Empirical Mode Decomposition revisited by multicomponent non smooth convex optimization [1]

Nelly Pustelnik, Pierre Borgnat, and Patrick Flandrin *(EURASIP Fellow)*

Laboratoire de Physique, ENS de Lyon, CNRS, Université Claude Bernard Lyon I, Université de Lyon, UMR CNRS 5672, F-69364 Lyon cedex 7, France
Corresponding author e-mail : nelly.pustelnik@ens-lyon.fr

Abstract

This work deals with the decomposition of a signal into a collection of intrinsic mode functions. More specifically, we aim to revisit Empirical Mode Decomposition (EMD) based on a sifting process step, which highly depends on the choice of an interpolation method, the number of inner iterations, and that does not have any convergence guarantees. The proposed alternative to the sifting process is based on non-smooth convex optimization allowing to integrate flexible in the criterion we aim to minimize. We discuss the choice of the criterion, we describe the proposed algorithm and its convergence guarantees, we propose an extension to deal with multivariate signals, and we figure out the effectiveness of the proposed method compared to the state-of-the-art.

Keywords: Empirical Mode Decomposition (EMD); Intrinsic Mode Functions (IMF); Trend-fluctuation; Convex optimization; Proximal algorithms; ℓ_p-norm.

[1] Part of this work appeared in the conference proceedings of EUSIPCO 2012 [1].

1. Introduction

Numerous signals stemming from natural phenomena and/or man-made systems (e.g., in meteorology, oceanography, biology, or energy networks, to name but a few), exhibit non-linear and/or non-stationary behaviours [2, 3]. One useful class of models aimed at describing such signals can be expressed as a sum of AM-FM signals

$$(\forall t \in \mathbb{R}) \quad x(t) = \sum_{k=1}^{K} \alpha_k(t) \cos \theta_k(t) \tag{1.1}$$

where $\alpha_k(t)$ is assumed to be non-negative and smoother than $\cos \theta_k(t)$, and $\omega_k(t) = \frac{d}{dt}\theta_k(t) \geq 0$. The limit case where $\omega_K(t) = 0$ permits to account for a trend in the data represented by $\alpha_k(t)$.

According to such a model, the challenge is twofold: on the one hand, extract each component $x_k(t) = \alpha_k(t) \cos \theta_k(t)$ from $x(t)$ and, on the other hand, for every $k \in \{1, \ldots, K\}$, evaluate the instantaneous frequency $\omega_k(t)$ and the instantaneous amplitude $\alpha_k(t)$. A toy example is provided in Figure 1. Instantaneous frequency and amplitude of a mono-component AM-FM signal (i.e., $K = 1$) can be defined through the concept of analytic signal introduced by Ville [4]. When $K > 1$, an extraction of $(\alpha_k)_{1 \leq k \leq K}$ and $(\omega_k)_{1 \leq k \leq K}$ based on the analytic signal is not any more feasible. Consequently, during the last decades, numerous solutions have been investigated such as zero crossing methods [5, 6], reassignment methods [7], synchrosqueezing [8], or Empirical Mode Decomposition (EMD) [9]. These methods can be split in two classes:

1. Methods that evaluate instantaneous frequencies and amplitudes without extracting the components $(x_k)_{1 \leq k \leq K}$ from x.

2. Methods that first extract the components $(x_k)_{1\leq k\leq K}$ from x and then evaluate instantaneous frequency and amplitude from each x_k.

Synchrosqueezing and reassignment belong to the first class of methods and EMD to the second one. Reassignment aims at sharpening a time-frequency representation by allocating each value to a different point in the time-frequency plane according to its local behaviour. The synchrosqueezing proposed by Daubechies *et al.* [8] appears to be a particular case of reassignment that operates along the frequency direction only.[2] This restriction allows to reconstruct each component individually if the components are locally well-separated in frequency. However, in order to extract the components, their number and the definition of their domains is required and it needs to be supervised [11, 12].

From a different perspective, EMD extracts iteratively the intrinsic mode functions of a signal. Instantaneous frequency and amplitude can then be obtained directly from each mode as well as correlations between modes or other kinds of mode analysis. Nonetheless, this technique faces the difficulty of having no mathematical definition besides its algorithm and thus no convergence properties. The purpose of this work is to propose an efficient alternative to EMD based on a variational approach and convex optimization methods which allows us to have convergence guarantees, to be flexible regarding the signal to be analysed, to deal with signals having zero-valued segments without creating artificial oscillations, and to be robust to sampling effects.

Section 2 aims at recalling the EMD principle, together with alterna-

[2]The synchrosqueezing theory has been developed initially through the wavelet formalism, but it can be formulated as well for the Short-Time Fourier Transform [10].

tives that have been proposed in the literature. The proposed variational approach is detailed in Section 3. Section 4 evaluates the performance of the proposed solution on simulated and real data.

2. State of the art

2.1. Empirical Mode Decomposition

We consider a signal $x = (x[n])_{1 \leq n \leq N} \in \mathbb{R}^N$ of length N which may be written as

$$x = \sum_{k=1}^{K} d_k + a_K, \tag{2.1}$$

where, for every $k \in \{1, \ldots, K\}$, $d_k \in \mathbb{R}^N$ denotes the intrinsic mode function (IMF) of order k, and $a_K \in \mathbb{R}^N$ denotes the trend of order K. To this end, an average envelope is defined as the mean of an upper envelope and a lower envelope, both constructed (*via* a cubic splines interpolation) from local maxima and minima, respectively. The IMF characterization imposes this average envelope to be almost zero everywhere.

EMD is a procedure in K steps allowing to find $(d_k)_{1 \leq k \leq K}$ and a_K from x. At each step, the trend and the IMF of order $k \geq 1$, respectively denoted by a_k and d_k, are extracted from the trend of order $k-1$ (note that $a_0 = x$). This decomposition stage is known as the *sifting process* and it consists in:

i. set $i = 0$ and initialize a temporary variable $s^{(i)} = a_{k-1}$,

ii. identify all extrema of $s^{(i)}$,

iii. interpolate between minima (resp. maxima) ending up with some envelope $e_{\min}$ (resp. $e_{\max}$),

iv. compute the mean envelope $m = \frac{e_{\min} + e_{\max}}{2}$,

v. extract the residual $s^{(i+1)} = s^{(i)} - m$ and increment i,

vi. iterate Steps (ii)-(v) until the residual $s^{(i)}$ achieves a zero mean envelope,

vii. let the IMF of order k be $d_k = s^{(i)}$ and the trend of order k be $a_k = a_{k-1} - d_k$.

This adaptive approach proved its efficiency for analysing signals through numerous applications (see [3] and references therein). However, the result of this method is highly dependent on the interpolation process in Step iii., it has also been pointed out to be sensitive to sampling effects [13], and this method is not efficient to deal with signals having null segments. Worst of all, no proof of convergence can be established for this technique.

Before detailing the solution we propose, we will first briefly review recent EMD variants which focus on a modal decomposition with convergence guarantees that rely on convex optimization tools.

2.2. Modal decomposition with convex optimization tools

A first intuitive variation on EMD introducing convex optimization is the work by Huang and Kunoth [14], aiming at modifying the way the envelopes are computed. The goal is to avoid intersection between the upper (resp. lower) envelope and the signal that leads to an over (resp. under) shooting effect in Step iii. of the EMD. This method replaces the interpolation step with an estimation of the upper (resp. lower) envelope based on an optimization scheme that can be solved with classical convex optimization tools such as interior point method or projected gradient. In [15], the estimation of the lower and upper envelopes is based on parabolic partial differential equation (PDE). The resulting envelopes are piecewise cubic polynomial curves interpolating the maxima (resp. minima) of the signal. In [15], the authors have also proposed a PDE based method allowing to directly estimate the mean envelope through the inflection points. In [16], a

genetic algorithm is proposed in order to efficiently estimate the envelopes. It is also interesting to refer to the work by Hawley et al. [17] where the cubic spline interpolation is replaced by a trigonometric interpolation. If this interpolation method is less accurate than a cubic one it has the advantage of proposing a theoretical framework from which it is possible to derive convergence guarantees of the mean envelope to zero.

Another interesting approach was proposed by Meignen and Perrier in [18, 19]. The authors have replaced the sifting process by a convex optimization procedure. More precisely, the authors extract the trend of order k by solving the following minimization problem:

$$a_k = \underset{a \in \mathbb{R}^N}{\arg\min} \, \|a\|_2 \quad \text{subj. to} \quad a \in \Pi \cap C_{a_{k-1}} \tag{2.2}$$

where Π denotes the space of cubic spline functions, and $C_{a_{k-1}}$ denotes a constraint onto the dynamic range of the mean envelope at the location of the extrema of a_{k-1}. To be more specific, we have to specify that the variable to be optimized in (2.2) is not the trend but the coefficients associated to the Hermite interpolant of a. In [19], the dynamic range constraint is replaced by a constraint which imposes the symmetry of the upper and lower envelopes of $a_{k-1} - a$. The main limitation of this approach is that a first approximation $\overline{a}$ of a_k is required in order to involve a convex constraint. This second approach also looks for the mean envelope in the space of spline functions. The IMF of order k is then deduced from a_k and a_{k-1} (i.e., $d_k = a_{k-1} - a_k$). This method will be named OMD (for Optimization based Mode Decomposition) in the following.

More recently a dictionary learning procedure combined to sparse optimization has been proposed for modal decomposition [20, 22]. This method

aims at solving the following minimization problem:

$$\text{minimize} \quad K \qquad \text{subj. to} \quad \begin{cases} x = \sum_{k=1}^{K} \alpha_k \cos\theta_k, \\ (\forall k \in \{1, \ldots, K\}) \quad \alpha_k \cos\theta_k \in \mathcal{D}, \end{cases}$$

where $\mathcal{D}$ is a highly redundant dictionary requiring α_k to be smoother than $\cos\theta_k$ and the derivative of θ_k (instantaneous frequency) to be positive (i.e., $\theta_k' \geq 0$). The first solution, proposed by Hou and Shi [20], deals with the decomposition of x into its local median α_0 and an IMF $\alpha_1 \cos\theta_1$ considering a variational approach based on the total variation (TV) [21]. This minimization problem can be solved by a Gauss-Newton technique considering an additive variable. The resulting algorithm involves to update alternatively (α_0, α_1) and the phase function θ_1 that highly depends on the initialization and thus, in general, does not have any convergence guarantees to a global minimizer. The second solution, proposed by Hou and Shi [22], appears to be an extension of the previous one to K components and to be more robust to noise. The resolution of the above minimization problem is similar to the previous formulation. In [23], a convergence proof has been proposed under restrictive assumptions.

3. Multicomponent variational problem

The sifting process aims at extracting a trend and a fluctuation of order k or, in other words, elementary structures of the signal. This problem appears to be highly related to texture-geometry decomposition methods developed in the image processing field. In the context of denoising, this decomposition has been achieved in [21] with a total variation penalty in order to extract a piecewise smooth component. In [24, 25], a criterion combining total variation and G-norm, modelling strong oscillations, has been considered in

order to perform texture-geometry decomposition. In most recent works, the oscillating patterns have been extracted in considering ℓ_1-norm applied on frame coefficients [26, 27, 28].

In this work, we propose to adapt the texture-geometry framework in order to replace the sifting steps. The successive extraction of modes requires an iterative procedure based on specific choice for the involved functionals. In other words, we proceed K sequential steps allowing to extract $(a_k, d_k)_{k\in\{1,\dots,K\}}$ by solving a minimization problem. Note that in usual texture-geometry extraction, the procedure is not sequential but simultaneous, i.e., a variational problem is formulated in order to extract simultaneously $(d_1, \dots, d_K, a_K)$. In order to be comparable to usual EMD we adopt a sequential approach here. The general Sequential Variational Modal Decomposition (SeqVMD) can be formulated as below :

Problem 3.1

We denote $x \in \mathbb{R}^N$ the signal to be decomposed in $(d_1, \dots, d_K, a_K)$. Let $d_0 = 0$ and let $a_0 = x$. For every $k \in \{1, \dots, K\}$,

$$(a_k, d_k) \in \underset{a\in\mathbb{R}^N,\, d\in\mathbb{R}^N}{\operatorname{Argmin}} f(a, d; a_{k-1}) + g(a) + h(d) \tag{3.1}$$

where $g\colon \mathbb{R}^N \to]-\infty, +\infty]$ and $h\colon \mathbb{R}^N \to]-\infty, +\infty]$ are potentials promoting the properties of the trend and fluctuation components of order k separately, while $f(\cdot, \cdot; a_{k-1})\colon \mathbb{R}^N \times \mathbb{R}^N \to]-\infty, +\infty]$ is a coupling term modelling their interaction.

