - \operatorname{soft}_{1/a}(Dv^{(i)})\big) \tag{20}$$

$$v^{(i+1)} = \arg\min_{v \in \mathbb{R}^N} \left\{ \frac{1}{2}\|q + \gamma_1 a z^{(i)} - v\|_2^2 + \gamma_1 \|Dv\|_1 \right\}, \tag{21}$$

converges to the unique minimizer of $\mathcal{J}_1(v)$.

For the proof of convergence we refer the reader to [16].

We observe that the backward step (21) is a standard 1-dimensional TV denoising which can be calculated exactly in finite-time [33]. In the forward step (20) the soft threshold

function $\mathrm{soft}\colon \mathbb{R} \to \mathbb{R}$ with threshold parameter $\lambda \geq 0$ is defined as

$$\mathrm{soft}_\lambda(y) := \begin{cases} 0, & |y| \leq \lambda \\ (|y| - \lambda)sign(y), & |y| \geq \lambda. \end{cases} \tag{22}$$

If the soft threshold function is applied to a vector, then it is applied component-wise.

As observed in [16], the signal z^* computed by (20) upon convergence of the algorithm behaves as a jump-detector. In fact, by applying the soft threshold function in (22), we get

$$z^* = \begin{cases} \Delta v, & |Dv| \leq \frac{1}{a} \\ 0, & |Dv| \geq \frac{1}{a}. \end{cases} \tag{23}$$

Hence, the signal z can be interpreted as the response to a Laplacian (Lap) operator ($\Delta = D^T D$) on the piecewise-constant signal v. The Lap operator takes the second derivative of the signal: where the signal is basically flat, the Lap will give zero, in correspondence of jumps with minimal jump height lower than $1/a$, the signal z will zero-crossing the x-axis.

As illustrated in Figure 7(left), if we consider a smooth jump edge v with an increasing slope ($Dv > 0$), the inflection point of the step edge happened to be the point in which the slope of the edge (first derivative - Figure 7 middle -) stops increasing, reaches its maximum then starts decreasing, while the second order derivative - Figure 7 bottom - vanishes ($\Delta v = 0$). In case of piecewise-constant signals v, which are of interest in our work, we illustrate in Figure 7(right) the resulting z^* obtained in stage 1 upon convergence by processing the original signal f_1. The jump positions of v are located sufficiently close to the zero values of z^* between two successive peaks.

The *jump width* effects the amplitude of derivatives: the wider jump has the smaller derivative amplitude, and the effect becomes more noticeable at higher derivative orders. In general, it is found that the amplitude of the nth derivative of a jump is inversely proportional to the nth power of its width, for signals having the same shape and amplitude. The amplitude of a derivative of a jump also depends on the shape of the jump and is directly proportional to its *jump height*.

This motivated us to use the side result z^* from stage 1 to identify the jump locations and, accordingly, define a suitable vector d for the refinement of $\bar{v}$.

Moreover, the value of a can be interpreted, according to (23), as a threshold of the desired sensitivity to the jump height, over which the signal oscillation will not be considered a jump. Once a is given, the parameter γ_1 is carried out to satisfy (19) in order to preserve the strong convexity of the problem.

B. Stage 2: Residual Aided Refinement

Numerical solution of the least squares problem (8) in the refinement of $\bar{v}$ is obtained by solving the associated normal equation system $A^T Ax = A^T b$ with $A := DM_c^T \in \mathbb{R}^{N \times K}$ and $b := -DM^T y \in \mathbb{R}^N$. Being A a full rank matrix, the refinement for $\bar{v}$ has a unique least square solution.

The smoothing quadratic minimization problem (9) is a convex optimization problem which unique minimizer is explicitly given by imposing the first optimality conditions which lead to the solution of the following linear system of equations

$$(I_N + \alpha D^T D)w = \bar{w} + r_v \ , \tag{24}$$

with symmetric positive definite coefficient matrix.

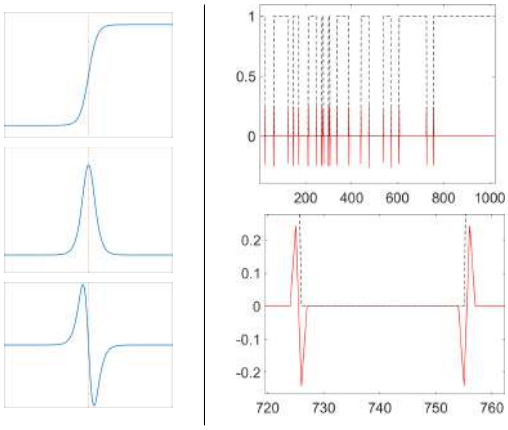
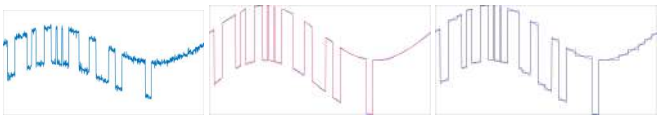


Fig. 7. Left: a smooth sigmoid signal (top), and its first (middle) and second (bottom) derivatives; Right: a non-smooth signal v (dashed black) and z^* signal in solid red (top); detail shown in the bottom.



input data f_1 $\quad$ $v^* + w^*$ of JOT $\quad$ v of [15]

Fig. 8. Decomposition results of the input data f_1 (left). JOT result in (middle, red), $RMSE = 0.0112$, and Potts result (right, black), $RMSE = 0.3534$, against the ground truth (blue).

