Published in Image Processing On Line on 2023–05–21.
Submitted on 2022–07–18, accepted on 2023–03–20.
ISSN 2105–1232 © 2023 IPOL & the authors CC–BY–NC–SA
This article is available online with supplementary materials,
software, datasets and online demo at
https://doi.org/10.5201/ipol.2023.417

A Two-stage Signal Decomposition into Jump, Oscillation and Trend using ADMM

Martin Huska[1], Antonio Cicone[2,4,5], Sung Ha Kang[3], Serena Morigi[1]

[1]Department of Mathematics, University of Bologna, Italy ({martin.huska,serena.morigi}@unibo.it)
[2]DISIM, University of L'Aquila, Italy (antonio.cicone@univaq.it)
[3]School of Mathematics, Georgia Institute of Technology, Georgia, US (kang@math.gatech.edu)
[4]Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy
[5]Istituto di Astrofisica e Planetologia Spaziali, INAF, Rome, Italy

Communicated by Laurent Oudre *Demo edited by* Martin Huska

Abstract

We present a thorough implementation of the two-stage framework proposed in [A. Cicone, M. Huska, S.H. Kang and S. Morigi, JOT: a Variational Signal Decomposition into Jump, Oscillation and Trend, IEEE Transactions on Signal Processing, 2022]. The method assumes as input a 1D signal represented by a finite-dimensional vector in $\mathbb{R}^N$. In the first stage the signal is decomposed into Jump (piece-wise constant), Oscillation, and Trend (smooth) components, and in the second stage the results are refined using residuals of other components. We propose an efficient numerical solution for the first stage based on alternating direction method of multipliers, and a solid algorithm for the solution of the second stage.

Source Code

The MATLAB source code and documentation for this algorithm are available from the web page of this article[1]. Usage instructions are included in the `README.txt` file of the archive.

Keywords: signal decomposition; variational model; non-convex optimization; jumps; trend; oscillating signals; ADMM

1 Introduction

Given a composite signal, the objective of signal decomposition is to extract and separate meaningful semantic features such as, for example, trend, oscillating, and impulse features. 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;

[1]https://doi.org/10.5201/ipol.2023.417

see [2] for a survey of linear filtering methods in trend estimation. Many advances in signal processing often aim to analyze the signal in terms of oscillating or non-oscillating features via frequency analysis. Time-Frequency analysis methods, such as Short Time Fourier Transform, wavelet, Synchrosqueezing wavelet [4], Hilbert Huang Transform, or IMFogram [1], can fail when abrupt changes and jump discontinuities appear in the signal. 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 started from the seminal work of Y. Meyer [7]. In particular in [5], the authors extended the idea to decompose the given image into a piecewise-constant part, a harmonic part, and a noisy part.

In [3], a variational framework separating the given signal into Jump, Oscillation, and Trend (JOT) is proposed, and in this work, we present a thorough description of the implementation of the JOT algorithm. The main idea is to model a given finite-length discrete 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 meaningful oscillatory components.

This paper is organized as follows. In Section 2, we present the two-stage algorithm. The first stage is described in Section 2.1, and the second stage in Section 2.2. In each subsection, we present both the variational model and the related algorithmic details. In Section 2.3 parameter selection strategies are discussed. In Section 3 we present the numerical experiments.

2 The JOT Decomposition Framework and Implementation

The workflow of the proposed framework is illustrated in Figure 1. Given a composite signal f as input, the first stage separates it into the components $\bar{v}, \bar{w}, \bar{n}$, which are then further refined in v^*, w^* and n^* as result of the second stage. Each signal $\bar{v}, \bar{w}, \bar{n}, v^*, w^*$ and n^* is presented in red and superposed against the ground truth signals in blue to show the effects of each stage.

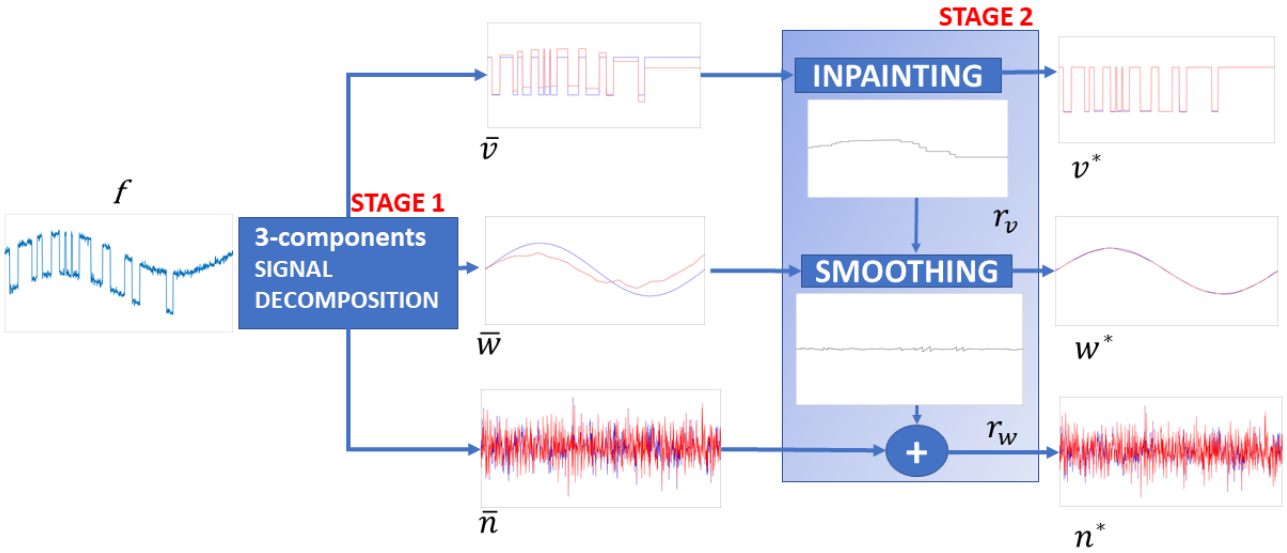

Figure 1: Decomposition workflow for a signal f.

2.1 Stage 1: Three-Component Signal Decomposition

Given the signal f, according to the decomposition model (1), the first step of the signal decomposition framework is performed via the following non-convex minimization problem

$$\{\bar{v}, \bar{w}, \bar{n}\} \leftarrow \arg \min_{v,w,n \in \mathbb{R}^N} \mathcal{J}(v, w, n), \tag{2}$$

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

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 first is the fidelity term imposing consistency with the observed signal f; the second term penalizes discrete derivative of v to promote piece-wise constant component; the third term is designed to penalize the second order derivative to capture the smooth w component. Finally, the fourth energy norm, which approximates Meyer's G-norm [7], is aimed to recover highly oscillating components in n. In particular,

- the operators D and H represent first and second-order derivatives, respectively. 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

$$D = \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{(N-1)\times N}. \tag{3}$$

  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

$$H = \begin{bmatrix} -1 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -1 \end{bmatrix} \in \mathbb{R}^{N\times N}. \tag{4}$$

  The matrices D and H are discrete approximations of the first and second derivatives, respectively, with Neumann homogeneous boundary conditions.