The functions f, g, and h have to be specified to obtain a EMD-like decomposition. While the IMF of order k is expected to have a zero-mean envelope and to be quasi-orthogonal to the IMF of order $j < k$, on the other hand, the trend of order k has to be a smooth signal or at least smoother

that the fluctuation of order k. We now propose to specify $f(\cdot,\cdot;a_{k-1})$, g, and h in order to model the listed properties that we want to impose onto each component.

3.1. Proposed solution

The first condition (zero-mean average envelope) is the most difficult to impose. We propose a method derived from [19] to deal with such a constraint. The location of the local extrema of a_{k-1} are denoted by $(t_k[\ell])_{1\leq\ell\leq L_k}$. Note that these extrema are alternatively minima and maxima. We can approximate the first condition by considering, for every $k\in\{1,\ldots,K\}$,

$$(\forall \ell\in\{1,\ldots,L_k\}) \quad \left| d\big[t_k[\ell]\big] + \frac{\alpha_\ell d\big[t_k[\ell-1]\big] + \beta_\ell d\big[t_k[\ell+1]\big]}{\alpha_\ell+\beta_\ell} \right| < \varepsilon_{k,\ell} \quad (3.2)$$

where $\alpha_\ell = t_k[\ell+1]-t_k[\ell]$, $\beta_\ell = t_k[\ell]-t_k[\ell-1]$, and $\varepsilon_{k,\ell}>0$. The coefficients α_ℓ and β_ℓ are chosen so that, in Eq. (3.2), the point $d\big[t_k[\ell]\big]$ (that is generally close to an extrema of d, e.g., a maximum) is approximately compared to its mirror-point on the would-be other envelope which would be locally defined thanks to $d\big[t_k[\ell-1]\big]$ and $d\big[t_k[\ell+1]\big]$. Note that no envelope is explicitly computed. Figure 2 illustrates this condition.

One could globally rewrite (3.2) as

$$\|D_k d\|_q^q \leq \varepsilon_k, \quad (3.3)$$

where $q\geq 1$, $\varepsilon_k = \sum_{\ell=1}^{L_k}\varepsilon_{k,\ell}^q$ and $D_k\in\mathbb{R}^{L_k\times N}$ denotes a matrix which models the penalization imposed on d at each location $t_k[\ell]$. Each row of D_k is sparse and the non-zero values are 1, $\alpha_\ell/(\alpha_\ell+\beta_\ell)$, and $\beta_\ell/(\alpha_\ell+\beta_\ell)$ at the locations $t_k[\ell]$, $t_k[\ell-1]$, and $t_k[\ell+1]$ respectively.

The smoothness condition on the trend can be achieved by imposing, for every $k \in \{1, \ldots, K\}$,

$$\|Aa\|_p^p \leq \eta_k, \tag{3.4}$$

where $\eta_k > 0$, $p \geq 1$ (typically $p = 1$ or $p = 2$), and A denotes a derivative operator (1st or 2nd order derivative). If $p = 1$, it corresponds to a total variation constraint that favours piecewise constant signals and more generally allows non-smooth behaviour.

We want to impose that the sum of the trend and the IMF of order k is close to a_{k-1}. This can be modeled by considering the constraint

$$a_{k-1} = d + a$$

or, if we want to reduce the sampling effects, we can use a ℓ_2-norm:

$$\|a_{k-1} - d - a\|_2^2.$$

At last, the quasi-orthogonality condition requires to impose a constraint taking the form, for every $k \in \{1, \ldots, K\}$,

$$(\forall j < k) \quad |\langle d, d_j \rangle| \leq \zeta_{k,j}, \tag{3.5}$$

where $\zeta_{k,j} > 0$. For some signals as the one presented in Figure 1 the quasi-orthogonality constraint is usually satisfied but we can only check it a posteriori while adding this constraint allows us to quantify a priori this point.

The criterion we propose to consider is summarized in Problem 3.2.

Problem 3.2

We denote $x \in \mathbb{R}^N$ the signal to be decomposed in $(d_1, \ldots, d_K, a_K)$. Let

$d_0 = 0$ and let $a_0 = x$. For every $k \in \{1, \ldots, K\}$,

$$(a_k, d_k) \in \underset{a \in \mathbb{R}^N,\, d \in \mathbb{R}^N}{\operatorname{Argmin}} \|a_{k-1} - d - a\|_2^2$$
$$\text{subj. to} \quad \begin{cases} \|Aa\|_p^p \leq \eta_k, \\ \|D_k d\|_q^q \leq \varepsilon_k, \\ (\forall j \in \{0, \ldots, k-1\}),\ |\langle d, d_j \rangle| \leq \zeta_{k,j}, \end{cases} \tag{3.6}$$

where $\eta_k > 0$, $\epsilon_k > 0$, $\zeta_{k,j} \geq 0$, $p \geq 1$, $q \geq 1$, $A \in \mathbb{R}^{N \times N}$, and $D_k \in \mathbb{R}^{L_k \times N}$.

Problem 3.2 is a specific case of Problem 3.1 with

$$\big(\forall (a, d) \in \mathbb{R}^N \times \mathbb{R}^N\big) \quad \begin{cases} f(a, d; a_{k-1}) = \|a_{k-1} - d - a\|_2^2, \\ g(a) = \iota_{\|A \cdot\|_p^p \leq \eta_k}(a), \\ h(d) = \iota_{\|D_k \cdot\|_p^p \leq \varepsilon_k}(d) + \sum_{j=0}^{k-1} \iota_{|\langle \cdot, d_j \rangle| \leq \zeta_{k,j}}(d), \end{cases} \tag{3.7}$$

In (3.7), we use the indicator notation that is defined for a Hilbert space $\mathcal{H}$ and a closed convex set $C \in \mathcal{H}$ as, for every $x \in \mathcal{H}$, $\iota_C = 0$ if x belongs to C and $+\infty$ otherwise.

One could have noticed that Problem 3.2 can be equivalently written in a regularized form, i.e.,

$$(a_k, d_k) \in \underset{a \in \mathbb{R}^N}{\operatorname{Argmin}} \|a_{k-1} - a - d\|_2^2 + \lambda_k \|Aa\|_p^p + \chi_k \|D_k d\|_q^q$$
$$\text{subj. to} \quad (\forall j \in \{0, \ldots, k-1\}),\ |\langle d, d_j \rangle| \leq \zeta_{k,j}, \tag{3.8}$$

where $\lambda_k \in \mathbb{R}$ and $\chi_k \in \mathbb{R}$ denotes the regularization parameters. The constrained and the regularized problems lead to the same solutions for specific values for η_k, ε_k, λ_k, and χ_k. However, the bounds (e.g., η_k, ε_k) appear to be generally easier to handle when several constraints are involved

because they have a physical meaning that λ_k and χ_k do not have [30, 31, 32, 33].

3.2. Related solutions

According to the general formulation given by Problem 3.1, one might derive numerous solutions that differ from Problem 3.2. We propose to describe some other solutions that are proposed in the literature or that could be derived and we discuss the reasons why we have preferred the formulation of Problem 3.2

- Trend estimation.

  Considering $f = 0$ and $h = 0$ in Problem 3.1 leads to the estimation of the trend alone. Then, the IMF of order k is obtained considering $d_k = a_{k-1} - a_k$. Smart choices of g have been proposed in [18, 19] as discussed in Section 2.2. Another interesting particular case is discussed in [29] and it corresponds to the following smooth criterion:

  $$a_k \in \underset{a \in \mathbb{R}^N}{\text{Argmin}} \|Aa\|_2^2 + \lambda \|D_k(a_{k-1} - a)\|_2^2$$

  where D_k is also designed from the extrema of a_{k-1}. This specific minimization problem leads to the closed form given below:

  $$a_k = \frac{\lambda D_k^* D_k a_{k-1}}{A^* A + \lambda D_k^* D_k}$$

  $$d_k = a_{k-1} - a_k.$$

  It can be noticed that this formulation is restrictive for large size signals due to the computation of $(A^*A + \lambda D_k^* D_k)^{-1}$. An iterative procedure could then be provided. However, one of the main limitation is that sharp features in the trend could not be recovered with a ℓ_2-norm and that sampling effects are not dealt with.

- Trend and fluctuation estimation

  In order to limitate the sampling effects, a criterion such as

  $$(a_k, d_k) \in \underset{a \in \mathbb{R}^N}{\text{Argmin}} \|a_{k-1} - a - d\|_2^2 + \lambda_1 \|Aa\|_2^2 + \lambda_2 \|D_k d\|_2^2$$

  could have been employed. The solution can be efficiently obtained considering

  $$\begin{aligned} a_k &= \frac{(2 + \lambda_1 A^* A) a_{k-1}}{\text{Id} + \lambda_1 A^* A} \\ d_k &= \frac{(2 + \lambda_2 D_k^* D_k) a_{k-1}}{\text{Id} + \lambda_2 D_k^* D_k}. \end{aligned}$$

  The previous remark related to inability to deal with sharp behaviour of the trend stays valid.

Numerous applications involve multivariate data or a set of univariate data [34, 36, 37, 38]. While the most straightforward approach to extract the modes from this class of signals is to apply EMD separately on each univariate signal [39, 40], efficient bivariate/multivariate EMD solutions have been first proposed in [42, 43, 44, 41, 36]. A part of these methods is based on the projection of the multivariate signal along multiple directions on hyper-spheres in order to calculate their envelopes and local means. The main limitation of this approach is the dimensionality that increases exponentially with the number of signal we want to deal with simultaneously. To overcome this restriction, an alternative solution that estimates the mean envelope as the signal which interpolates the barycenters of the elementary oscillations (i.e., piece of the signal defined between two consecutive local extrema) was proposed in [41] while in [36] the mean envelope is obtained by averaging the mean of the upper and lower envelopes for all the signals. However, all these methods still require a sifting process and do not have

convergence guarantees. In this part, we suggest an alternative based on our convex optimization formalism using a mixed norm [45] rather than a ℓ_p-norm for promoting the properties of the multivariate trend (i.e., favour correlations when they exist). The minimization problem we address is detailed below:

Problem 3.3

We denote $x = (x_i)_{1\leq i\leq I} \in (\mathbb{R}^N)^I$ the multivariate signal to be decomposed in $(d_1, \ldots, d_K, a_K) \in (\mathbb{R}^N)^I \times \ldots \times (\mathbb{R}^N)^I \times (\mathbb{R}^N)^I$. Let $d_0 = 0$ and let $a_0 = x$. For every $k \in \{1, \ldots, K\}$,

$$(a_k, d_k) \in \underset{(a,d)\in(\mathbb{R}^N)^I\times(\mathbb{R}^N)^I}{\operatorname{Argmin}} \sum_{i=1}^{I} \|a_i + d_i - a_{k-1,i}\|_2^2$$

$$\text{subj. to } \begin{cases} \sum_{n=1}^{N} \sqrt{\sum_{i=1}^{I} |(Aa_i)[n]|^2} \leq \eta_k, \\ \sum_{i=1}^{I} \|D_{k,i} d_i\|_q^q \leq \varepsilon_k, \\ (\forall i \in \{1, \ldots, I\})(\forall j \in \{0, \ldots, k-1\}),\ |\langle d_i, d_{j,i} \rangle| \leq \zeta_{k,j,i}, \end{cases}$$

where, for every $i \in \{1, \ldots, I\}$ and $k \in \{1, \ldots, K\}$, $\eta_k > 0$, $\epsilon_k > 0$, $\zeta_{k,j,i} \geq 0$, $p \geq 1$, $A \in \mathbb{R}^{N\times N}$, and $D_{k,i} \in \mathbb{R}^{L_{k,i}\times N}$.