IV. Numerical Examples

We present numerical examples for both artificial signals, where we know the ground truths, and real life signals to showcase the proposed JOT model when applied to the study of nonstationary signals. We compare with the Potts model [15], and with the model proposed in [13] which combines Total Variation and ℓ_2 regularizations to reconstruct piecewise smooth signals. We note here that [15] yields a jump segmentation without trends within the segments, i.e. it effects a piece-wise constant component recovery.

A. Synthetic examples

The first result is presented in Figure 2 for the synthetic signals f_1 and f_2. From the second column, the final components v^*, w^*, n^* produced by the two stages of JOT are presented. We observe that jumps as well as trends are well-separated from noisy oscillations. In Figure 8, we compare the JOT result (red) with the ground truth (blue), and the Potts model in [15] (black) for the signal f_1. By using the Potts model, which relies on a sparsity-inducing ℓ_0 pseudonorm penalty, the signal is decomposed into v and the residual $f_1 - v$. JOT captures the trend more closely to the ground truth, without staircase effect. This is confirmed by the Root Mean Squared Error (RMSE) reported in Figure 8.

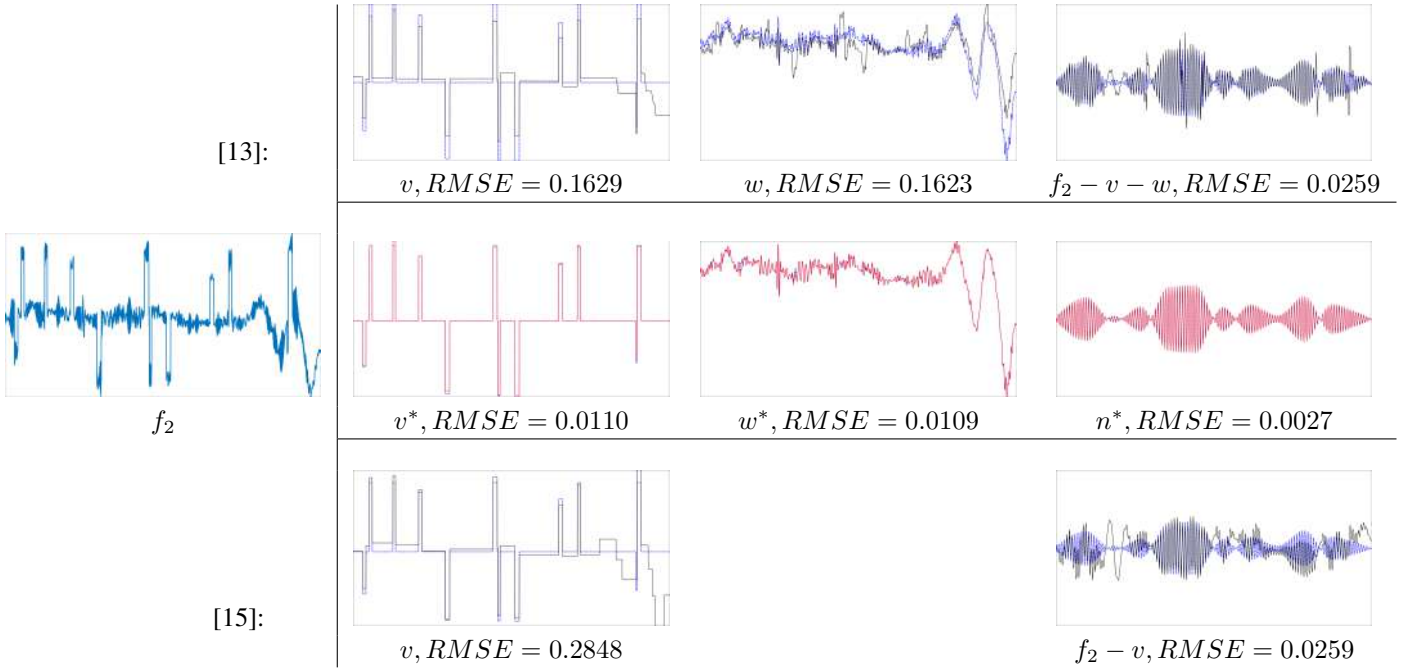


Fig. 9. Decomposition results for input data f_2 (left) where all graphs are superposed with the ground truth data (dashed blue): JOT (red, middle row), [13] (black, top row), and [15] (black, bottom row). RMSE values are lower for v^* and w^* of JOT.