  Different boundary conditions can be adopted according to the signal morphology. For example, in the case of data periodicity, the periodic boundary condition is more appropriate. The algorithm described in Section 2.1.1 is independent of the selected boundary conditions.

- the function $\phi(\cdot; a) : [0, +\infty) \to [0, 1]$ is a re-parameterized and re-scaled version of the minimax penalty, a non-convex sparsity promoting function, defined as

$$\phi(t; a) = \begin{cases} -\frac{a}{2} t^2 + \sqrt{2a}\, t & \text{for } t \in \left[0, \bar{a}\right), \\ 1 & \text{for } t \in \left[\bar{a}, +\infty\right) \end{cases} \tag{5}$$

  with $\bar{a} = \sqrt{2/a}$ being the point of transition from quadratic to constant. 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 penalty in (5) is defined as $\phi(t; a) = |t|$.

An oscillatory component n, which belongs to the negative Sobolev space $\mathcal{H}^{-1}(\Omega)$, presents minimal $\|\cdot\|_{\mathcal{H}^{-1}}$ norm defined as

$$\|n\|_{\mathcal{H}^{-1}} = \inf\left\{ \sqrt{\sum_i |g_i|^2} \quad \Big| \; n = D^T g \right\} \approx \|g\|_2, \tag{6}$$

where D is defined in (3) and D^T is the adjoint operator.

2.1.1 ADMM-based Solution of Stage 1

To proceed with the numerical solution for (2) we substitute n (2) with $D^T g$ from (6), and we resort to the variable splitting technique to deal with the non-differentiability of the non-convex penalty term $\phi(\cdot; a)$. By introducing the auxiliary variable $t := Dv \in \mathbb{R}^N$, we formulate the following constrained optimization problem

$$\{v^*, w^*, g^*, t\} \;\leftarrow\; \arg\min_{v,w,g,t} \mathcal{J}(v, w, g, t), \quad \text{s.t.} \quad t = Dv,$$

$$\mathcal{J}(v, w, g, t) = \gamma_1 \sum_{j=1}^{N} \phi\left(\|t_j\|_2; a\right) + \frac{\gamma_2}{2}\|Hw\|_2^2 + \frac{\gamma_3}{2}\|g\|_2^4 + \frac{1}{2}\left\|f - (v + w + D^T g)\right\|_2^2. \tag{7}$$

We present the details of the ADMM-based iterative algorithm for the minimization of (7).

The corresponding augmented Lagrangian functional for the optimization problem reads as

$$\begin{aligned}\mathcal{L}(v, w, g, t; \rho) =& \gamma_1 \sum_{j=1}^{N} \phi\left(\|t_j\|_2; a\right) + \frac{\gamma_2}{2}\|Hw\|_2^2 + \frac{\gamma_3}{2}\|g\|_2^4 \\ & - \langle \rho, t - Dv\rangle + \frac{\beta}{2}\|t - Dv\|_2^2 + \frac{1}{2}\left\|f - (v + w + D^T g)\right\|_2^2,\end{aligned} \tag{8}$$

where $\beta > 0$ is a penalty scalar parameter, and $\rho \in \mathbb{R}^N$ represents the vector of Lagrange multipliers associated with the linear constraint $t = Dv$.

To simplify notations, in the following we denote by $x := (v^T, w^T, g^T)^T$ the $(3N)$-dimensional column vector formed by stacking the three optimization variables v, w, $g \in \mathbb{R}^N$. We then consider the following saddle-point problem:

$$\begin{aligned} &\text{Find} && (x^*, t^*, \rho^*) \;\in\; \mathbb{R}^{3N} \times \mathbb{R}^N \times \mathbb{R}^N \\ &\text{s.t.} && \mathcal{L}(x^*, t^*; \rho) \;\le\; \mathcal{L}\left(x^*, t^*; \rho^*\right) \;\le\; \mathcal{L}\left(x, t; \rho^*\right), \quad \forall (x, t, \rho) \;\in\; \mathbb{R}^{3N} \times \mathbb{R}^N \times \mathbb{R}^N. \end{aligned} \tag{9}$$

An ADMM-based iterative scheme is applied to approximate the solution of the saddle-point problem (8)–(9). Having zero-initialized vectors $t^{(0)}$ and $\rho^{(0)}$, the k-th iteration of the proposed alternating iterative scheme reads as follows

$$x^{(k+1)} = \operatorname*{argmin}_{x \in \mathbb{R}^{3N}} \mathcal{L}(x, t^{(k)}; \rho^{(k)}), \tag{10}$$

$$t^{(k+1)} = \operatorname*{argmin}_{t \in \mathbb{R}^{N}} \mathcal{L}(x^{(k+1)}, t; \rho^{(k)}), \tag{11}$$

$$\rho^{(k+1)} = \rho^{(k)} - \beta(t^{(k+1)} - Dv^{(k+1)}). \tag{12}$$

For the **x-subproblem** (10), the first-order optimality conditions read as

$$\begin{cases} (v^{(k+1)} + w^{(k+1)} + D^T g^{(k+1)} - f) + D^T \rho^{(k)} - \beta D^T (t^{(k)} - Dv^{(k)}) & = 0 \\ (w^{(k+1)} + v^{(k+1)} + D^T g^{(k+1)} - f) + \gamma_2 H^T H w^{(k+1)} & = 0 \\ D(D^T g^{(k+1)} + v^{(k+1)} + w^{(k+1)} - f) + 2\gamma_3 \left\|g^{(k+1)}\right\|_2^2 g^{(k+1)} & = 0 \end{cases}. \tag{13}$$

By replacing the nonlinear term $\left\|g^{(k+1)}\right\|_2^2$ in the third equation with the value at the previous iteration k, (13) reduces to the following linear system of equations

$$Lx^{(k+1)} = y, \tag{14}$$

where

$$L = \begin{pmatrix} I + \beta D^T D & I & D^T \\ I & I + \gamma_2 H^T H & D^T \\ D & D & DD^T + 2\gamma_3 \left\|g^{(k)}\right\|_2^2 I \end{pmatrix}, \; y = \begin{pmatrix} f + \beta D^T (t^{(k)} - \frac{1}{\beta}\rho^{(k)}) \\ f \\ Df \end{pmatrix}, \tag{15}$$

which is solved for $x^{(k+1)} = ((v^{(k+1)})^T, (w^{(k+1)})^T, (g^{(k+1)})^T)^T$. The block in L containing the discretized operator $H^T H$ slightly worsens the conditioning of the linear system. A suitable approximate solution of (14) is determined by solving the following regularized system of equations

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

with a small scalar parameter $\kappa > 0$, that allows the system to be efficiently solved by applying an iterative preconditioned conjugate gradient linear solver.