It can be observed that the proposed multivariate SeqVMD is a slightly modified version of the univariate SeqVMD. The difference comes from the grouping behaviour of the trend that we impose through the $\ell_{2,1}$-norm, i.e., through the constraint

$$\left(\forall a = (a_i)_{1\leq i\leq I} \in (\mathbb{R}^N)^I\right) \qquad \sum_{n=1}^{N} \sqrt{\sum_{i=1}^{I} |(Aa_i)[n]|^2} \leq \eta_k.$$

.

3.3. Proposed algorithm

In order to design an efficient algorithm, we have to clearly identify the properties (convexity, differentiability,...) of the functions involved in Problem 3.2. First note that Problem 3.2 denotes a minimization problem under non-linear constraints that can be equivalently written

$$\min_{a\in\mathbb{R}^N,\, d\in\mathbb{R}^N} \|a_{k-1} - d - a\|_2^2 + \iota_{\|A\cdot\|_p^p \leq \eta_k}(a) + \iota_{\|D_k\cdot\|_q^q \leq \varepsilon_k}(d) + \sum_{j=0}^{k-1} \iota_{|\langle\cdot, d_j\rangle| \leq \zeta_{k,j}}(d). \tag{3.9}$$

It clearly appears that (3.9) requires to minimize a criterion involving convex but non-necessarily differentiable functions. According to the recent literature on non-smooth convex optimization (see [47, 48, 49] and the references therein) it is interesting to remark that (3.9) is a particular case of a more general problem that is

$$\min_{u\in\mathcal{H}} \sum_{r=1}^{R} \mathrm{f}_r(L_r u) \tag{3.10}$$

where $\mathcal{H}$ denotes a real Hilbert space (in the following it will be reduced to the product of Euclidean spaces), $(L_r)_{1\leq r\leq R}$ denote linear operators from $\mathcal{H}$ to $\mathbb{R}^{N_r}$, and $(\mathrm{f}_r)_{1\leq r\leq R}$ model convex, lower semi-continuous (l.s.c.), proper[3] functions from $\mathbb{R}^{N_r}$ to $]-\infty, +\infty]$. Indeed, in (3.10), the choices $R = k+1$, $u = (a, d) \in \mathbb{R}^N \times \mathbb{R}^N$, i.e, $\mathcal{H} = \mathbb{R}^N \times \mathbb{R}^N$ denotes a product space equipped with the usual vector space structure and the scalar product

$$(\forall (u_1, u_2) \in \mathcal{H}\times\mathcal{H}), \quad \langle u_1, u_2\rangle \mapsto \langle a_1|a_2\rangle + \langle d_1|d_2\rangle, \tag{3.11}$$

[3]The assumptions convex, lower semi-continuous, proper are the usual assumptions in convex optimization [48].

and

$$\begin{cases} \mathrm{f}_1 \colon (a,d) \in \mathbb{R}^N \times \mathbb{R}^N \mapsto \|a_{k-1} - d - a\|_2^2 \\ \mathrm{f}_2 \colon (\tilde{a},\tilde{d}) \in \mathbb{R}^N \times \mathbb{R}^{L_k} \mapsto \iota_{\|\cdot\|_p^p \leq \eta_k}(\tilde{a}) + \iota_{\|\cdot\|_q^q}(\tilde{d}) \\ \mathrm{f}_3 \colon (a,d) \in \mathbb{R}^N \times \mathbb{R}^N \mapsto \iota_{|\langle \cdot, d_0 \rangle| \leq \zeta_{k,0}}(d) \\ \vdots \\ \mathrm{f}_R \colon (a,d) \in \mathbb{R}^N \times \mathbb{R}^N \mapsto \iota_{|\langle \cdot, d_{k-1} \rangle| \leq \zeta_{k,k-1}}(d) \\ L_1 = L_3 = \ldots = L_R = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \mathrm{Id} \end{pmatrix} \\ L_2 = \begin{pmatrix} A & 0 \\ 0 & D_k \end{pmatrix}. \end{cases} \tag{3.12}$$

lead to the minimization problem (3.9). Note that $(\mathrm{f}_r)_{2 \leq r \leq R}$ denotes non-differentiable convex functions, consequently usual gradient-based algorithms cannot be used to solve (3.9).

For several years, numerous algorithms have been designed to efficiently (in terms of computational time and convergence guarantees) deal with convex but non-smooth optimization techniques. These algorithms often called proximal algorithms are based on the proximity operator [46], that is defined, for every $u \in \mathcal{H}$,

$$\mathrm{prox}_{\mathrm{f}} \colon u \mapsto \underset{v \in \mathcal{H}}{\arg\min}\ \mathrm{f}(v) + \frac{1}{2}\|u - v\|^2. \tag{3.13}$$

where $\mathcal{H}$ denotes a real Hilbert space and f is a convex, l.s.c., and proper function from $\mathcal{H}$ to $]-\infty, +\infty]$. A particular example of proximity operator is the projection onto a nonempty closed convex set $C \subset \mathcal{H}$. Indeed, if $\mathrm{f} = \iota_C$ (the indicator function of C) then the projection operator P_C is prox_{ι_C}. For a detailed account of the theory of proximity operators, see

[47] and references therein. In what follows we recall some of the proximity operators required for solving Problem 3.2.

The projection onto the convex set $C_p^\kappa = \{u \in \mathbb{R}^N \mid \|u\|_p^p \leq \kappa\}$ when $p = 1$ can be computed iteratively by considering [57] or by using epigraphical projection techniques as detailed in [58]. For $p = 2$, the projection onto the convex set $C_p^\kappa = \{u \in \mathbb{R}^N \mid \|u\|_2^2 \leq \kappa\}$ is given in [59] :

$$(\forall u \in \mathbb{R}^N), \quad P_{C_2^\kappa} u = \begin{cases} u, & \text{if} \quad \|u\|_2^2 \leq \kappa \\ \frac{\kappa u}{\|u\|_2^2}, & \text{otherwise.} \end{cases} \tag{3.14}$$

For every $r \in \{1, \ldots, k-1\}$, the projection onto the constraint set $E_r^{\zeta_{r,k}} = \{u \in \mathbb{R}^N \mid |\langle u, d_r \rangle| \leq \zeta_{r,k}\}$ is given in [59] and such that, for every $u \in \mathbb{R}^N$,

$$P_{E_r^{\zeta_{r,k}}} u = \begin{cases} u, & \text{if } |\langle u, d_r \rangle| \leq \zeta_{r,k} \\ u + \frac{\zeta_{r,k} - \langle u, d_r \rangle}{\|d_r\|_2^2} d_r, & \text{otherwise.} \end{cases} \tag{3.15}$$

Let $\tau \in]0, +\infty[$, the proximity operator of $\mathrm{f} = \tau \|\cdot\|_2^2$ is given in [60] by, for every $u \in \mathbb{R}^N$,

$$\mathrm{prox}_{\tau\|\cdot\|_2^2} u = \frac{u}{1 + 2\tau}. \tag{3.16}$$

The proximal algorithms include ADMM [50], Split Bregman iterations [51, 52], or primal-dual algorithms such as Chambolle-Pock algorithm [53, 54, 55, 56]. The large interest for proximal tools enabled to develop a large panel of algorithms to efficiently solve problems being formulated as problem (3.10).

The class of proximal algorithms can be roughly split in two groups: the primal algorithms [47] and the primal-dual algorithms [55]. To summarize,

the primal algorithms generally require to compute the inverse of $\sum_r L_r^* L_r$ whereas the primal-dual iterations only involve the computation of L_r and itd adjoint L_r^*.

Due to the difficulty to invert efficiently $D_k^* D_k$, for every $k \in \{1, \dots, K\}$, and thus the difficulty to invert $\sum_r L_r^* L_r$, Problem 3.2 will be solved with a primal-dual proximal algorithm derived from [62]. For the paper to be self-content, we first recall the iterations and the convergence properties of the algorithm M+SFBF proposed in [62] in Algorithm 3.4 and its convergence guarantees in Proposition 3.5.

Algorithm 3.4 M+SFBF [62]

Initialization

- Let $(\omega_r)_{1 \leq r \leq R}$ be real numbers in $]0, 1]$ such that $\sum_{r=1}^R \omega_r = 1$
- Let $\beta = \max_{1 \leq r \leq R} \|L_r\|$
- Let $\varepsilon \in]0, 1/(\beta + 1)[$ and let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $[\varepsilon, (1 - \varepsilon)/\beta]$
- For every $r \in \{1, \dots, R\}$, $v_{2,r}^{(0)} \in \mathcal{H}$ and $v_{1,r}^{(0)} \in \mathcal{G}_r$

For $n = 0, 1, \dots$

- $u^{(n)} = \sum_{r=1}^R \omega_r v_{2,r}^{(n)}$
- For $r = 1, \dots, R$
  - $y_{1,r}^{(n)} = v_{1,r}^{(n)} - \gamma_n L_r^* v_{2,r}^{(n)}$
  - $y_{2,r}^{(n)} = v_{2,r}^{(n)} + \gamma_n L_r v_{1,r}^{(n)}$
- $p_1^{(n)} = \sum_{r=1}^R \omega_r y_{1,r}^{(n)}$
- For $r = 1, \dots, R$
  - $p_{2,r}^{(n)} = y_{2,r}^{(n)} - \gamma_n \operatorname{prox}_{\gamma_n^{-1} \mathrm{f}_r}(\gamma_n^{-1} y_{2,r}^{(n)})$
  - $q_{1,r}^{(n)} = p_1^{(n)} - \gamma_n L_r^* p_{2,r}^{(n)}$
  - $q_{2,r}^{(n)} = p_{2,r}^{(n)} + \gamma_n L_r p_1^{(n)}$
  - $u_r^{(n+1)} = u_r^{(n)} - y_{1,r}^{(n)} + q_{1,r}^{(n)}$
  - $v_r^{(n+1)} = v_r^{(n)} - y_{2,r}^{(n)} + q_{2,r}^{(n)}$

Proposition 3.5 *[62, Proposition 4.4]*

Let $\mathcal{H}$ be a finite dimensional Hilbert space. For every $r \in \{1, \ldots, R\}$, let $\mathcal{G}_r$ be an Hilbert space, let φ_r be a convex, l.s.c., and proper function from $\mathcal{G}_r$ to $]-\infty, +\infty]$ and let $L_r : \mathcal{H} \to \mathcal{G}_r$ be a bounded linear operator. Assume that the set of solutions of $\min \sum_{r=1}^{R} \varphi_r \circ L_r$ is not empty. The sequence $(u^{(n)})_{n\in\mathbb{N}}$ generated by Algorithm 3.4 converges to a minimizer of $\sum_{r=1}^{R} \varphi_r \circ L_r$.

We may easily derive from Algorithm 3.4 and Proposition 3.5, the proposed Seq-VMD algorithm (cf. Figure 3) and the associated convergence guarantees describe in Proposition 3.6.

Proposition 3.6

For every $k \in \{1, \ldots, K\}$, assume that the set of solutions of (3.9) is not empty, then the sequence $(u_k^{(n)})_{n\in\mathbb{N}}$ generated by Algorithm 6.1 (cf. Figure 3) converges to (a_k, d_k).

The convergence result in Proposition 3.6 is obtained by applying [62, Proposition 4.4] in $\mathcal{H} = \mathbb{R}^N \times \mathbb{R}^N$ and using Eq. (3.11).

In Algorithm 6.1, the constraints sets $(E_r^{\zeta_{r,k}})_{1\leq r\leq k-1}$, $C_p^{\eta_k}$, and $C_q^{\varepsilon_k}$ are involved. The projection operators associated to it have been defined previously.

Note that other multicomponent proximal algorithms could be derived for solving Problem 3.2. The purpose of this work is however not to compare all proximal methods, but to figure out the efficiency of one of them. In the next section, the experimental results will show that despite the iterations which may appear tedious, the solution can be obtained in a few seconds.

4. Experimental results

In order to evaluate the proposed method and compare it to the state-of-the-art, we first analyse the performances of the proposed approach on simulated data and then we analyse the behaviour on real data. In our experiments we consider that the value of K is known. However, one can employ the method proposed in [12] in order to estimate the number of components.