In Figure 9, we compare the decomposed components v^*, w^*, n^* by JOT (middle row), the method proposed in [13] (bottom row) and Potts model (top row) for the signal f_2. The superiority of the minimax penalty (3) applied in JOT with respect to the Total Variation penalty used in the variational model in [13] is appreciable in the v^* components. The advantage of using a specific $\|n\|_{\mathcal{H}^{-1}}$ term to capture the oscillatory component, instead of a L_2 norm of the residual as used in [13], is visible in the last column. Residual structures from v and w can be noted in the first and third row even with an optimal tuning of the parameters involved.

In Figure 10 first row, we report the raw signal f_2, the Potts model result, and the w^* term from JOT decomposition framework. In the second row we show the corresponding Short Time Fourier Transform, a.k.a. spectrogram. In the third and fourth row the time frequency plots obtained via CWT based Synchrosqueezing (SST) [1] and the IMFs based IMFogram [5] are shown respectively. The IMFs have been produced using the FIF algorithm [21]. The CWT of these data sets are shown in Figure 3.

All the time-frequency plots confirm that the proposed method better remove jumps from the signal. This is particularly evident in the spectrogram plots (second row of Figure 10). In the middle panel, we see leftover yellow vertical lines, which correspond to the jumps, as explained for instance in [6]. Whereas these vertical lines disappear on the right panel. The other time-frequency plots, which are more focused and less prone to artifacts than the Spectrogram, in particular the SST and the IMFogram, allow to notice that also low frequency contributions coming from the jumps are removed properly from the plots. In particular, around roughly 10 Hz the right panels in the third and fourth row are empty at any time, whereas the corresponding plots from

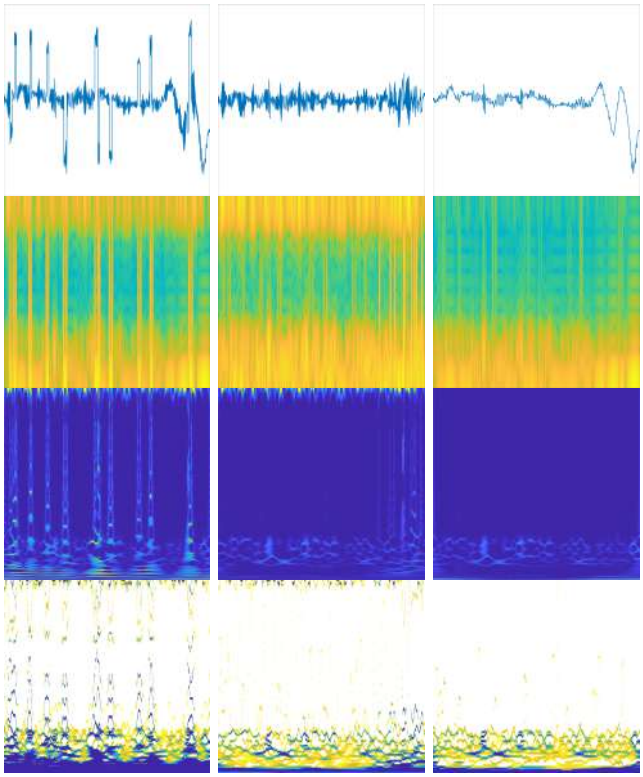


Fig. 10. The top row shows the signal f_2, the Potts result v and the JOT result w^*. The rows below show their time-frequency representations. From the second to last row, Spectrograms, the CWT based Synchrosqueezing, and the IMFogram are presented respectively.

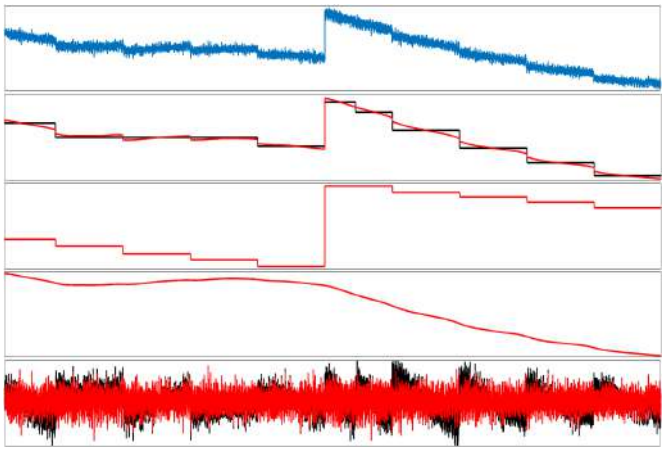


Fig. 11. Decomposition results of JOT (stage 1) compared with Potts method. Top row shows the input data "CSES_Exd" f_3. Second row: $\bar{v}+\bar{w}$ of JOT (red) compared with the Potts $\bar{v}$ component (black). Third and fourth row depict, respectively, $\bar{v}$ and $\bar{w}$ produced by JOT. Last row: $\bar{n}$ of the proposed method (red) and the Potts $\bar{n}$ component (black).