The ***t*-subproblem** (11) can be written omitting the constant terms as

$$t^{(k+1)} = \underset{t \in \mathbb{R}^N}{\operatorname{argmin}} \left\{ \gamma_1 \sum_j \phi(|t_j|; a) - \langle \rho, t \rangle + \frac{\beta}{2} \|t - Dv\|_2^2 \right\}. \tag{17}$$

The minimization problem in (17), rewritten in component-wise form, is equivalent to the following N independent 1-dimensional problems of the form

$$t_j^{(k+1)} = \underset{t \in \mathbb{R}}{\operatorname{argmin}} \left\{ \frac{1}{\lambda} \phi(|t|; a) + \frac{1}{2} \|t - q_j\|_2^2 \right\}, \qquad j = 1, \ldots, N, \tag{18}$$

with $\lambda = \beta/\gamma_1$ and $q_j = (Dv^{(k)})_j + \rho_j^{(k)}/\beta$. Necessary and sufficient conditions for strong convexity of the cost functions in (18) are demonstrated in [6]. In particular, the problems in (18) are strongly convex if and only if the following condition holds

$$a < \lambda \implies \beta > a\gamma_1 \implies \beta = \tau a \gamma_1, \quad \text{for} \quad \tau \in \mathbb{R},\ \tau > 1. \tag{19}$$

Under the assumption (19), the unique solutions of problems in (18) can be obtained in closed form as

$$t_j^{(k+1)} = \min(\max(\nu - \zeta/|q_j|, 0), 1)\, q_j, \tag{20}$$

where $\nu = \dfrac{\lambda}{\lambda - a}$ and $\zeta = \dfrac{\sqrt{2a}}{\lambda - a}$.

We remark that the condition on β defined in (8) only ensures the convexity conditions (19) of t-subproblem (18), but does not guarantee convergence of the overall ADMM scheme. The proposed two-block ADMM-based minimization algorithm is summarized in Algorithm 1.

2.2 Stage 2: Residual Aided Refinement

Since the regularization parameters γ_1, γ_2, γ_3 strongly depend on the signal morphology, in practice the three resulting components from stage 1 may be slightly mixed: some trend can be found in $\bar{v}$, and some noise (highly oscillating behaviour) in the smoothed component $\bar{w}$. The trend contribution

Algorithm 1: Stage 1 Decomposition

input : f
output : $\bar{v}$, $\bar{w}$, $\bar{n}$ components
parameters: $\bar{a}$, γ_1, γ_2, γ_3, τ
Generate discrete operators D in (3) and H in (4).
$a = 2/\bar{a}^2$, $\beta = \tau a \gamma_1$
while $k < iter$ **and** $r > th$ **do**
  $x^{(k+1)} \leftarrow$ solve $(L + \kappa I)x^{(k+1)} = y$, using L, y in (15) — *subproblem for* $x = (v, w, g)$
  $t^{(k+1)} \leftarrow$ compute (20) — *subproblem for* t
  $\rho^{(k+1)} = \rho^{(k)} - \beta(t^{(k+1)} - Dv^{(k+1)})$ — *subproblem for* ρ
  Update L in (15)
  $r = \|x^{(k+1)} - x^{(k)}\|/\|x^{(k)}\|$
  $k = k + 1$
Get $\{\bar{v}, \bar{w}, g\}$ from x
$\bar{n} = D^T g$

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 $\bar{n}$, as is illustrated in the diagram scheme in Figure 1.

Therefore, the goal of this second stage is to refine the results $\bar{v}$, $\bar{w}$, $\bar{n}$ obtained from stage 1 to satisfy the specific application target when needed. We remark that stage 2 can be eventually not necessary in case the results from stage 1 are already of acceptable quality.

Refinement of $\bar{v}$. The refined v^* component is obtained from $\bar{v}$, solution of problem (7), by subtracting residual parts

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

where r_v is carried out by filling in $\bar{v}$ the intervals corresponding to the jumps with piece-wise constant samples (hence named *inpainting*). In particular, having a partitioning $C = \{C_i\}$ of $\bar{v}$ domain such that, in an alternating manner, the sub-intervals C_i, represent jump and non-jumps, under the assumption that C_1 is not a jump, r_v is defined as follows

$$r_v = \begin{cases} \bar{v}|_{C_i} & \text{if } i \text{ odd} \\ 0.5(mean(\bar{v})|_{C_{i-1}} + mean(\bar{v})|_{C_{i+1}}) & \text{if } i \text{ even.} \end{cases} \tag{21}$$

In this way, over odd-indexed C_i, i.e. away from the jump sub-intervals, $r_v = \bar{v}$, while inside jumps our goal is to obtain a piece-wise constant inpainting for the even-indexed C_i. This can be directly achieved by setting the r_v values to the constant mean value of the surrounding sub-intervals.

The following two steps describe in detail our approach to building the set C and r_v, and Figure 2 shows the main ingredients of our procedure for two exemplary samples (column-wise) extracted from $\bar{v}$ (first row).

1. Partition of $\bar{v}$ into sub-intervals $C = \{C_i\}$.
   We compute the second order derivative $z = D^T D\ \bar{v}$, reported in the second row of Figure 2, which has the following behavior: z is zero where the signal is flat, numerically $z \leq 10^{-4}$, while z zero-crosses the x-axis in correspondence of jump discontinuities.

   We then partition the domain of $\bar{v}$ into sub-intervals $\{\bar{C}_i\}$ bounded by the zero-crossings of signal z.