4.1. Simulated data

The first experiment is dedicated to evaluate the behavior of the proposed method for signals with non-smooth components. In Figure 4, we deal with the decomposition of a sum of a triangular trend and a AM-FM signal. We run Seq-VMD with $K = 1$, 3.10^4 iterations, $\eta_1 = 0.1$, $\varepsilon_1 = 1$, and $p = q = 1$ in order to achieve convergence and attain a feasible solution. The time computational cost of EMD [9] is less than 1 second, OMD [19] requires about 5 seconds, and Seq-VMD about 40 seconds. According to the results, the proposed method allows us to efficiently recover sharp features compared to state-of-the-art methods. Note that classical EMD (thanks to sifting process) and the proposed approach allow us to limitate sampling effects but OMD does not take them into account. The normalized mean square errors for each method are summarized in Figure 4. The impact of the choice of η_1, ε_1, p, and q is evaluated in Tables 1 and 2. The proposed method is clearly better for small values of η_1 and ε_1.

A second example is illustrated in Figure 5. It consists in the sum of a rectangular signal and a AM-FM signal. We run Seq-VMD with $K = 1$, 2.10^4 iterations, $\eta_1 = 0.9$, $\varepsilon_1 = 0.2$, and $p = q = 1$ in order to achieve convergence and attain a feasible solution. The time computational cost of

EMD is 0.3 seconds, OMD requires 3 seconds, and Seq-VMD 22 seconds. One can notice that the proposed method performs better than the other ones. EEMD [63] could have been employed to deal with the null sections but this method also suffers from a lack of mathematical background. The good quantitative performances in term of normalized mean square errors of the proposed method are presented in Figure 5. The impact of the choice of η_1, ε_1, p, and q is evaluated in Tables 3 and 4.

The third experiment aims at recovering a sum of two chirps. This is a classical example where usual filter banks cannot perform the decomposition. The results are provided in Figure 6. We run Seq-VMD with $K = 1$, 10^3 iterations, $\eta_1 = 5.6$, $\varepsilon_1 = 0.8$, $p = 2$, and $q = 1$ in order to achieve convergence and attain a feasible solution. The time computational cost of EMD is 0.3 seconds, OMD requires around 6 seconds, and Seq-VMD around 1 second. The performance is quite similar from one method to the other. The impact of the choice of η_1, ε_1, p, and q is evaluated in Tables 5 and 6.

In Figure 7, we deal with a bivariate decomposition. We consider two signals that each consists in the sum of a triangular trend and a AM-FM signal, yet with different frequencies. With this example, we want to illustrate the performance of the $\ell_{2,1}$-norm applied on the two-component trend as compared to a ℓ_1-norm applied on each trend. We run Seq-VMD with $K = 1$, 10^4 iterations, $\eta_1 = 0.8$, and $\varepsilon = 0.06$ in order to achieve convergence and attain a feasible solution. The time computational cost of the bivariate EMD needs less than 1 second, Seq-VMD applied on each component requires around 10 seconds, and the bivariate Seq-VMD takes around 6 seconds. The resulting IMFs of order 1 are similar from one approach to the other, but the trend appears to be better estimated as indicated in the blue boxes.

We present the decomposition results obtained with Seq-VMD for $K = 2$ and $N = 500$ in Figure 8. The first step of Seq-VMD takes 10^3 iterations and the second step takes 10^4 iterations, the parameters are such that $\eta_1 = 6.5$, $\varepsilon_1 = 1.2$, $\eta_2 = 0.1$, $\varepsilon_2 = 2$, $\zeta_{1,2} = 10$, $p = 2$, and $q = 1$ in order to achieve convergence and attain a feasible solution. The time computational cost of the bivariate EMD needs around 1 second, Seq-VMD requires less than 50 seconds, and the OMD takes around 1 minute. The results of the proposed approach are similar to EMD except in the flat areas where the proposed method has a better behaviour. Moreover, the decomposition is better achieved with Seq-VMD than with OMD where residual oscillations can be observed in a_2.

In Tables 1, 2, 3, 4, 5, and 6, one could note that the black cases denote the unfeasible solutions, i.e., solutions for which all the constraints cannot be satisfied simultaneously. Such solutions occur for small values of η_k and ε_k that lead to restrictive constraints.

4.2. Real data

In this section we evaluate the performance of the proposed approach on a real signal that corresponds to a light consumption in a building [64, 65]. This signal is interesting because it presents some discontinuities (signal equals zero when the light is shut down) that are usually difficulty to deal with usual mode decomposition methods. The results are provided in Figure 9. We run Seq-VMD with $K = 2$, $\eta_1 = 0.8$, $\varepsilon_1 = 0.001$, $\eta_2 = 0.03$, $\varepsilon_2 = 0.1$, $\zeta_{1,2} = 10^4$, $p = q = 1$ (for $k = 1$) and $p = 2$ and $q = 1$ (for $k = 2$) in order to achieve convergence and that a feasible solution exists. The computational cost of the EMD is 0.4 seconds, the VMD requires around 4 seconds, and the Seq-VMD less than 13 second. Thanks to its behaviour on

non-smooth components, where EMD and OMD suffer from mode mixing (i.e., attempt at decomposing flat part in oscillations), the proposed method leads to a better mode decomposition (i.e., flat area are preserved on the first IMF, no mode mixing between the first and the second IMF, the trend is correctly extracted).

5. Conclusions and discussion

In this work we have proposed an efficient alternative to EMD in order to deal with mode extraction. The empirical sifting process step is replaced by a convex optimization procedure having convergence guarantees. Although the proposed solution is related to texture-geometry decomposition ideas developed in image processing, we incorporate EMD spirit considering extrema-based constraints, quasi-orthogonality constraint, and iterative mode extraction rather than a simultaneous mode extraction as proposed in [24] in a context of texture-geometry extraction. The good performances of the proposed algorithm are evaluated on signals where usual filterbanks would not be suited (e.g., sum of chirps) and on signals having non-smooth behaviour which may be often encountered in real data such as the one considered in building energy consumption.

The constraint formulation allows us to facilitate the selection of the parameters, however it also leads to the problem of unfeasible solutions that may be handled carefully.

6. Acknowledgments

The authors would like to thank Romain Fontugne, Hideya Ochiai, and Hiroshi Esaki of the University of Tokyo for the real data of building energy consumption from [64, 65].

This work is supported in part by Agence Nationale de la Recherche under grant ANR-13-BS03-0002-01

References

[1] N. Pustelnik, P. Borgnat, P. Flandrin, A multicomponent proximal algorithm for Empirical Mode Decomposition, in: Proc. Eur. Sig. Proc. Conference, Bucharest, Romania, 2012, pp. 1880–1884.

[2] M. Priestley, Non-linear and non-stationary time series analysis, Academic Press, 1989.

[3] N. E. Huang, S. S. Shen (Eds.), Hilbert-Huang transform and its applications, Vol. 5, World Scientific, Singapore, 2005.

[4] J. Ville, Théorie et applications de la notion de signal analytique, Câbles et Transmissions 2A (1948) 61–74.

[5] W. K. Meville, Wave modulation and breakdown, J. Fluid Mech. 128 (1983) 489–506.

[6] S. O. Rice, Mathematical analysis of random noise, Bell Syst. Tech. J. 23 (1944) 282–310.

[7] P. Flandrin, F. Auger, E. Chassande-Mottin, Time-frequency re-assignment — From principles to algorithms, in: A. Papandreou-Suppappola (Ed.), Applications in Time- Frequency Signal Processing, CRC Press, Boca Raton, FL, 2003, Ch. 5, pp. 179–203.

[8] I. Daubechies, J. Lu, H.-T. Wu, Synchrosqueezed wavelet transforms: An Empirical Mode Decomposition-like tool, Appl. Comp. Harm. Analysis 30 (2) (2011) 243–261.

[9] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, H. H. Liu, The Empirical Mode Decomposition and the Hilbert

spectrum for nonlinear and nonstationary time series analysis, Proc. R. Soc. Lond. A 454 (1998) 903–995.

[10] G. Thakur, H.-T. Wu, Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples, SIAM J. Math. Anal. 43 (2011) 2078–2095.

[11] S. Meignen, T. Oberlin, S. McLaughlin, On the mode synthesis in the synchrosqueezing method, in: Proc. Eur. Sig. Proc. Conference, Bucharest, Romania, 2012, pp. 1865–1869.

[12] N. Saulig, N. Pustelnik, P. Borgnat, P. Flandrin, and V. Sucic, Instantaneous counting of components in nonstationary signals, in: Proc. Eur. Sig. Proc. Conference, Marrakech, Morocco, Sept. 9-13, 2013, 5p.

[13] G. Rilling, P. Flandrin, Sampling effects on the Empirical Mode Decomposition, Adv. Adapt. Data Anal. 1 (1) (2009) 43–59.

[14] B. Huang, A. Kunoth, An optimization-based Empirical Mode Decomposition scheme, J. Comput. Appl. Math. 240 (2013) 174–183.

[15] E. Deléchelle, J. Lemoine, O. Niang, Empirical Mode Decomposition: an analytical approach for sifting process, IEEE Signal Process. Lett. 12(11) (2005) 764–767.

[16] Y. Kopsinis, S. McLaughlin, Investigation and performance enhancement of the Empirical Mode Decomposition method based on a heuristic search optimization approach, IEEE Trans. Signal Process. 56(1) (2008) 1–13.

[17] S.D. Hawley, L.E. Atlas, H.J. Chizeck, Some properties of an empirical mode type signal decomposition algorithm, IEEE Signal Process. Lett. 17(1) (2010) 24–27.

[18] S. Meignen, V. Perrier, A new formulation for Empirical Mode Decomposition based on constrained optimization, IEEE Signal Process. Lett. 14 (12) (2007) 932–935.

[19] T. Oberlin, S. Meignen, V. Perrier, An alternative formulation for the Empirical Mode Decomposition, IEEE Trans. Signal Process. 60 (5) (2012) 2236–2246.

[20] T. Y. Hou, Z. Shi, Adaptive data analysis via sparse time-frequency representation, Adv. Adapt. Data Anal. 3 (1-2) (2011) 1–28.

[21] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1-4) (1992) 259–268.

[22] T. Y. Hou, Z. Shi, Data-driven time-frequency analysis, Appl. Comp. Harm. Anal. (2013). To appear.

[23] T. Y. Hou, Z. Shi, P. Tavallali, Convergence of a data-driven time-frequency analysis methods (2013). Preprint, http://arxiv.org/abs/1303.7048.

[24] J.-F. Aujol, G. Gilboa, T. Chan, S. Osher, Structure-texture image decomposition - modeling, algorithms, and parameter selection, Int. J. Comp. Vis. 67 (1) (2006) 111–136.

[25] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, A. Chambolle, Image decomposition into a bounded variation component and an oscillating component, Int. J. Comp. Vis. 22 (2005) 71–88.

[26] L. M. Briceño-Arias, P. L. Combettes, J.-C. Pesquet, N. Pustelnik, Proximal algorithms for multicomponent image processing, J. Math. Imag. Vis. 41 (1) (2011) 3–22.

[27] J.-L. Starck, M. Elad, D. Donoho, Image decomposition via the combination of sparse representations and a variational approach, IEEE Trans. Image Process. 14 (2005) 1570–1582.

[28] S. Anthoine, E. Pierpaoli, I. Daubechies, Deux méthodes de déconvolution et séparation simultanées; application à la reconstruction des amas de galaxies, Trait. Signal 23 (5-6) (2006) 439–447.

[29] M. Colominas, G. Schlotthauer, M. Torres, An unconstrained optimization approach to Empirical Mode Decomposition, Tech. rep., Universidad Nacional de Entre Ríos (2013).

[30] D. C. Youla, H. Webb, Image restoration by the method of convex projections. Part I - theory, IEEE Trans. Med. Imag. 1 (2) (1982) 81–94.

[31] H. J. Trussell, M. R. Civanlar, The feasible solution in signal restoration, IEEE Trans. Acous., Speech Signal Process. 32 (2) (1984) 201–212.

[32] P. L. Combettes, Inconsistent signal feasibility problems : least-squares solutions in a product space, IEEE Trans. Signal Process. 42 (11) (1994) 2955–2966.

[33] K. Kose, V. Cevher, A. E. Cetin, Filtered variation method for denoising and sparse signal processing, in: Proc. Int. Conf. Acoust., Speech Signal Process., Kyoto, Japan, 2012, pp. 3329 – 3332.