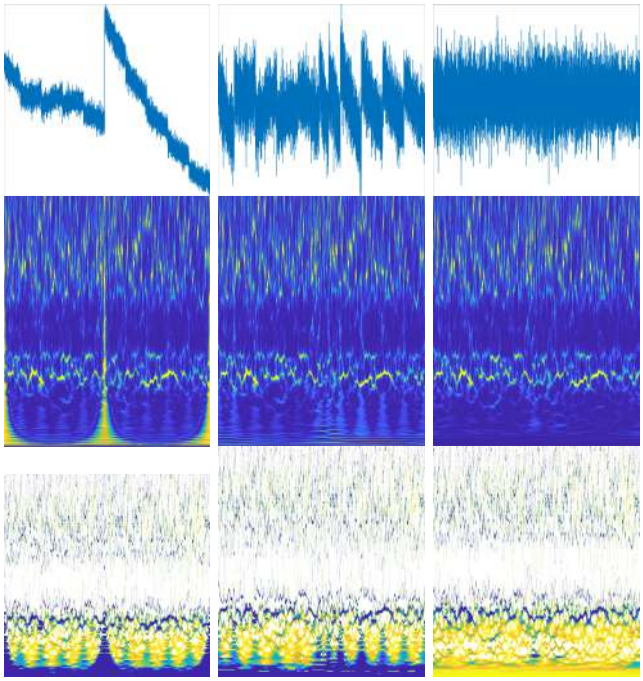


Fig. 12. The first row shows the given signal CSES f_3, $f_3 - v$ from [15] and $\bar{n}$ from JOT. Second and third rows illustrate, column-wise, their associated time-frequency representations. The second row shows CWT based Synchrosqueezing, and the bottom row, the IMFogram.

the central column present still some contributions towards the end of the data set. Furthermore, the proposed approach has the advantage allowing to separate "for free" also the high frequency components contained in the signal, as shown in the time-frequency plots. In fact, from 50 Hz above the right column panels are completely empty.

JOT framework allows to pre-process data, and, as a consequence, to produce cleaner and crisper time-frequency plots of a signal by separating jumps and high frequency contributions in an accurate way.

B. Real life Signals

We illustrate the decomposition of three real signals (IV-B1-IV-B3) using the proposed JOT framework. We demonstrate the efficacy of the JOT signal decomposition as a pre-processing tool in signal analysis.

1) Earth's Electric Field: The study of the Earth magnetic and electric field is an active research field in Geophysics. Among the open questions that are still waiting to be addressed, we mention the identification and removal of the magnetic and electric background, i.e. the trend and low frequency oscillations, the analysis and classification of chirps, i.e. quick changes in frequency of one of more components contained in the signal, and other transient phenomena [34].

In Figure 11 first row, we show a sample of the Earth's electric field measured by the CSES-01 satellite. CSES-01 is a sun-synchronous satellite flying at an altitude of $\sim$ 507 km [7]. The orbital descending node time is around 14:00 LT and the revisiting period is of 5 days. In Figure 11, we present the stage 1 of JOT results showing $\bar{v}, \bar{w}$ and $\bar{n}$, respectively in the third, fourth and last row. In the second row of Figure 11, we compare with Potts. The proposed method better separates the jumps producing a cleaner high frequency oscillations remainder $\bar{n}$ than the $f_3 - v$ of [15], as shown in the last row of the same figure.

This clear separation is also confirmed by the time-frequency plots in Figure 12, which are associated with the raw signal f_3, compared with the ones associated with the $f_3 - v$ and $\bar{n}$ components. In particular, even if both approaches allow to properly remove the main jump present in the middle of the raw signal, only the proposed approach allows to remove completely the secondary jumps present in f_3. This is shown by the presence of vertical bars in the CWT of $f_3 - v$ and blue columns in its IMFogram, central column of Figure 12. These structures disappear, if we study the $\bar{n}$ component of JOT, as shown in the right column plots of Figure 12.

Furthermore, the proposed method allows to identify and separate the electric field background (i.e. trend and low frequency oscillations), depicted in the fourth row of Figure 11. JOT framework proves to be valuable for the pre-processing of this kind of data.

2) CPU emanated EM signals: As a second real life example, we present, in Figure 13 first row, the signal f_4 representing electromagnetic side-channel emanations of a computer processor, i.e. Emanated EM signal. This signal was measured with a near field magnetic probe (Aaronia H2 near-field magnetic probe) located around the pins of the processor of a target device (in this particular case is a A13-OLinuXino). The processor of this specific device is ARM Cortex A8 whose operating clock frequency is 1 GHz. For recording the measured data, Keysight UXA signal analyzer is used. The signal is recorded with 1.28 GHz sampling frequency around 1 GHz center frequency and 500 MHz bandwidth. For more details on the data and their measurement the interested reader can refer to [9], [35].

The emanated EM signals are measured to estimate the side-channel leakage, which is a consequence of program execution in a computer processor. Understanding the relationship be-

tween code execution and information leakage is a necessary step in estimating information leakage and its capacity limits.