   Then, we analyse the behavior of $\bar{v}$ over $\bar{C}_i$ in order to generate C considering the following rules:

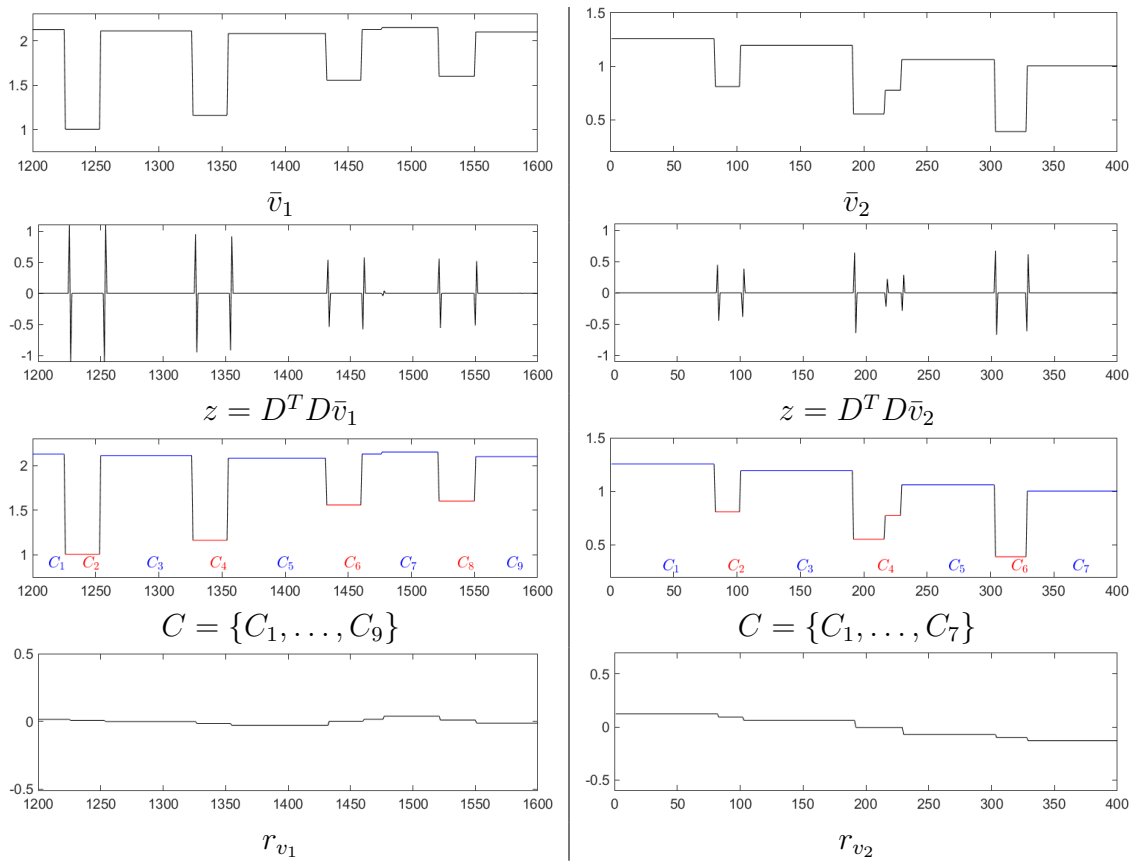

Figure 2: Details of the inpainting of r_v on two signals $\bar{v}_1$, $\bar{v}_2$. The sub-intervals C_2, C_4, C_6 and C_8 define four jumps (red), while C_7 (blue) contains negligible jump in $\bar{v}_1$ (left). The sub-intervals C_2, C_4, C_6, identify three jumps in the signal $\bar{v}_2$ where C_4 contains a composite jump (right).

- Two adjacent sub-intervals $\bar{C}_k$, $\bar{C}_{k+1}$ are unified into the sub-interval C_i if they satisfy

$$\left|mean(\bar{v})|_{\bar{C}_k} - mean(\bar{v})|_{\bar{C}_{k+1}}\right| < th_v, \tag{22}$$

  where $th_v > 0$ is the maximum variation in $\bar{v}$ allowed as negligible jump, as illustrated in Figure 2 (first column, third row) where $\bar{C}_k$ and $\bar{C}_{k+1}$ were joined into C_7.

- When $\bar{C}_k$, $\bar{C}_{k+1}$ do not satisfy (22), the jump may be composed by one or more sub-intervals. Therefore, starting from $\bar{C}_k$, which is assumed to be a non-jump sub-interval, we seek for the nearest non-jump sub-interval $\bar{C}_{k+j}$ that is

$$\left|mean(\bar{v})|_{\bar{C}_k} - mean(\bar{v})|_{\bar{C}_{k+j}}\right| < th_v. \tag{23}$$

  This indicates that the in-between sub-intervals $\bar{C}_{k+1}, \ldots, \bar{C}_{k+j-1}$ can be joined into a single sub-interval. This situation is illustrated in Figure 2 (second column, third row) where $\bar{C}_4, \bar{C}_5$ were joined into C_4.

Thus, we finally get a sequence of sub-intervals $C = \{C_i\}$ where, conventionally, the first sub-interval C_1 does not belong to a jump, and the sub-intervals with even index represent jumps.

2. Setup r_v
   The values of r_v are obtained according to (21) and the $mean(r_v)$ is subtracted to not alter $mean(v^*)$ w.r.t. $\bar{v}$ as illustrated in Figure 2 (last row).

Refinement of $\bar{w}$. The spurious trend eventually included in r_v in (21) should be added to $\bar{w}$. The final component w^* is then computed by smoothing the signal $\bar{w} + r_v$, which may contain spurious oscillatory residual, via solving the following optimization problem

$$w^* = \arg\min_{w \in \mathbb{R}^N} \frac{1}{2}\|w - (\bar{w} + r_v)\|_2^2 + \alpha \|Dw\|_2^2, \tag{24}$$

where the regularization parameter $\alpha > 0$ controls the level of smoothness of w^*. The smoothing quadratic minimization problem (24) is a convex optimization problem whose unique minimizer is explicitly given by imposing the first optimality conditions which lead to the solution of the following linear system of equations

$$(I_N + \alpha D^T D)w = \bar{w} + r_v, \tag{25}$$

with symmetric positive definite coefficient matrix.

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 and the residual from stage 1 $\bar{r}$

$$n^* = \bar{n} + r_w + \bar{r}, \tag{26}$$

where $\bar{r} = f - \bar{v} - \bar{w} - \bar{n}$, and r_w is given by

$$r_w = (\bar{w} + r_v) - w^*. \tag{27}$$

In case the signal f contains additional components besides the three considered here, the residual $\bar{r}$ should not be added in (26).

Algorithm 2 summarizes the main steps in stage 2. The user can freely decide according to the application if stage 2 needs to be performed entirely or only for a selected set of components; this is controlled via flag variables in the implementation.

Algorithm 2: Stage 2: Refinement

input : $\bar{v}$, $\bar{w}$, $\bar{n}$ components, $\bar{r} = f - \bar{v} - \bar{w} - \bar{n}$ residual
output : v^*, w^*, n^* components
parameters: th_v, α
Generate discrete operator D in (3)
Solve inpainting problem for v^*
- $z = D^T D\, \bar{v}$
- Compute the set $\{\bar{C}_i\}$ from z
- Generate the set $\{C_i\}$ from $\{\bar{C}_i\}$ and $\bar{v}$
- set r_v according to (21)
- $v^* = \bar{v} - r_v$

Solve smoothing problem for w^*
- Solve (25) for w^*
- $r_w = \bar{w} + r_v - w^*$

Refinement for n^*
- $n^* = \bar{n} + r_w + \bar{r}$

2.3 Parameter Selection

In this section we describe in detail the strategies used for parameter selection assuming the range of f to be similar to the examples presented. A general rule-of-thumb for fine-tuning the parameters is provided with examples in Section 3.1.