[34] N. Ur Rehman, Y. Wia, D. Mandic, Application of multivariate Empirical Mode Decomposition for seizure detection in EEG signals, in: IEEE Conf. Eng. Med. Biol. Soc., Buenos Aires, Argentina, 2010, pp. 1650–1653.

[35] A. Mutlu, S. Aviyente, Multivariate Empirical Mode Decomposition for quantifying multivariate phase synchronization, EURASIP J. Adv. Signal Process. 615717 (2011) 1–13.

[36] J. Fleureau, A. Kachenoura, L. Albera, J.C. Nunes, L. Senhadji, Multivariate Empirical Mode Decomposition and application to multichannel filtering, Signal Process. 91(12) (2011) 2783–2792.

[37] A. Mutlu, S. Aviyente, Multivariate Empirical Mode Decomposition for quantifying multivariate phase synchronization, EURASIP J. Adv. Signal Process. 615717 (2011) 1–13.

[38] F. Fontugne, J. Ortiz, D. Culler, H. Esaki, Empirical Mode Decomposition for

intrinsic-relationship extraction in large sensor deployments, in: IoT-App12, Workshop on Internet of Things Applications, Beijing, China, 2012, pp. 1–6

[39] T. B. J. Kuo, C. C. H. Yang, N. E. Huang, Quantification of respiratory sinus arrythmia using hilbert-huang transform, Adv. Adapt. Data Anal. 1 (2) (2009) 295–307.

[40] X. Chen, Z. Wu, N. E. Huang, The time-dependent intrinsic correlation based on the Empirical Mode Decomposition, Adv. Adapt. Data Anal. 2 (2) (2010) 233–265.

[41] J. Fleureau, J.C. Nunes, A. Kachenoura, L. Albera, L. Senhadji, Turning tangent Empirical Mode Decomposition: A framework for mono- and multivariate signals, IEEE Trans. Signal Process. 59(3) (2011) 1309–1316.

[42] T. Tanaka, D. P. Mandic, Complex Empirical Mode Decomposition, IEEE Signal Process. Lett. 14(2) (2007) 101–104.

[43] G. Rilling, P. Flandrin, P. Goncalves, J. M. Lilly, Bivariate Empirical Mode Decomposition, IEEE Signal Process. Lett. 14 (2007) 936–939.

[44] N. Rehman, D. Mandic, Multivariate Empirical Mode Decomposition, Proc. R. Soc. A 466 (2010) 1291–1302.

[45] M. Kowalski, Sparse regression using mixed norms, Appl. Comp. Harm. Analysis 27 (3) (2009) 303–324.

[46] J. J. Moreau, Proximité et dualité dans un espace hilbertien, Bull. Soc. Math. France 93 (1965) 273–299.

[47] P. L. Combettes, J.-C. Pesquet, Proximal splitting methods in signal processing, in: H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, H. Wolkowicz (Eds.), Fixed-Point Algorithms for Inverse Problems in Science and Engineering, Springer-Verlag, New York, 2010, pp. 185–212.

[48] H. H. Bauschke, P. L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer, New York, 2011.

[49] F. Bach, R. Jenatton, J. Mairal, G. Obozinski, Optimization with sparsity-inducing penalties, Foundations and Trends in Machine Learning 4 (1) (2012) 1–106.

[50] M. V. Afonso, J. M. Bioucas-Dias, M. A. T. Figueiredo, An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems, IEEE Trans. Image Process. 20 (3) (2011) 681–695.

[51] T. Goldstein, S. Osher, The split Bregman method for ℓ_1-regularized problems, SIAM J. Imaging Sci. 2 (2009) 323–343.

[52] S. Setzer, G. Steidl, T. Teuber, Deblurring Poissonian images by split Bregman techniques, J. Visual Communication and Image Representation 21 (3) (2010) 193–199.

[53] A. Chambolle, T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, J. Math. Imag. Vis. 40 (1) (2011) 120–145.

[54] B. C. Vũ, A splitting algorithm for dual monotone inclusions involving cocoercive operators, Adv. Comput. Math. 38 (2011) 667–681.

[55] P. L. Combettes, J.-C. Pesquet, Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators, Set-Valued Var. Anal. 20(2) (2012) 307–330.

[56] L. Condat, A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms, J. Optim. Theory Appl. To appear.

[57] E. Van Den Berg, M. P. Friedlander, Probing the Pareto frontier for basis pursuit solutions, SIAM J. Sci. Comput. 31 (2) (2008) 890–912.

[58] G. Chierchia, N. Pustelnik, J.-C. Pesquet, B. Pesquet-Popescu, A proximal approach for constrained cosparse modelling, in: Proc. Int. Conf. Acoust., Speech Signal Process., Kyoto, Japan, 2012, pp. 3433–3436.

[59] R. T. Rockafellar, Conjugate duality and optimization, SIAM, Philadelphia, PA, 1974.

[60] C. Chaux, P. L. Combettes, J.-C. Pesquet, V. R. Wajs, A variational formulation for frame-based inverse problems, Inverse Problems 23 (4) (2007) 1495–1518.

[61] D. L. Donoho, De-noising by soft-thresholding, IEEE Trans. Inform. Theory 41 (3) (1995) 613–627.

[62] L. M. Briceño-Arias, P. L. Combettes, A monotone + skew splitting model for composite monotone inclusions in duality, SIAM J. Opt. 21 (4) (2011) 1230–1250.

[63] G. Schlotthauer, M. Torres, H. Rufiner, P. Flandrin, EMD of Gaussian white noise: Effects of signal length and sifting number on the statistical properties of intrinsic mode functions, Adv. Adapt. Data Anal. 1 (4) (2009) 517–527.

[64] R. Fontugne, J. Ortiz, N. Tremblay, P. Borgnat, P. Flandrin, K. Fukuda, D. Kuller, H. Esaki, Strip, bind, and search: A method for identifying abnormal energy consumption in buildings, in: 12th ACM/IEEE Conf. on Info. Proc. and Sensor Networks, Philadelphia, PA, 2013, pp. 1–12.

[65] R. Fontugne, N. Tremblay, P. Borgnat, P. Flandrin, H. Esaki, Mining anomalous electricity consumption using Ensemble Empirical Mode Decomposition, in: Proc. Int. Conf. Acoust., Speech Signal Process., Vancouver, Canada, 2013, pp. 1–5.

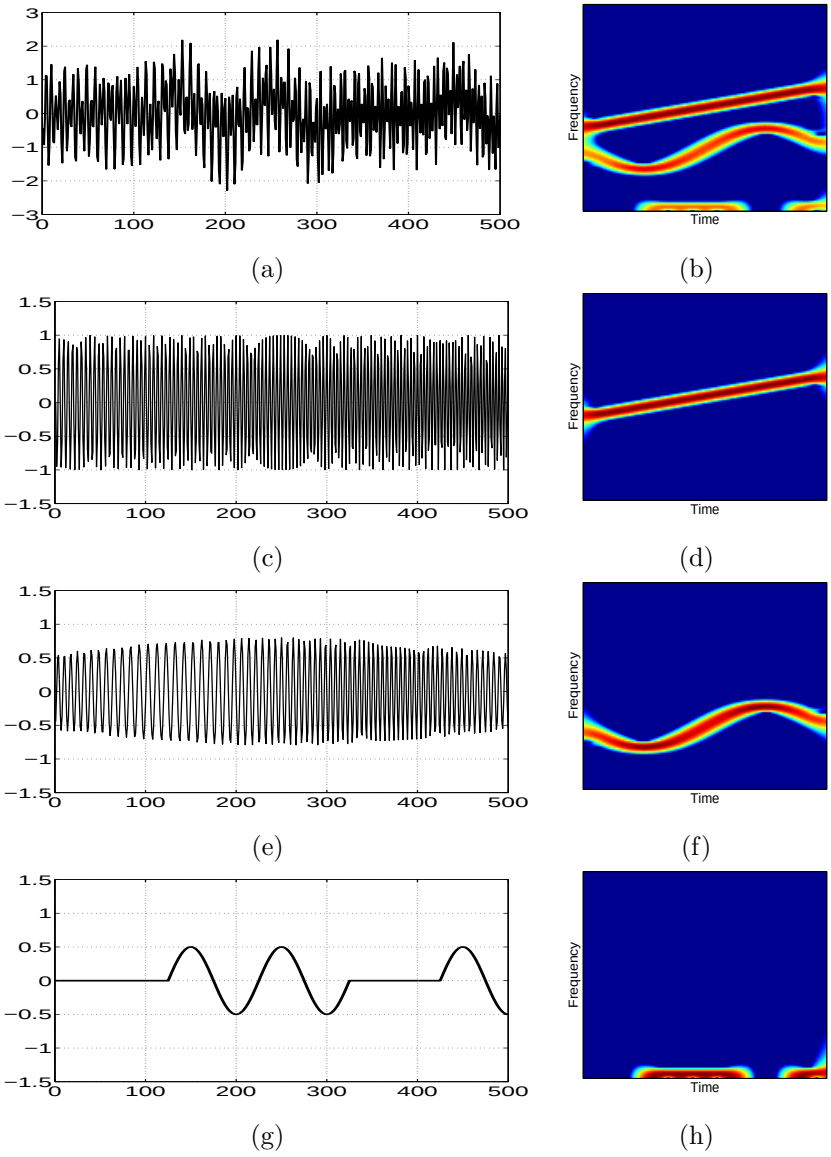


Figure 1: (a) Signal $x(t)$ that is a sum of $K = 3$ AM-FM signals and (b) its spectrogram. (c) First AM-FM component $x_1(t) = \alpha_1(t)\cos\theta_1(t)$ and (d) its spectrogram. (e) Second AM-FM component $x_2(t) = \alpha_2(t)\cos\theta_2(t)$ and (f) its spectrogram. (g) Third AM-FM component $x_3(t) = \alpha_3(t)\cos\theta_3(t)$ and (h) its spectrogram.

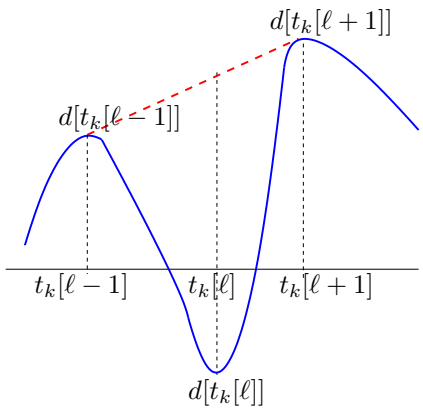


Figure 2: Illustration of the zero-mean envelope constraint

Algorithm 6.1 For $k = 1, \dots, K$

Initialization

Let $(\omega_r)_{1\leq r\leq R}$ be real numbers in $]0,1]$ such that $\sum_{r=1}^{R}\omega_r = 1$

Evaluate D_k according to a_k and let $\beta = \max\left\{1, \max\{\|A\|, \|D_k\|\}\right\}$

Let $\varepsilon \in]0, 1/(\beta+1)[$ and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (1-\varepsilon)/\beta]$

$\big(\forall r \in \{1,\dots,k-1\}\big)$, $v_{2,r}^{(0)} = \left(\overline{v}_{2,r}^{(0)}, \widetilde{v}_{2,r}^{(0)}\right) \in \mathbb{R}^N \times \mathbb{R}^N$ and $v_{1,r}^{(0)} = \left(\overline{v}_{1,r}^{(0)}, \widetilde{v}_{1,r}^{(0)}\right) \in \mathbb{R}^N \times \mathbb{R}^N$

$v_{2,k}^{(0)} \in \mathbb{R}^N$ and $v_{1,k}^{(0)} = \left(\overline{v}_{1,k}^{(0)}, \widetilde{v}_{1,k}^{(0)}\right) \in \mathbb{R}^N \times \mathbb{R}^N$

$v_{2,k+1}^{(0)} = \left(\overline{v}_{2,k+1}^{(0)}, \widetilde{v}_{2,k+1}^{(0)}\right) \in \mathbb{R}^N \times \mathbb{R}^N$ and $v_{1,k+1}^{(0)} = \left(\overline{v}_{1,k+1}^{(0)}, \widetilde{v}_{1,k+1}^{(0)}\right) \in \mathbb{R}^N \times \mathbb{R}^{L_{k+1}}$

Fix the parameters η_k, ε_k, and, for every $r \in \{1,\dots,k-1\}$, $\zeta_{r,k}$.