The data collected present many sudden drops due to, probably, other activities of the motherboard. In Figure 13, the JOT decomposition captures the sudden jumps as v^* (second row) and the low frequency component as w^* (third row), and the data of interest n^* (fourth row) in red. Superposed in black color, we provide comparison with components v, w and $f_4 - v - w$ obtained by [13]. Stage 2 of JOT does not only provide better alignment in v^*, but the superior penalty function we use in stage 1 recovers better the amplitude. The last two rows of Figure 13 provide comparison of JOT with the Potts model [15]: The component $v^* + w^*$ is shown together with component v of [15] and components n^* with $f_4 - v$ of [15] respectively. Due to the Potts model not capturing the smooth trend, The component $f_4 - v$ is affected by spurious data.

The drops, represented by v^* influence the time-frequency plots introducing artifacts at any frequency level, as shown in Figure 14. In the first row, we present the raw data f_4 and its associated time-frequency representation plots, whereas in the second row, we plot the time-frequency representations associated with the n^* signal of the proposed JOT.

By comparing the first and second row time-frequency plots of Figure 14, IMFogram time-frequency representations show how v^* and w^* components introduce artifacts at every frequency level. These artifacts mask the actual active frequencies which instead become well visible after pre-processing of the data by means of the JOT decomposition algorithm, Figure 14 second row. In particular, the cycles which are present in the frequency range $[3000,\ 4000]$ kHz can be easily recognized after the JOT pre-processing, in the second row right IMFogram in Figure 14.

3) ECG Data with moving electrodes: In Figure 15, we plot the signal f_5 which represents an ECG recording created starting from a clean recording from the MIT-BIH Arrhythmia Database, to which calibrated amounts of noise has been added, to simulate electrode motion artifacts [8]. Noise was added beginning after the first 5 minutes of each record, during two-minute segments alternating with two-minute clean segments. The signal-to-noise ratio (SNR) measured in dB during the noisy segments of this record is -6 dB[1].

In this example is evident how the proposed method is able to produce a better jump identification even in a scenario in which the jumps can be represented as a piece-wise constant function.

C. Convergence and CPU time

We present an empirical investigation on the numerical convergence of the proposed Stage 1 minimization scheme. In Figure 16, we present the convergence plots showing the decay of the energy functional $\mathcal{J}$ versus the number of iteration of the numerical optimization scheme for all the signals presented in this paper. The plots are normalized for an overall visualization.

[1]The data and more details on how they have been produced can be found in the Physionet database at https://physionet.org/content/nstdb/1.0.0/

	f_1	f_2	f_3	f_4	f_5
N	1024	4450	20000	330000	650000
time(s)	0.0031	0.0286	0.0487	0.9667	1.9069
a	8.00e+2	1.39e+4	1.25e+7	5.00e+5	8.68e+8
γ_1	3.84e-2	2.50e-6	6.00e+1	1.20e+0	4.17e+3
γ_2	7.52e+2	2.31e+1	2.40e+8	1.15e+2	–
γ_3	1.80e-4	1.50e-3	1.00e-12	2.90e-11	2.00e-20

TABLE I
CPU TIMING IN SECONDS PER ITERATION.

For an insight onto the computational complexity of the proposed optimization Stage 1, we report in Table I the cpu time per iteration in terms of the signal dimension N for every signal considered in the experiments illustrated in section IV. The timings refer to a naive implementation in MATLAB environment on an Intel®Core™i7-8565U CPU with 1.99GHz and 16GB RAM.

An overall timing of stage 1 of JOT framework come straightforwardly by combining data in Table I with number of iterations to achieve convergence reported in Figure 16.

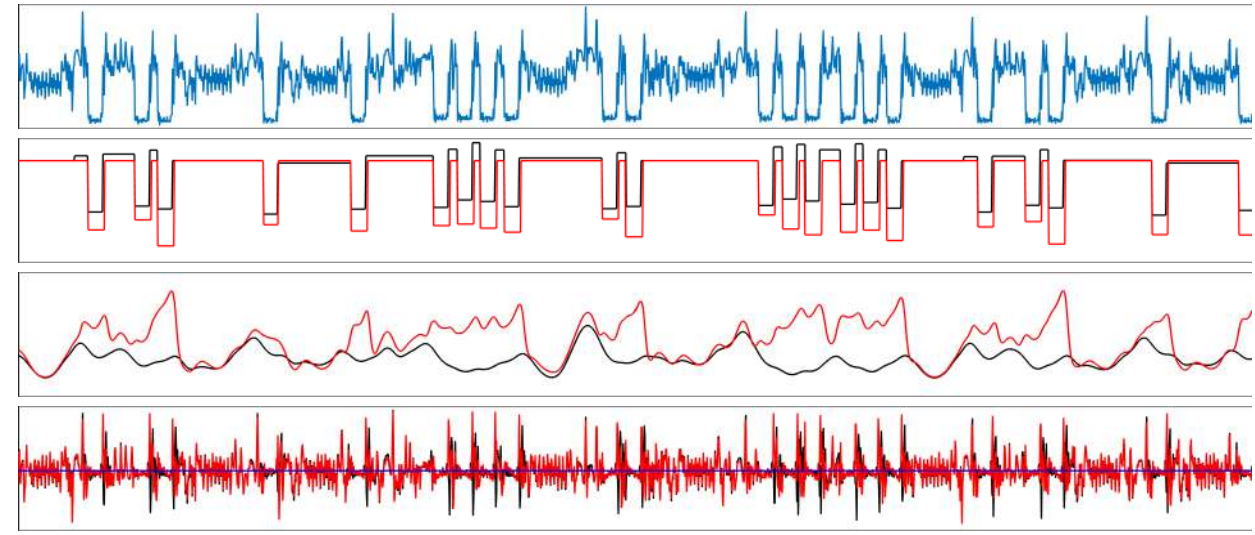
The computational cost of Stage 2 of the proposed JOT framework consists mainly in the solution of two linear systems (8) and (9) (via (24)), where the inpaining system is of dimension $K << N$ while the smoothing linear system is highly sparse and involves N unknowns. The overall computational cost of stage 2 corresponds at most to one iteration of Stage 1 reported in Table I.