The fixed parameters of stage 1 are:

- $iter = 1500$, the maximum number of ADMM iterations as stopping criterion,
- $th = 10^{-6}$, the solution relative change as stopping criterion,
- $\kappa = 10^{-7}$, the scalar parameter in (16). In case the linear system (16) becomes unstable, the user is welcome to increase its value,

while the parameters that the user can modify are

- γ_1, γ_2, $\gamma_3 > 0$, regularization parameters,
- $\tau > 1$ in (19) related to the β parameter proportionally (in the range $1.1 - 50$, default 10),
- $\bar{a}$, the expected minimal jump height (default 0.3).

In order for the contributions of each regularization term to be balanced, one should aim to set γ_1, γ_2, γ_3 approximately to the inverse of their respective penalty term values such that to balance the contribution of the three terms. In the following, we provide a description of how to estimate these values.

The value $\bar{a}$ indicates the minimal height value above which every jump $|(Dv)_i| > \bar{a}$ becomes penalized equally, therefore, we propose the user to set $\bar{a}$ as the expected minimal jump height, eventually observed from the input data, from which we estimate a automatically as $a = 2/\bar{a}^2$.

Consequently, if every jump in v identified as non-zero derivative $|(Dv)_i| > 0$ also satisfies $|(Dv)_i| > \bar{a}$, the resulting value of the v-regularization term will become the number of non-zero derivatives in v, i.e. the number of jumps. Therefore, we propose to set γ_1 to

$$\gamma_1 = 1/\text{number of expected jumps.}$$

The w-regularization term is represented by squared ℓ_2 norm of second-order derivatives Hw of a relatively smooth component. In general the squared ℓ_2 norm is a small number, therefore the γ_2 value needs to be relatively high. We propose $\gamma_2 > 10$.

Differently from the w penalty, for the g-regularization term, its penalty values grow more rapidly with N, therefore we suggest using small values of γ_3, i.e. of order at most 10^{-1}.

Table 1 lists the optimal parameter values for the signals presented in Section 3. Finally, we note a successful decomposition is not driven by the precision of values γ_1, γ_2, and γ_3 in the model (7), but rather by mutual ratios between the parameters themselves.

A general rule of thumb for adjusting the parameters based on the visual inspection of the decomposition obtained is described in Section 3.1.

The parameters of stage 2 are:

- $th_v > 0$ threshold for negligible jump in $\bar{v}$.
- $\alpha > 0$ smoothness regularization parameter for w^* (in the range $(0, 1]$).

Stage 2 performs the alignment of jumps in $\bar{v}$, the smoothing of $\bar{w} + r_v$ and adds the remaining residual to $\bar{n}$.

The threshold th_v is the maximum variation in $\bar{v}$ allowed between the mean $\bar{v}$ values at sub-intervals adjacent to a jump. In case the value th_v is smaller than a possible negligible jump, the effect on v^* is that not all jumps would be aligned. Otherwise, if the value th_v is greater than an edge forming the jump, these may be then included in r_v and consequently worsen each v^*, w^* and n^* w.r.t. $\bar{v}$, $\bar{w}$, $\bar{n}$ of stage 1.

The parameter α forces the smoothness of w^* and its value, although normalized to signal length, depends on how much the user wants to further smooth the component $\bar{w}$. The higher α, the smoother resulting w^* component. In case the α value is too large, the refinement of $\bar{w}$ may produce the so-called *over-smoothed(-regularized)* component w^*, and consequently worsening n^*.

3 Numerical Experiments

We present examples that showcase the proposed JOT model when applied to the study of non-stationary signals.

In Table 1 we report the optimal values of the parameters used in the experiments.

	f_1 ($N = 1024$)	f_2 ($N = 4451$)	f_3 ($N = 4096$)	f_4 ($N = 20000$)	f_5 ($N = 2051$)
$\bar{a}$	0.20	0.20	0.45	0.35	0.35
γ_1^*	5e-2	5e-2	2e-1	1e-2	2.5e-2
γ_2^*	1e+3	2.5e+1	1e+4	5e+5	5e+1
γ_3^*	5e-2	1e-2	5e-2	5e-8	5e-5
τ	10	10	50	50	10
th_v	0.25	0.25	0.25	-	0.25
α	0.4	1e-4	0.4	-	1e-4

Table 1: Parameter values used in experiments. Stage 1 first block, stage 2 second block.

In Table 2, we report the signal to noise ratio (SNR) values of the resulting components for the synthetic input signals f_1, f_2 and f_3 after both stage 1 and stage 2, computed as

$$SNR(x, y) = 10 \log_{10} \frac{\|y - mean(y)\|_2^2}{\|y - x\|_2^2},$$

for a proper comparison of the resulting components with the ground truth; due to the equivalency of solutions $\mathcal{J}(v, w, n) = \mathcal{J}(v + t, w - t, n)$, we shift the components (v, w) by a constant t, such that the mean values of $\bar{v}$ and v^* correspond to the mean of the ground truth component v.

	$SNR(\bar{v}, v)$	$SNR(\bar{w}, w)$	$SNR(\bar{n}, n)$	$SNR(v^*, v)$	$SNR(w^*, w)$	$SNR(n^*, n)$
f_1	8.57	6.66	5.63	25.19	30.06	14.01
f_2	15.45	13.61	11.25	29.37	27.78	21.44
f_3	9.54	14.77	2.58	30.67	39.23	16.75

Table 2: Signal to noise ratio for f_1, f_2 and f_3 components after stage 1 and stage 2 w.r.t. ground truth components.

In Figures 3–7 we report in the first row the input signal f, in the second row the resulting components $\bar{v}$, $\bar{w}$, $\bar{n}$ of stage 1 and in the third row (if present) the refined components v^*, w^*, n^* of stage 2 in red color. Over-posed in black color the ground-truth components when known.