For $n = 0, 1, \dots$

$u_k^{(n)} = \sum_{r=1}^{k+1}\omega_r v_{2,r}^{(n)}$

For $r = 1, \dots, k-1$

$y_{1,r}^{(n)} = v_{1,r}^{(n)} - \gamma_n v_{2,r}^{(n)}$

$y_{2,r}^{(n)} = v_{2,r}^{(n)} + \gamma_n v_{1,r}^{(n)}$

$y_{1,k}^{(n)} = \big(\overline{v}_{1,k}^{(n)} + \widetilde{v}_{1,k}^{(n)}\big) - \gamma_n v_{2,k}^{(n)}$

$y_{2,k}^{(n)} = v_{2,k}^{(n)} + \gamma_n\big(\overline{v}_{1,k}^{(n)}, \widetilde{v}_{1,k}^{(n)}\big)$

$y_{1,k+1}^{(n)} = \big(\overline{v}_{1,k+1}^{(n)} - \gamma_n A^*\overline{v}_{2,k+1}^{(n)}, \widetilde{v}_{1,k+1}^{(n)} - \gamma_n D_k^*\widetilde{v}_{2,k+1}^{(n)}\big)$

$y_{2,k+1}^{(n)} = \big(\overline{v}_{2,k+1}^{(n)} + \gamma_n A\overline{v}_{1,k+1}^{(n)}, \widetilde{v}_{2,k+1}^{(n)} + \gamma_n D_k\widetilde{v}_{1,k+1}^{(n)}\big)$

$p_1^{(n)} = \sum_{r=1}^{k+1}\omega_r y_{1,r}^{(n)}$

For $r = 1, \dots, k-1$

$p_{2,r}^{(n)} = y_{2,r}^{(n)} - \gamma_n P_{E_r^{\zeta_{r,k}}}(\gamma_n^{-1}y_{2,r}^{(n)})$

$p_{2,k}^{(n)} = y_{2,k}^{(n)} - \gamma_n \operatorname{prox}_{\gamma_n^{-1}\|\cdot - a_{k-1}\|_2^2}(\gamma_n^{-1}y_{2,k}^{(n)})$

$p_{2,k+1}^{(n)} = \big(\overline{y}_{2,k+1}^{(n)} - \gamma_n P_{C_p^{\eta_k}}(\gamma_n^{-1}\overline{y}_{2,k+1}^{(n)}), \widetilde{y}_{2,k+1}^{(n)} - \gamma_n P_{C_q^{\varepsilon_k}}(\gamma_n^{-1}\widetilde{y}_{2,k+1}^{(n)})\big)$

For $r = 1, \dots, k-1$

$q_{1,r}^{(n)} = p_{1,r}^{(n)} - \gamma_n p_{2,r}^{(n)}$

$q_{2,r}^{(n)} = p_{2,r}^{(n)} + \gamma_n p_{1,r}^{(n)}$

$q_{1,k}^{(n)} = \big(\overline{p}_{1,k}^{(n)} + \widetilde{p}_{1,k}^{(n)}\big) - \gamma_n p_{2,k}^{(n)}$

$q_{2,k}^{(n)} = p_{2,k}^{(n)} + \gamma_n\big(\overline{p}_{1,k}^{(n)}, \widetilde{p}_{1,k}^{(n)}\big)$

$q_{1,k+1}^{(n)} = \big(\overline{p}_{1,k+1}^{(n)} - \gamma_n A^*\overline{p}_{2,k+1}^{(n)}, \widetilde{p}_{1,k+1}^{(n)} - \gamma_n D_k^*\widetilde{p}_{2,k+1}^{(n)}\big)$

$q_{2,k+1}^{(n)} = \big(\overline{p}_{2,k+1}^{(n)} + \gamma_n A\overline{p}_{1,k+1}^{(n)}, \widetilde{v}_{2,k+1}^{(n)} + \gamma_n D_k\widetilde{p}_{1,k+1}^{(n)}\big)$

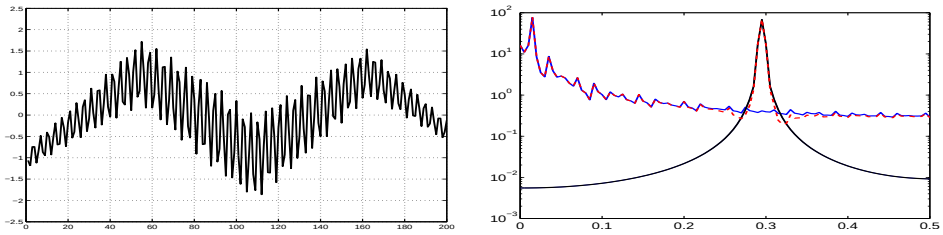
For $r = 1, \dots, k+1$

$u_r^{(n+1)} = u_r^{(n)} - y_{1,r}^{(n)} + q_{1,r}^{(n)}$

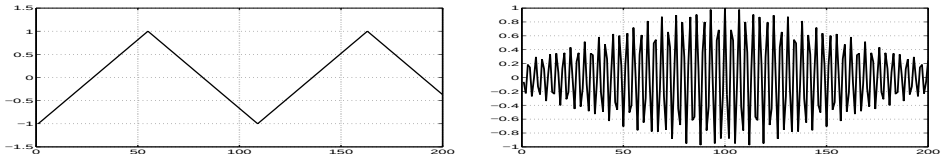
$v_r^{(n+1)} = v_r^{(n)} - y_{2,r}^{(n)} + q_{2,r}^{(n)}$

$(a_k, d_k) = \lim_{n\to\infty} u_k^{(n)}$

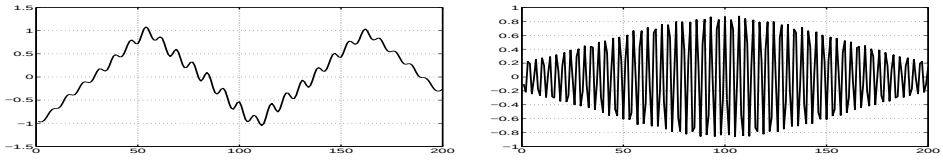
Figure 3: Seq-VMD algorithm



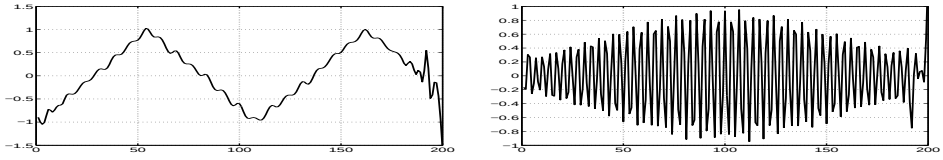
Original signal x | DFT a_1 (blue), d_1 (black), x (red)



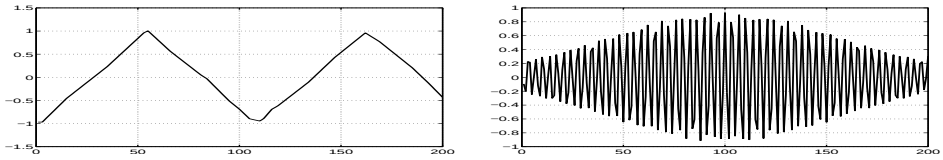
Original signal a_1 | Original signal d_1



EMD a_1 (EQM = 0.018) | EMD d_1 (EQM = 0.023)



OMD a_1 (EQM = 0.047) | OMD d_1 (EQM = 0.058)



Seq-VMD a_1 (**EQM = 0.002**) | Seq-VMD d_1 (**EQM = 0.016**)

Figure 4: Decomposition of a sum of a triangular trend and a AM-FM signal.

	0.05	0.10	0.50	1.00	5.00	10.00	50.00
0.05	0.0167	0.0173	0.0300	0.0469	0.1868	0.3078	0.3078
0.10	0.0232	0.0240	0.0358	0.0563	0.1909	0.3102	0.3102
0.50	0.0421	0.0458	0.0610	0.0824	0.2021	0.3141	0.3141
1.00	0.0631	0.0670	0.0855	0.1081	0.2273	0.3439	0.3439
5.00	0.1851	0.1861	0.1929	0.2087	0.3080	0.4192	0.4192
10.00	0.3919	0.3922	0.3962	0.4024	0.4701	0.5565	0.5565
50.00	0.3919	0.3922	0.3962	0.4024	0.4701	0.5565	0.5565
0.05	0.0136	0.0141	0.0244	0.0382	0.1518	0.2501	0.2501
0.10	0.0188	0.0195	0.0291	0.0457	0.1552	0.2521	0.2521
0.50	0.0342	0.0372	0.0496	0.0669	0.1643	0.2553	0.2553
1.00	0.0512	0.0544	0.0695	0.0879	0.1847	0.2795	0.2795
5.00	0.1505	0.1512	0.1568	0.1696	0.2503	0.3407	0.3407
10.00	0.3185	0.3187	0.3220	0.3271	0.3820	0.4522	0.4522
50.00	0.3185	0.3187	0.3220	0.3271	0.3820	0.4522	0.4522

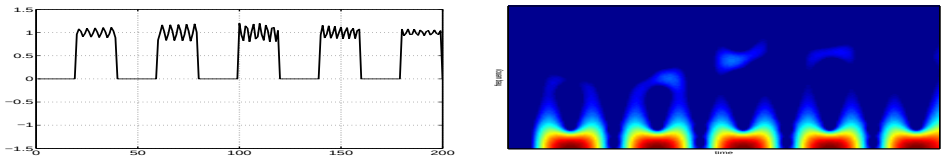
	0.05	0.10	0.50	1.00	5.00	10.00	50.00
0.05			0.0180	0.0176	0.0408	0.0837	0.3078
0.10	0.0278	0.0288	0.0295	0.0297	0.0474	0.0915	0.3102
0.50	0.0399	0.0412	0.0443	0.0460	0.0740	0.1261	0.3141
1.00	0.0579	0.0594	0.0635	0.0662	0.1004	0.1543	0.3439
5.00	0.1882	0.1884	0.1894	0.1911	0.2116	0.2483	0.4192
10.00	0.3929	0.3930	0.3934	0.3935	0.4032	0.4263	0.5565
50.00	0.3929	0.3930	0.3934	0.3935	0.4032	0.4263	0.5565
0.05			0.0146	0.0143	0.0332	0.0680	0.2501
0.10	0.0226	0.0234	0.0240	0.0241	0.0385	0.0743	0.2521
0.50	0.0324	0.0335	0.0360	0.0374	0.0601	0.1025	0.2553
1.00	0.0470	0.0483	0.0516	0.0538	0.0816	0.1254	0.2795
5.00	0.1530	0.1531	0.1539	0.1553	0.1720	0.2018	0.3407
10.00	0.3193	0.3194	0.3197	0.3198	0.3277	0.3464	0.4522
50.00	0.3193	0.3194	0.3197	0.3198	0.3277	0.3464	0.4522

Table 1: Normalized mean square error for different values of η_1 (different lines) and ε_1 (different columns), $p = 2$, $q = 2$ (top) and $q = 1$ (bottom). The light gray cases correspond to results better than EMD and OMD while the dark gray cases match the results only better than OMD. The black cases denote the unfeasible solutions. It corresponds to the decomposition result of a sum of a triangular trend and a AM-FM signal. The top table presents the normalized MSE for d_1 while the bottom table lists the results for a_1.