V. CONCLUDING REMARKS

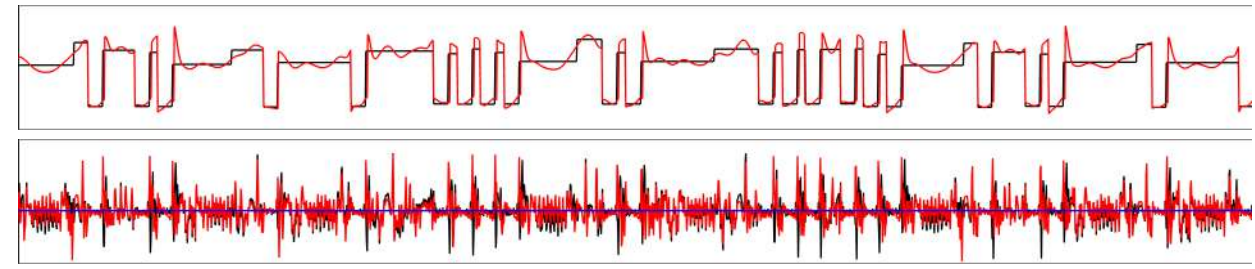
Many state of the art methods for time-frequency analysis aim to extract frequencies hidden in nonstationary signals. The estimation of the signal trend is of fundamental importance in many applied fields of research, like physics, economy, medicine, and many more. However, when the signal under investigation is contaminated by anomalies/jumps, any state of the art signal processing technique ends up producing time-frequency representations which are corrupted by artifacts, and the proper estimation of the signal trend becomes extremely challenging.

In this paper, we proposed a two-stage signal decomposition method, called JOT, which allows to separate piecewise-constant jumps, as well as the oscillatory component, from the non-oscillating trend. The framework relies on the solution of convex and non-convex optimization problems by means of an efficient and robust alternating optimization algorithm and FBS solver. The theoretical convergence is currently only supported by robust empirical tests. We plan to explore this in the future. The numerical procedure proposed for the solution of the three component decomposition (stage 1) is new, and differs from the ADMM-based numerical solution described in [29] for image decomposition. This new alternating procedure allows for an automatic estimation of the jump locations, and an exact evaluation of the 1-dimensional TV denoising which is not available for the 2-dimensional case.

To demonstrate the effectiveness of the proposed framework, we compared its results on synthetic and real life signals with the ones obtained by the decomposition models presented in [13] and [15]. Extending from decomposition

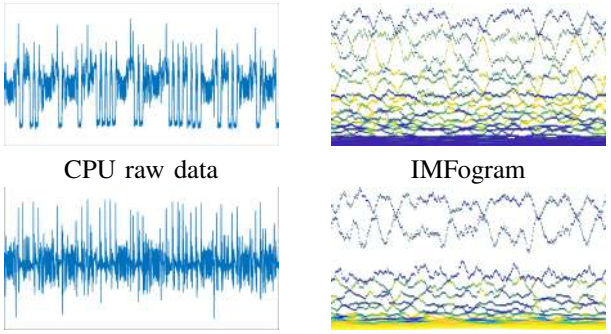


Results of [13] (black) and JOT (red).



Potts [15] (black) and JOT (red) results

Fig. 13. The first row shows the given CPU signal f_4 The second to fourth row show the JOT results v^*, w^*, n^* in red respectively, superposed in black v, w and $f_4 - v - w$ of [13]. The fifth and sixth row illustrate JOT result $v^* + w^*$ and n^* in red, superposed in black the Potts result v $f_4 - v$ of [15].



CPU raw data IMFogram

CPU JOT pre-processed data IMFogram

Fig. 14. Time-frequency representations of the CPU signal f_4 in Figure 13: The second row shows JOT result, subtracting the v^* and n^* components from f. Notices IMFogram shows more clear separation of the signal for JOT.