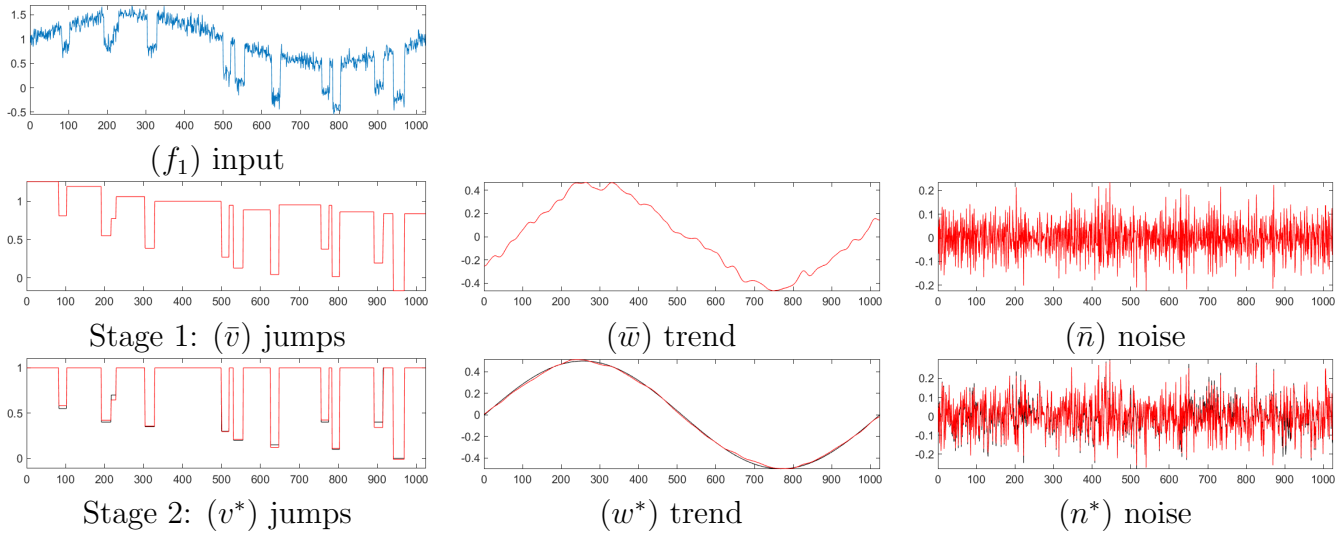

Figure 3: Decomposition results of the JOT method: (first row) the given signal f_1; (second row) the results of stage 1: jumps $\bar{v}$, the trend $\bar{w}$ and noise $\bar{n}$ are separated from the given signal f_1; (third row) stage 2: the components after refinement over-imposed with the ground truth in black.

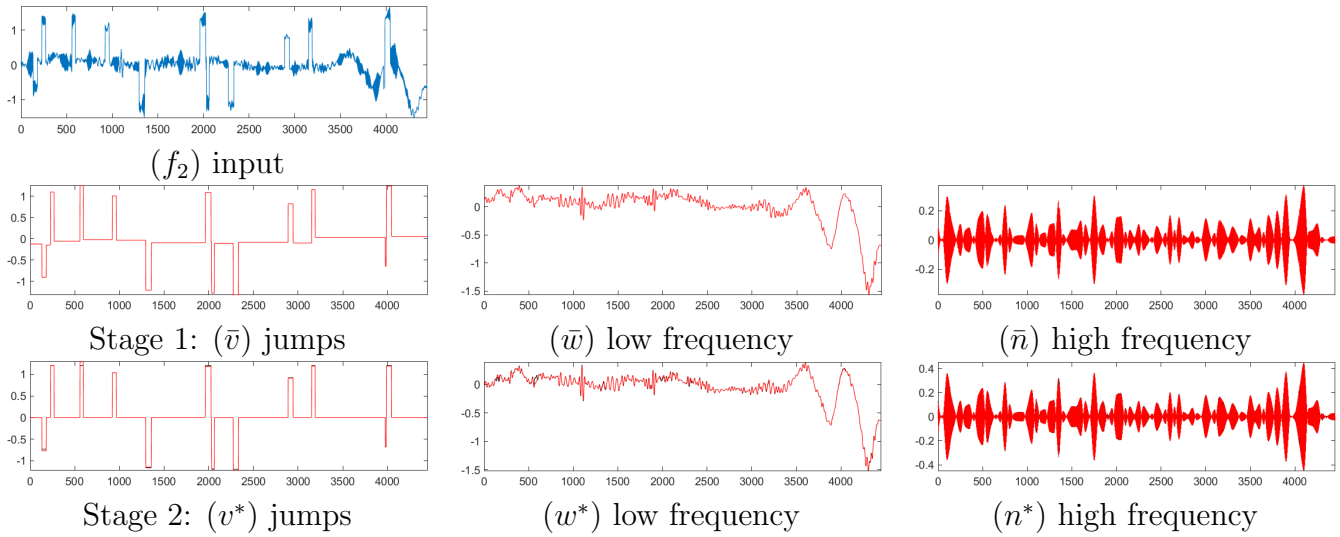

Figure 4: Decomposition results of the JOT method: (first row) the given signal f_2; (second row) the results of stage 1: jumps $\bar{v}$, the trend $\bar{w}$ and noise $\bar{n}$ are separated from the given signal f_2; (third row) stage 2: the components after refinement over-imposed with the ground truth in black.

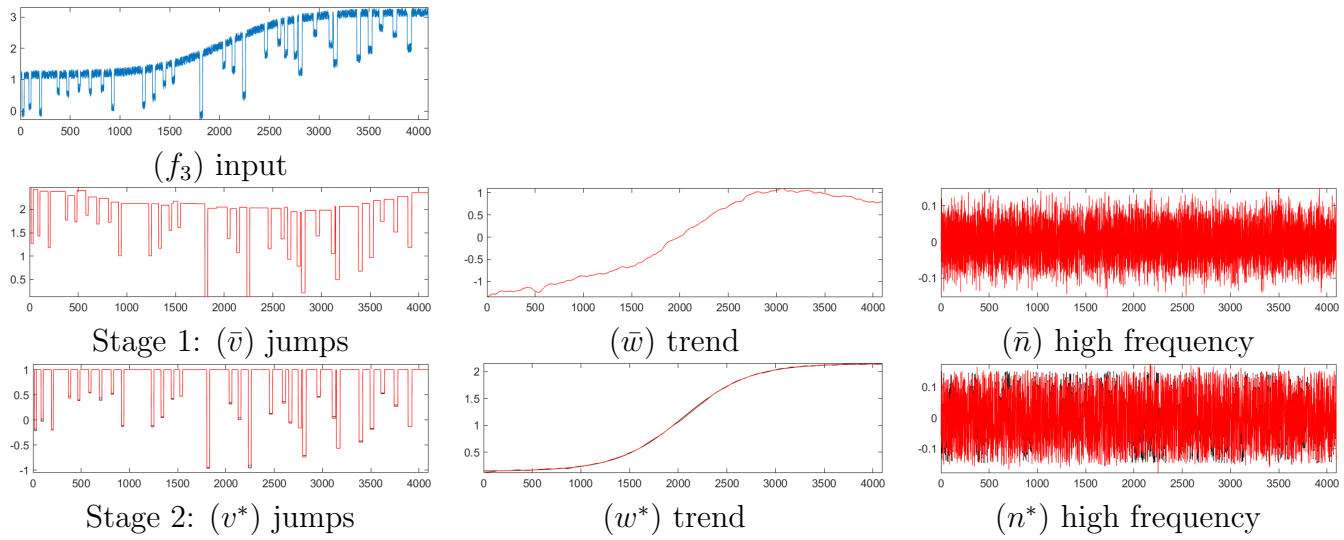

Figure 5: Decomposition results of the JOT method: (first row) the given signal f_3; (second row) the results of stage 1: jumps $\bar{v}$, the trend $\bar{w}$ and noise $\bar{n}$ are separated from the given signal f_3; (third row) stage 2: the components after refinement over-imposed with the ground truth in black.