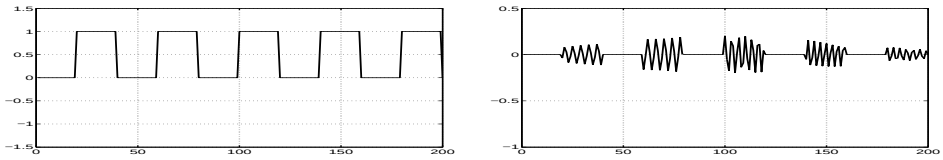
	0.05	0.10	0.50	1.00	5.00	10.00	50.00
0.05					0.1659	0.3073	0.3073
0.10			0.0099	0.0285	0.1673	0.3072	0.3072
0.50			0.0184	0.0362	0.1728	0.3077	0.3077
1.00		0.0120	0.0242	0.0410	0.1780	0.3086	0.3086
5.00	0.0365	0.0383	0.0548	0.0744	0.2041	0.3262	0.3262
10.00	0.0754	0.0789	0.1004	0.1222	0.2523	0.3634	0.3634
50.00	0.3919	0.3922	0.3962	0.4024	0.4701	0.5565	0.5565
0.05					0.1348	0.2498	0.2498
0.10			0.0080	0.0232	0.1360	0.2497	0.2497
0.50			0.0150	0.0294	0.1404	0.2501	0.2501
1.00		0.0097	0.0196	0.0333	0.1447	0.2508	0.2508
5.00	0.0297	0.0312	0.0445	0.0604	0.1658	0.2651	0.2651
10.00	0.0613	0.0641	0.0816	0.0993	0.2050	0.2953	0.2953
50.00	0.3185	0.3187	0.3220	0.3271	0.3820	0.4522	0.4522

	0.05	0.10	0.50	1.00	5.00	10.00	50.00
0.05							0.3073
0.10	0.0165	0.0159	0.0127	0.0100	0.0040	0.0513	0.3072
0.50	0.0171	0.0168	0.0141	0.0119	0.0181	0.0658	0.3077
1.00	0.0192	0.0188		0.0151	0.0338	0.0747	0.3086
5.00	0.0361	0.0373	0.0393	0.0400	0.0615	0.1066	0.3262
10.00	0.0682	0.0696	0.0738	0.0761	0.1053	0.1535	0.3634
50.00	0.3929	0.3930	0.3934	0.3935	0.4032	0.4263	0.5565
0.05							0.2498
0.10	0.0019	0.0018	0.0016	0.0015	0.0032	0.0417	0.2497
0.50	0.0039	0.0039	0.0037	0.0040	0.0147	0.0535	0.2501
1.00	0.0083	0.0083		0.0090	0.0275	0.0607	0.2508
5.00	0.0293	0.0303	0.0320	0.0325	0.0500	0.0866	0.2651
10.00	0.0554	0.0565	0.0600	0.0619	0.0856	0.1248	0.2953
50.00	0.3193	0.3194	0.3197	0.3198	0.3277	0.3464	0.4522

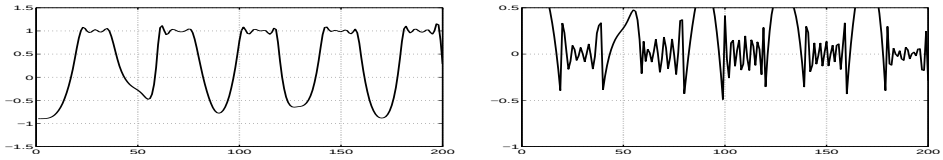
Table 2: Normalized mean square error for different values of η_1 (different lines) and ε_1 (different columns), $p = 2$, $q = 2$ (top) and $q = 1$ (bottom). The light gray cases correspond to results better than EMD and OMD while the dark gray cases match the results only better than OMD. The black cases denote the unfeasible solutions. It corresponds to the decomposition result of a sum of a triangular trend and a AM-FM signal. The top table presents the normalized MSE for d_1 while the bottom table lists the results for a_1.



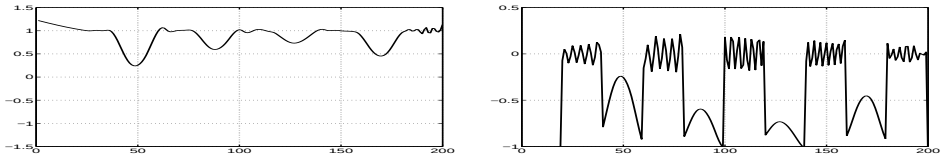
Original signal x | Spectrogram



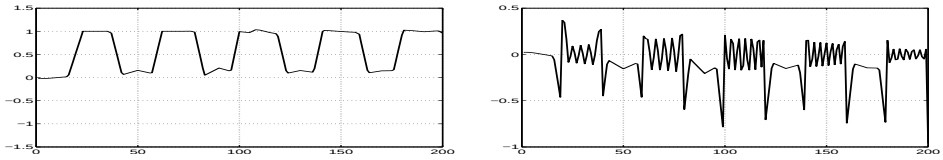
Original signal a_1 | Original signal d_1



EMD a_1 (EQM = 0.303) | EMD d_1 (EQM = 29.1)



OMD a_1 (EQM = 0.658) | OMD d_1 (EQM = 63.3)



Seq-VMD a_1 (**EQM = 0.086**) | Seq-VMD d_1 (**EQM = 8.31**)

Figure 5: Decomposition of a sum of a rectangular trend and a chirp signal.

	0.05	0.10	0.50	1.00	5.00	10.00	50.00
0.05	74.9329	73.0283	75.1258	76.3650	71.6330	22.3897	31.4454
0.10	58.7398	56.9335	57.3676	58.8151	13.6045	21.4195	26.5218
0.50	4.2408	4.6040	5.7935	7.0201	14.2755	19.8392	24.2024
1.00	4.9829	5.3223	6.7908	8.2000	15.6040	19.7077	24.2024
5.00	11.0635	11.2055	11.8293	12.2245	15.8395	19.7077	24.2024
10.00	11.0635	11.2055	11.8293	12.2245	15.8395	19.7077	24.2024
50.00	11.0635	11.2055	11.8293	12.2245	15.8395	19.7077	24.2024
0.05	0.7970	0.7759	0.7814	0.7943	0.7451	0.2329	0.3271
0.10	0.6110	0.5922	0.5967	0.6118	0.1415	0.2228	0.2759
0.50	0.0441	0.0479	0.0603	0.0730	0.1485	0.2064	0.2517
1.00	0.0518	0.0554	0.0706	0.0853	0.1623	0.2050	0.2517
5.00	0.1151	0.1166	0.1230	0.1272	0.1648	0.2050	0.2517
10.00	0.1151	0.1166	0.1230	0.1272	0.1648	0.2050	0.2517
50.00	0.1151	0.1166	0.1230	0.1272	0.1648	0.2050	0.2517

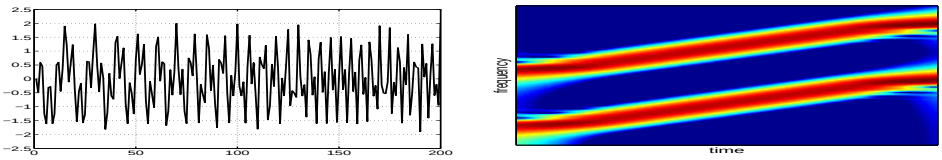
	0.05	0.10	0.50	1.00	5.00	10.00	50.00
0.05					97.7348	92.2843	31.4454
0.10		50.3829	57.9646	57.3807	70.6997	11.3666	26.5218
0.50	3.3318	3.5169	4.2410	4.7184	6.7169	9.8981	24.2024
1.00	3.9052	4.1315	4.9170	5.4150	7.6817	11.1345	24.2024
5.00	10.6108	10.7798	11.1961	11.4939	12.2361	13.8879	24.2024
10.00	10.6791	10.7798	11.1961	11.4939	12.2361	13.8879	24.2024
50.00	10.6791	10.7798	11.1961	11.4939	12.2361	13.8879	24.2024
0.05					1.0166	0.9599	0.3271
0.10		0.5241	0.6029	0.5969	0.7354	0.1182	0.2759
0.50	0.0347	0.0366	0.0441	0.0491	0.0699	0.1030	0.2517
1.00	0.0406	0.0430	0.0511	0.0563	0.0799	0.1158	0.2517
5.00	0.1104	0.1121	0.1165	0.1196	0.1273	0.1445	0.2517
10.00	0.1111	0.1121	0.1165	0.1196	0.1273	0.1445	0.2517
50.00	0.1111	0.1121	0.1165	0.1196	0.1273	0.1445	0.2517

Table 3: Normalized mean square error for different values of η_1 (different lines) and ε_1 (different columns), $p = 2$, $q = 2$(top) and $q = 1$ (bottom). The light gray cases correspond to results better than EMD and OMD while the dark gray cases match the results only better than OMD. The black cases denote the unfeasible solutions. It corresponds to the decomposition result of a sum of a rectangular trend and a chirp signal. The top table presents the normalized MSE for d_1 while the bottom table lists the results for a_1.

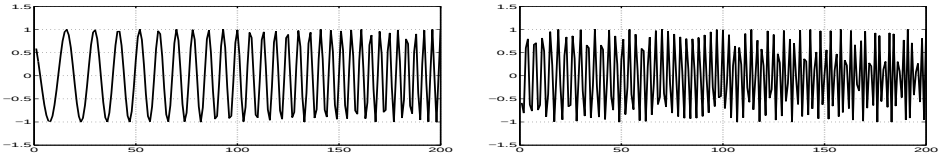
	0.05	0.10	0.50	1.00	5.00	10.00	50.00
0.05				66.0872	60.6300	49.3995	60.3747
0.10			51.2752	61.4812	50.5244	42.0365	52.0189
0.50	13.7013	12.7633	10.0532	9.2128	13.4984	21.5087	28.2099
1.00	5.8716	5.4771	5.1856	5.8417	12.4204	20.6839	25.8148
5.00	3.1546	3.4285	4.6822	5.8738	13.4777	19.6350	24.2024
10.00	4.6287	5.0376	6.8458	8.3462	15.6907	19.7077	24.2024
50.00	11.0635	11.2055	11.8293	12.2245	15.8395	19.7077	24.2024
0.05				0.6874	0.6307	0.5138	0.6280
0.10			0.5334	0.6395	0.5255	0.4373	0.5411
0.50	0.1425	0.1328	0.1046	0.0958	0.1404	0.2237	0.2934
1.00	0.0611	0.0570	0.0539	0.0608	0.1292	0.2151	0.2685
5.00	0.0328	0.0357	0.0487	0.0611	0.1402	0.2042	0.2517
10.00	0.0481	0.0524	0.0712	0.0868	0.1632	0.2050	0.2517
50.00	0.1151	0.1166	0.1230	0.1272	0.1648	0.2050	0.2517

	0.05	0.10	0.50	1.00	5.00	10.00	50.00
0.05					87.3942	80.7942	60.3747
0.10					73.0993	58.1447	52.0189
0.50		17.4980	14.8712	13.5323	9.1307	10.1273	28.2099
1.00	7.7535	7.4459	6.7447	6.3295	5.9550	7.7938	25.8148
5.00	2.3915	2.6020	3.3330	3.7256	5.4025	8.3263	24.2024
10.00	3.4154	3.6975	4.5959	5.1275	7.4074	10.9861	24.2024
50.00	10.6791	10.7798	11.1961	11.4939	12.2361	13.8879	24.2024
0.05					0.9090	0.8404	0.6280
0.10					0.7604	0.6048	0.5411
0.50		0.1820	0.1547	0.1408	0.0950	0.1053	0.2934
1.00	0.0806	0.0774	0.0702	0.0658	0.0619	0.0811	0.2685
5.00	0.0249	0.0271	0.0347	0.0388	0.0562	0.0866	0.2517
10.00	0.0355	0.0385	0.0478	0.0533	0.0771	0.1143	0.2517
50.00	0.1111	0.1121	0.1165	0.1196	0.1273	0.1445	0.2517

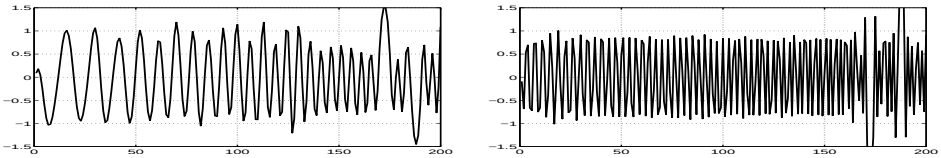
Table 4: Normalized mean square error for different values of η_1 (different lines) and ε_1 (different columns), $p = 1$, $q = 2$(top) and $q = 1$ (bottom). The light gray cases correspond to results better than EMD and OMD while the dark gray cases match the results only better than OMD. The black cases denote the unfeasible solutions. It corresponds to the decomposition result of a sum of a rectangular trend and a chirp signal. The top table presents the normalized MSE for d_1 while the bottom table lists the results for a_1.