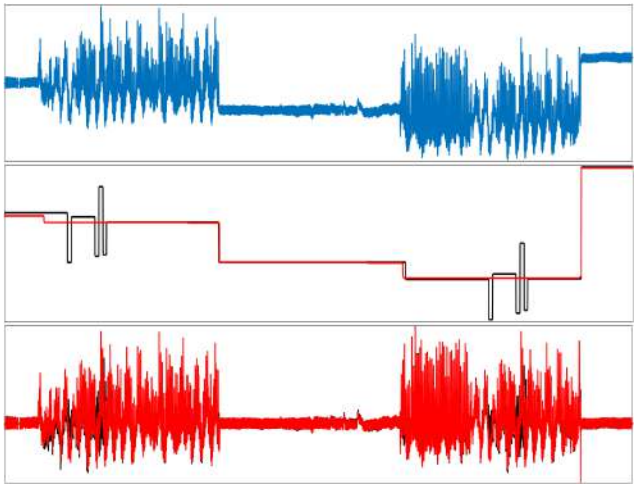


Fig. 15. Decomposition of the input signal "118e_6m" f_5 (first row) into $\bar{v}$ (second row) and $\bar{n}$ (third row) assuming $\bar{w} = 0$. In red, the results of stage 1 in comparison to the Potts model result (black).

to signal recovery or segmentation, one can consider higher order Mumford-Shah model for signal separation, such as [12]. In general, we found that (i) the proposed JOT method identifies not only big but also small step jumps more stably, e.g., Figure 11 and 15, and (ii) having a dedicated term (the $\mathcal{H}^{-1}$-norm term) in the functional (2) makes a difference in capturing highly oscillating components (representing either noise or signal component). Without such a term, in general, the residue from a fitting term collects spurious data from the other components, e.g. see Fig.9, third column and [29].

JOT framework proves to be a reliable and robust method for the separation of jumps and low frequency oscillations/trend from a nonstationary signal. JOT decomposition can become a valuable pre-processing algorithm in real life applications for the detrending and time-frequency analysis of signals containing anomalies and jumps.

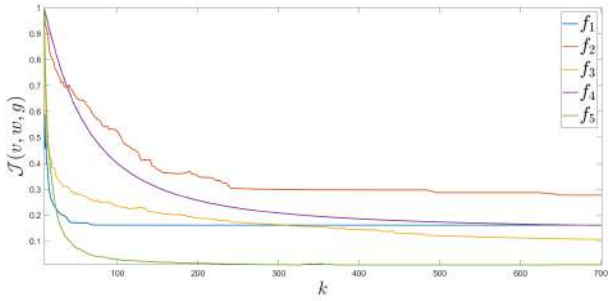


Fig. 16. Plots providing empirical evidence of numerical convergence of the stage 1, for the signals analyzed.

VI. Acknowledgement

This work made use of the data from CSES mission, a project funded by China National Space Administration (CNSA) and China Earthquake Administration in collaboration with Italian Space Agency and Istituto Nazionale di Fisica Nucleare. CSES data are freely available through the CSES mission at http://www.leos.ac.cn/#/home. Research by SM, MH, and AC was supported in part by the National Group for Scientific Computation (GNCS-INDAM), Research Projects 2021. SHK was supported in part by the Siomns Foundation Grant 584960. AC thanks the Italian Space Agency, for the financial support under the contract ASI "LIMADOU scienza +" n° 2020-31-HH.0, and the ISSI-BJ project "the electromagnetic data validation and scientific application research based on CSES satellite" and Dragon 5 cooperation 2020-2024 (ID. 59236).

References

[1] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Applied and computational harmonic analysis*, vol. 30, no. 2, pp. 243–261, 2011.

[2] 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," *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[3] P. Singh, "Novel fourier quadrature transforms and analytic signal representations for nonlinear and non-stationary time-series analysis," *Royal Society open science*, vol. 5, no. 11, p. 181131, 2018.

[4] J. Antoni, "The spectral kurtosis: a useful tool for characterising non-stationary signals," *Mechanical systems and signal processing*, vol. 20, no. 2, pp. 282–307, 2006.

[5] P. Barbe, A. Cicone, W. Suet Li, and H. Zhou, "Time-frequency representation of nonstationary signals: the imfogram," *Pure and Applied Functional Analysis*, 2021.

[6] A. Stallone, A. Cicone, and M. Materassi, "New insights and best practices for the successful use of empirical mode decomposition, iterative filtering and derived algorithms," *Scientific Reports*, vol. 10, p. 15161, 2020.

[7] P. Diego, J. Huang, M. Piersanti, D. Badoni, Z. Zeren, R. Yan, G. Rebustini, R. Ammendola, M. Candidi, Y.-B. Guan *et al.*, "The electric field detector on board the china seismo electromagnetic satellite—in-orbit results and validation," *Instruments*, vol. 5, no. 1, p. 1, 2021.

[8] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[9] B. B. Yilmaz, M. Prvulovic, and A. Zajić, "Electromagnetic side channel information leakage created by execution of series of instructions in a computer processor," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 776–789, 2019.

[10] L. J. Christiano and T. J. Fitzgerald, "The band pass filter*," *International Economic Review*, vol. 44, no. 2, pp. 435–465, 2003.