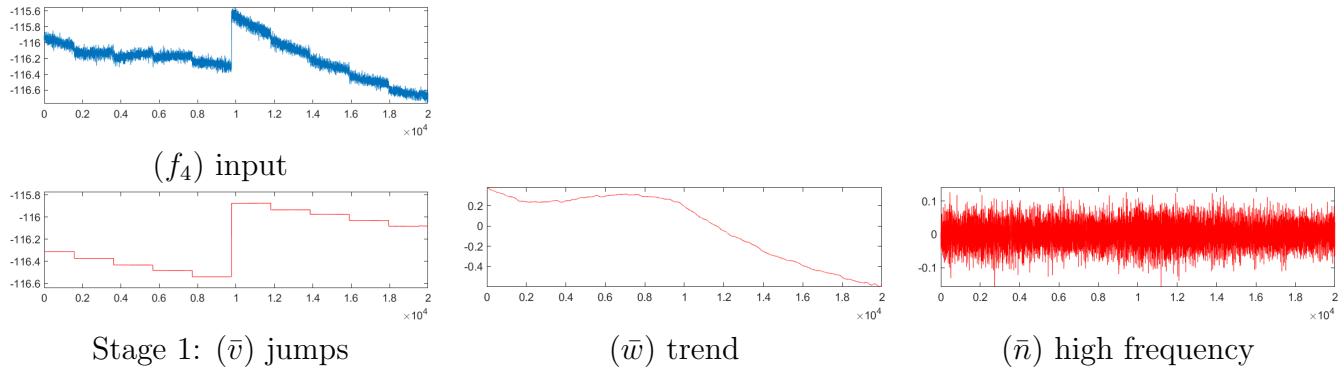

Figure 6: Decomposition results of the JOT method: (first row) the given signal f_4; (second row) the results of stage 1: jumps $\bar{v}$, the trend $\bar{w}$ and high frequency $\bar{n}$ are separated from the given signal f_4.

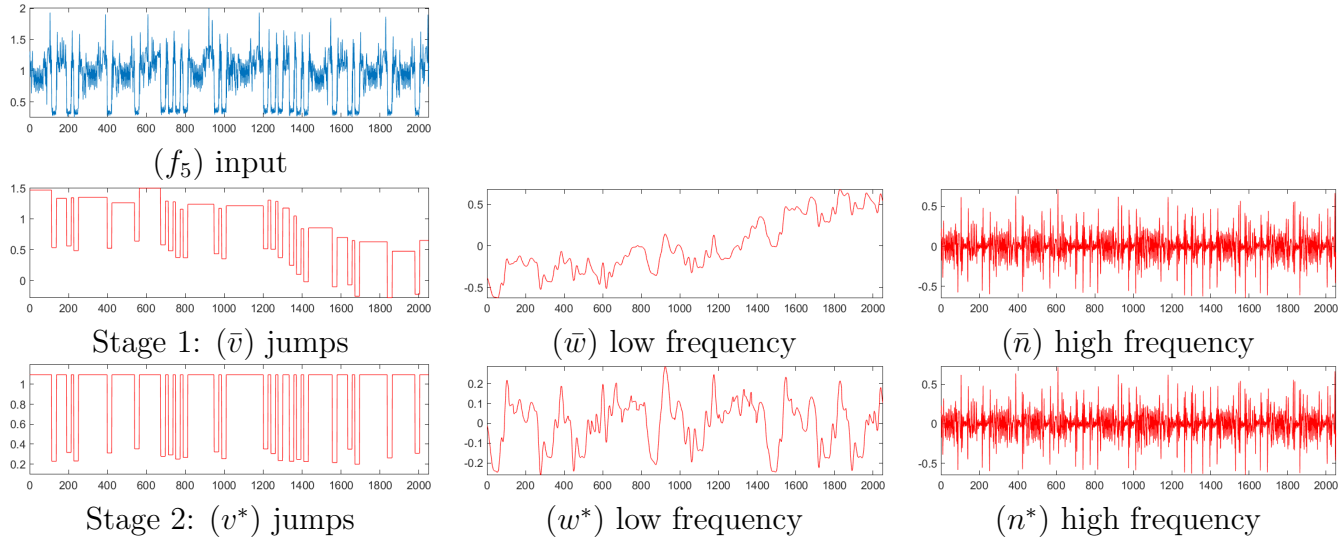

Figure 7: Decomposition results of the JOT method: (first row) the given signal f_5; (second row) the results of stage 1: jumps $\bar{v}$, the trend $\bar{w}$ and high frequency $\bar{n}$ are separated from the given signal f_5; (third row) stage 2: the components after refinement.

3.1 Effect of Penalty Parameters γ_1, γ_2, γ_3

In this section we revisit the parameter selection. In particular, we discuss the possible decomposition results in case the values of the parameters γ_1, γ_2 and γ_3 are either too large or too small w.r.t. the optimal values reported in Table 1.

The effects discussed in the rest of this section are illustrated for the f_3 input signal, Figure 5 first row. The visual effect that we expect is that, for each component $\bar{v}$, $\bar{w}$, $\bar{n}$, if the corresponding regularization parameter γ_1, γ_2, γ_3 is too large, residuals that are pushed away from one component can appear in one or both remaining components. On the other hand, if the parameter value is too small, the corresponding component can absorb parts of the other two components.

As illustrated in Figure 8 first row, for too large values of γ_1, many piece-wise constant jumps are lost, since they are pushed into the $\bar{n}$ component. The second row of Figure 8 shows the effect of γ_1 too small, the corresponding energy term is too low and some oscillations from $\bar{n}$ were absorbed as piece-wise constant jumps into $\bar{v}$.