[11] R. J. Hodrick and E. C. Prescott, "Postwar u.s. business cycles: An empirical investigation," *Journal of Money, Credit and Banking*, vol. 29, no. 1, pp. 1–16, 1997.

[12] M. Storath, L. Kiefer, and A. Weinmann, "Smoothing for signals with discontinuities using higher order mumford–shah models," *Numerische Mathematik*, vol. 143, no. 2, pp. 423–460, 2019.

[13] A. Gholami and S. Hosseini, "A balanced combination of tikhonov and total variation regularizations for reconstruction of piecewise-smooth signals," *Signal Processing*, vol. 93, no. 7, pp. 1945 – 1960, 2013.

[14] S.-J. Kim, K. Koh, S. Boyd, and D. Gorinevsky, "ℓ_1 trend filtering," *SIAM Review*, vol. 51, no. 2, pp. 339–360, 2009.

[15] M. Storath, A. Weinmann, and L. Demaret, "Jump-sparse and sparse recovery using potts functionals," *IEEE Transactions on Signal Processing*, vol. 62, no. 14, pp. 3654–3666, 2014.

[16] I. Selesnick, A. Lanza, S. Morigi, and F. Sgallari, "Non-convex total variation regularization for convex denoising of signals," *Journal of Mathematical Imaging and Vision*, vol. 62, no. 6-7, pp. 825–841, Jul. 2020.

[17] G. Cai, I. W. Selesnick, S. Wang, W. Dai, and Z. Zhu, "Sparsity-enhanced signal decomposition via generalized minimax-concave penalty for gearbox fault diagnosis," *Journal of Sound and Vibration*, vol. 432, pp. 213–234, 2018.

[18] C. Zhang, "Nearly unbiased variable selection under minimax concave penalty," *Ann. Statist.*, vol. 38, no. 2, pp. 894–942, 04 2010.

[19] P. Flandrin, *Time-frequency/time-scale analysis*. Academic press, 1998.

[20] L. Lin, Y. Wang, and H. Zhou, "Iterative filtering as an alternative algorithm for empirical mode decomposition," *Advances in Adaptive Data Analysis*, vol. 1, no. 4, pp. 543–560, 2009.

[21] A. Cicone, "Iterative filtering as a direct method for the decomposition of nonstationary signals," *Numerical Algorithms*, pp. 1–17, 2020.

[22] A. Cicone, W. S. Li, and H. Zhou, "New theoretical insights in the decomposition and time-frequency representation of nonstationary signals: the imfogram algorithm." *arXiv preprint*, 2021.

[23] Y. Meyer and D. Lewis, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*, ser. Memoirs of the American Mathematical Society. American Mathematical Society, 2001.

[24] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, no. 1-4, pp. 259–268, 1992.

[25] J. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle, "Image decomposition into a bounded variation component and an oscillating component," *Journal of Mathematical Imaging and Vision*, vol. 22, no. 1, pp. 71–88, 2005.

[26] J. Aujol and S. Kang, "Color image decomposition and restoration," *J. Visual Communication and Image Representation*, vol. 17, pp. 916–928, 2006.

[27] L. Lieu and L. Vese, "Image restoration and decomposition via bounded total variation and negative hilbert-sobolev spaces," *Applied Mathematics and Optimization*, vol. 58, no. 2, p. 167, May 2008.

[28] S. Osher, A. Solé, and L. Vese, "Image decomposition and restoration using total variation minimization and the H^{-1} norm," *Multiscale Modeling & Simulation*, vol. 1, no. 3, pp. 349–370, 2003.

[29] M. Huska, S. H. Kang, A. Lanza, and S. Morigi, "A variational approach to additive image decomposition into structure, harmonic, and oscillatory components," *SIAM Journal on Imaging Sciences*, vol. 14, no. 4, pp. 1749–1789, 2021. [Online]. Available: https://doi.org/10.1137/20M1355987

[30] M. Huska, A. Lanza, S. Morigi, and I. Selesnick, "A convex-nonconvex variational method for the additive decomposition of functions on surfaces," *Inverse Problems*, vol. 35, no. 12, p. 124008, 2019.

[31] I. W. Selesnick and I. Bayram, "Sparse signal estimation by maximally sparse convex optimization," *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1078–1092, 2014.

[32] J. Aujol and A. Chambolle, "Dual norms and image decomposition models," *Int. J. Comput. Vision*, vol. 63, no. 1, pp. 85–104, 2005.

[33] L. Condat, "A direct algorithm for 1-d total variation denoising," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1054–1057, 2013.

[34] P. Picozza, R. Battiston, and G. e. a. Ambrosi, "Scientific goals and in-orbit performance of the high-energy particle detector on board the cses," *The Astrophysical Journal Supplement Series*, vol. 243, no. 1, p. 16, 2019.

[35] N. Sehatbakhsh, A. Nazari, M. Alam, F. Werner, Y. Zhu, A. Zajić, and M. Prvulovic, "Remote: Robust external malware detection framework by using electromagnetic signals," *IEEE Transactions on Computers*, vol. 69, no. 3, pp. 312–326, 2019.