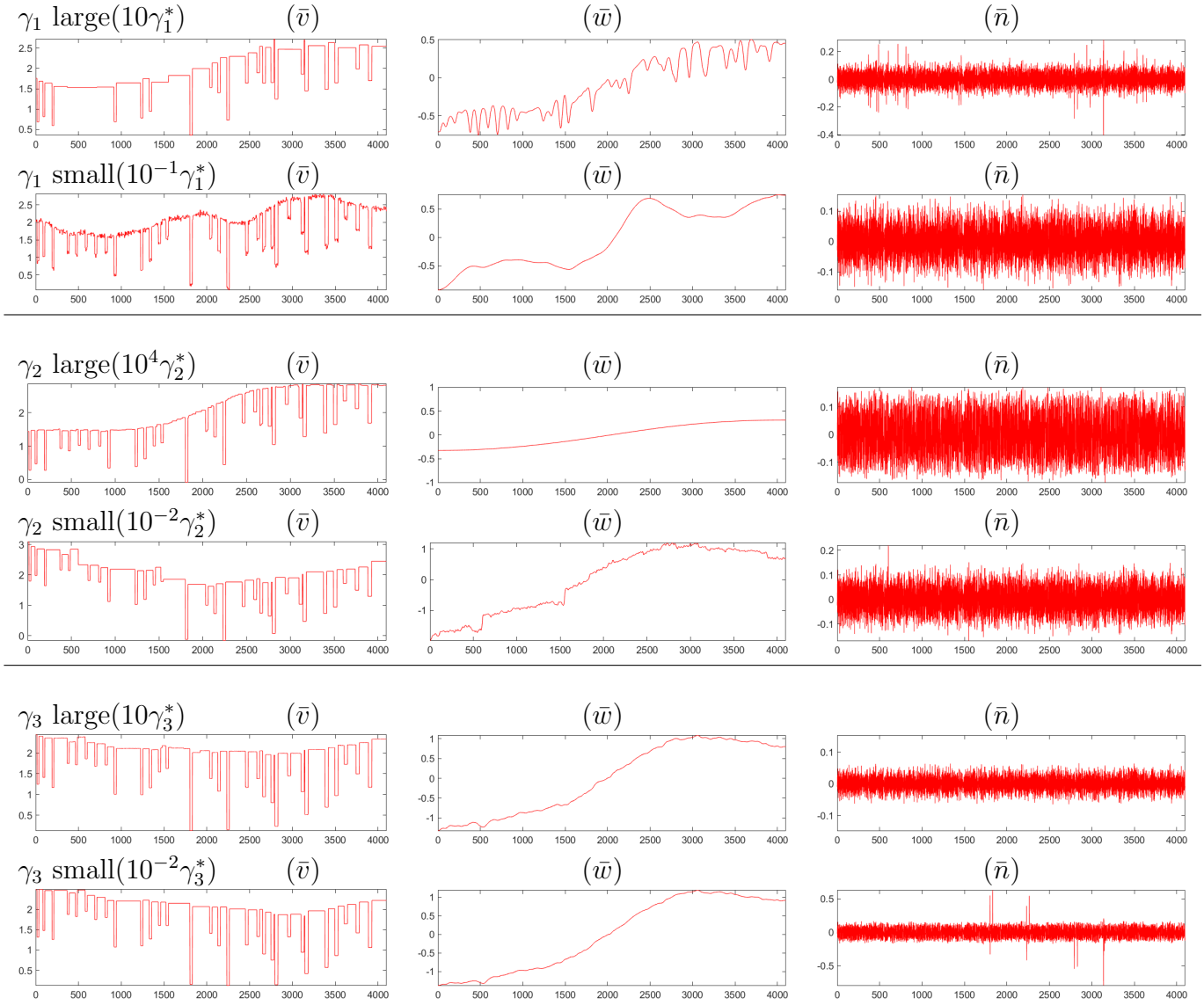

Figure 8: Effect of the parameter values on the decomposition results. From top to bottom: decomposition results in case the values of γ_1, γ_2, and γ_3 are too high and too low, respectively, with respect to the optimal values reported in Table 1.

The effect of the γ_2 value is reported in the third and the fourth rows of Figure 8. Large values of γ_2 force the minimization of the second order derivatives in $\bar{w}$ and tend to push the smooth trend

into $\bar{v}$ as a stair-case effect. On the other hand, small values of γ_2 can absorb parts of the piece-wise constant jumps as well as oscillations from $\bar{n}$.

In case the γ_3 value is too large, the overall variance of data captured in $\bar{n}$ is decreased, see $\bar{n}$ amplitudes in Figure 8, fifth row, compared to the amplitudes in Figure 5, left column. The residual from $\bar{n}$ is captured by $\bar{w}$ in this case. On the other hand, for small values of γ_3, Figure 8 sixth row, the highest-amplitude piece-wise constant jumps from $\bar{v}$ are captured in $\bar{n}$.

We notice that when a component x absorbs parts of a component y, the situation can be caused either by γ_x being too small or γ_y being too large. The decision of which parameter to tune should be done w.r.t. to the third component. For example, $\bar{v}$ in the second row of Figure 8 absorbed parts of $\bar{w}$, thus either γ_1 is too small or γ_2 too large. However, here $\bar{v}$ absorbed part of $\bar{n}$ as well, thus it would be more sensible to increase γ_1 value rather than decreasing γ_2.

References

[1] P. Barbe, A. Cicone, W. Suet L., and H. Zhou, *Time-Frequency Representation of Non-stationary Signals: the IMFogram*, Pure and Applied Functional Analysis, 7 (2022), pp. 27–39. http://yokohamapublishers.jp/online2/oppafa/vol7/p27.html.

[2] L. Christiano and T. J. Fitzgerald, *The Band Pass Filter*, International Economic Review, 44 (2003), pp. 435–465, https://doi.org/10.1111/1468-2354.t01-1-00076.

[3] A. Cicone, M. Huska, S.-H. Kang, and S. Morigi, *JOT: A Variational Signal Decomposition Into Jump, Oscillation and Trend*, IEEE Transactions on Signal Processing, 70 (2022), pp. 772–784, https://doi.org/10.1109/TSP.2022.3145665.

[4] I. Daubechies, J. Lu, and H.-T. Wu, *Synchrosqueezed Wavelet Transforms: An Empirical Mode Decomposition-Like Tool*, Applied and Computational Harmonic Analysis, 30 (2011), pp. 243–261, https://doi.org/10.1016/j.acha.2010.08.002.

[5] 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, 14 (2021), pp. 1749–1789. https://doi.org/10.1137/20M1355987.

[6] M. Huska, A. Lanza, S. Morigi, and I. Selesnick, *A Convex-Nonconvex Variational Method for the Additive Decomposition of Functions on Surfaces*, Inverse Problems, 35 (2019), p. 124008, https://doi.org/10.1088/1361-6420/ab2d44.

[7] Y. Meyer and D. Lewis, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*, Memoirs of the American Mathematical Society, American Mathematical Society, 2001. ISBN 9780821829202.