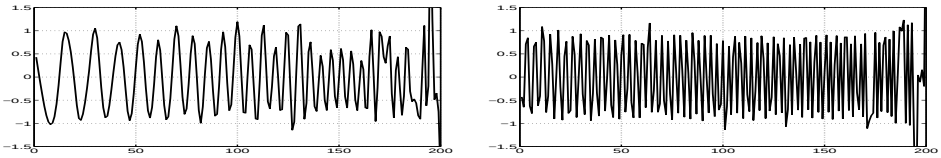
Original signal x | Spectrogram



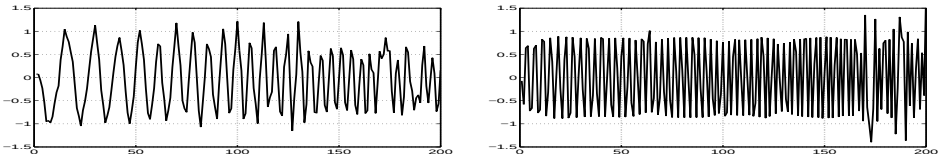
Original signal a_1 | Original signal d_1



EMD a_1 (EQM = 0.275) | EMD d_1 (EQM = 0.273)



OMD a_1 (EQM = 0.378) | OMD d_1 (EQM = 0.376)



Seq-VMD a_1 (**EQM = 0.187**) | Seq-VMD d_1 (**EQM = 0.186**)

Figure 6: Decomposition of a sum of two chirps.

	0.05	0.10	0.50	1.00	5.00	10.00	50.00
0.05	0.3849	0.3871	0.4026	0.4161	0.5016	0.5664	0.9944
0.10	0.3835	0.3857	0.4014	0.4150		0.6756	1.0505
0.50	0.3686	0.3712	0.3881	0.4024	0.4702	0.4724	0.9210
1.00	0.3452	0.3481	0.3654	0.3795	0.4630	0.5317	0.7564
5.00	0.2362	0.2496	0.1775	0.2025	0.2839	0.3928	0.3928
10.00	0.1491	0.1618	0.2017	0.2137	0.2895	0.3711	0.3711
50.00	0.3435	0.3459	0.3613	0.3746	0.4291	0.4981	0.4981
0.05	0.9981	0.9986	1.0000	1.0003	1.0443	1.0631	1.0021
0.10	0.9931	0.9936	0.9951	0.9955		0.9836	1.0586
0.50	0.9437	0.9445	0.9462	0.9455	0.9760	0.9158	0.9280
1.00	0.8674	0.8683	0.8679	0.8634	0.7508	0.5357	0.7622
5.00	0.2812	0.3039	0.1788	0.2040	0.2861	0.3958	0.3958
10.00	0.1503	0.1631	0.2033	0.2153	0.2917	0.3740	0.3740
50.00	0.3462	0.3486	0.3641	0.3775	0.4324	0.5019	0.5019

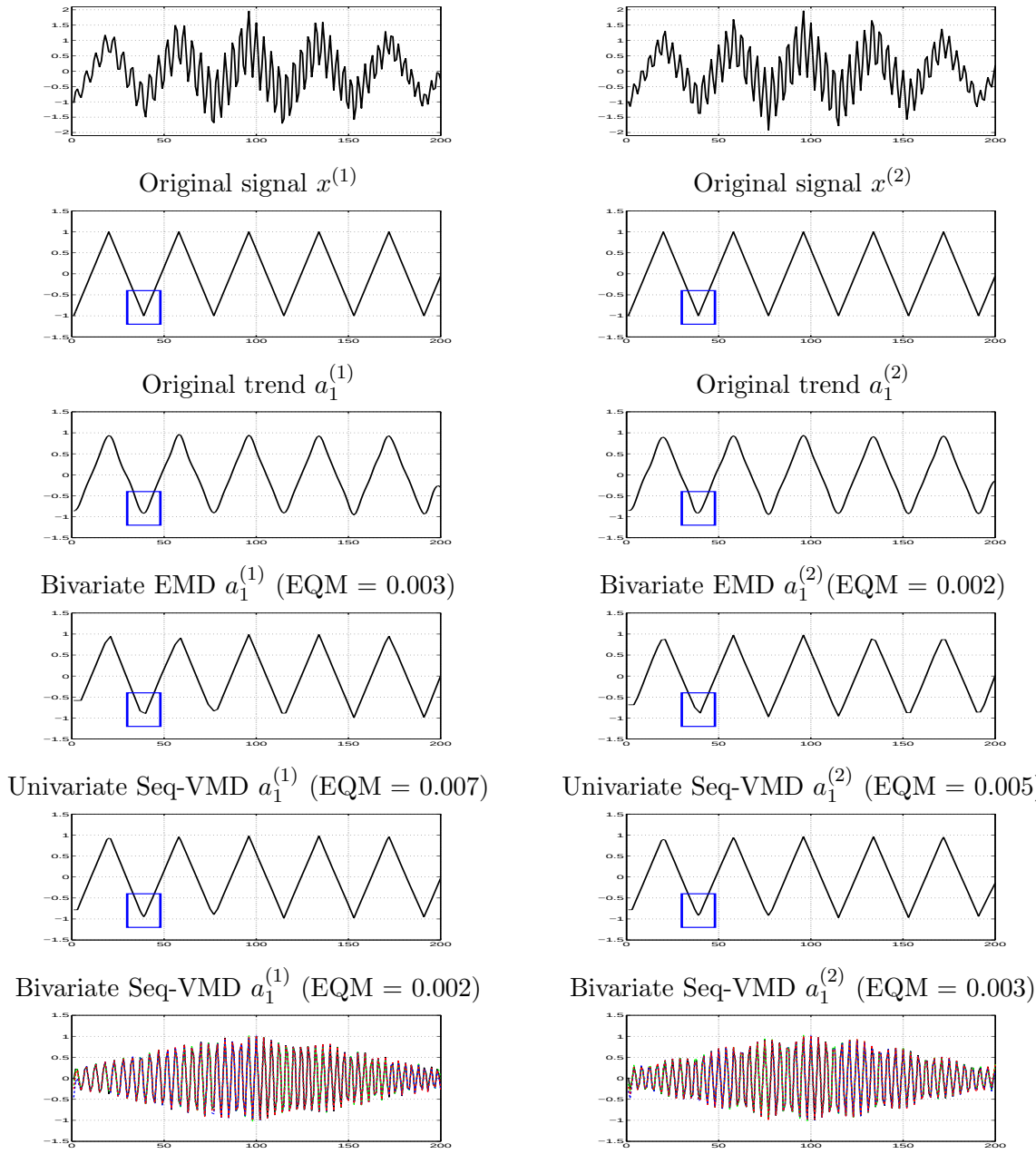
	0.05	0.10	0.50	1.00	5.00	10.00	50.00
0.05							0.9912
0.10							0.9999
0.50							0.9210
1.00							0.7564
5.00					0.1975	0.2261	0.3928
10.00	0.1470	0.1481	0.1571	0.1649	0.2231	0.2519	0.3711
50.00	0.3420	0.3431	0.3464	0.3501	0.3807	0.4085	0.4981
0.05							0.9988
0.10							1.0075
0.50							0.9280
1.00							0.7622
5.00					0.1990	0.2278	0.3958
10.00	0.1481	0.1493	0.1583	0.1662	0.2248	0.2538	0.3740
50.00	0.3446	0.3457	0.3490	0.3528	0.3836	0.4116	0.5019

Table 5: Normalized mean square error for different values of η_1 (different lines) and ε_1 (different columns), $p = 2$, $q = 2$(top) and $q = 1$ (bottom). The light gray cases correspond to results better than EMD and OMD while the dark gray cases match the results only better than OMD. The black cases denote the unfeasible solutions. It corresponds to the decomposition result of a sum of two chirps. The top table presents the normalized MSE for d_1 while the bottom table lists the results for a_1.

	0.05	0.10	0.50	1.00	5.00	10.00	50.00
0.05							0.9851
0.10							0.9841
0.50							0.9502
1.00							0.8789
5.00						0.5157	0.5650
10.00					0.3288	0.3989	0.3989
50.00	0.3214	0.3272	0.3446	0.3606	0.4203	0.4912	0.4912
0.05							0.9926
0.10							0.9917
0.50							0.9575
1.00							0.8856
5.00						0.5196	0.5693
10.00					0.3313	0.4020	0.4020
50.00	0.3238	0.3297	0.3472	0.3634	0.4235	0.4950	0.4950

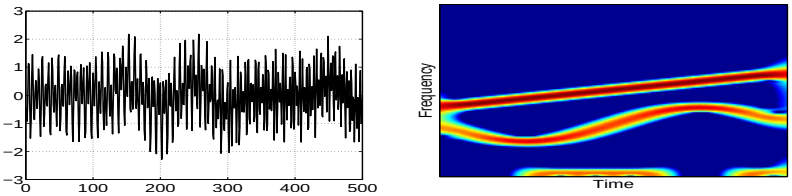
	0.05	0.10	0.50	1.00	5.00	10.00	50.00
0.05	0.3849	0.3871	0.4026	0.4161	0.5016	0.5664	0.9944
0.10	0.3835	0.3857	0.4014	0.4150		0.6756	1.0505
0.50	0.3686	0.3712	0.3881	0.4024	0.4702	0.4724	0.9210
1.00	0.3452	0.3481	0.3654	0.3795	0.4630	0.5317	0.7564
5.00	0.2362	0.2496	0.1775	0.2025	0.2839	0.3928	0.3928
10.00	0.1491	0.1618	0.2017	0.2137	0.2895	0.3711	0.3711
50.00	0.3435	0.3459	0.3613	0.3746	0.4291	0.4981	0.4981
0.05	0.9981	0.9986	1.0000	1.0003	1.0443	1.0631	1.0021
0.10	0.9931	0.9936	0.9951	0.9955		0.9836	1.0586
0.50	0.9437	0.9445	0.9462	0.9455	0.9760	0.9158	0.9280
1.00	0.8674	0.8683	0.8679	0.8634	0.7508	0.5357	0.7622
5.00	0.2812	0.3039	0.1788	0.2040	0.2861	0.3958	0.3958
10.00	0.1503	0.1631	0.2033	0.2153	0.2917	0.3740	0.3740
50.00	0.3462	0.3486	0.3641	0.3775	0.4324	0.5019	0.5019

Table 6: Normalized mean square error for different values of η_1 (different lines) and ε_1 (different columns), $p = 1$, $q = 2$(top) and $q = 1$ (bottom). The light gray cases correspond to results better than EMD and OMD while the dark gray cases match the results only better than OMD. The black cases denote the unfeasible solutions. It corresponds to the decomposition result of a sum of two chirps. The top table presents the normalized MSE for d_1 while the bottom table lists the results for a_1.

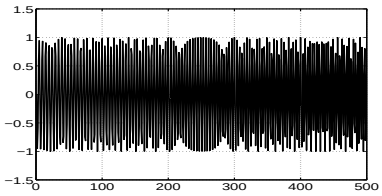


Superposition of the original IMF (solid black) with 3 plots that represent the extracted IMF of order 1 with bivariate EMD (dashed green), univariate Seq-VMD (dashed blue), and bivariate Seq-VMD (dashed red): $a_1^{(1)}$ (left) and $a_1^{(2)}$ (right)

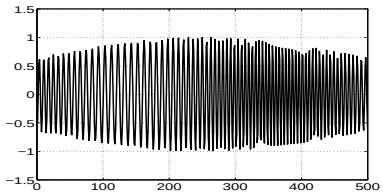
Figure 7: Bivariate decomposition.



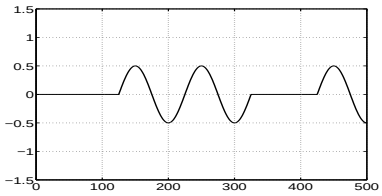
Original signal x (left) and Spectrogram (right)



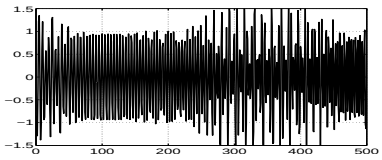
Original signal d_1



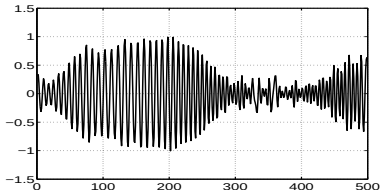
Original signal d_2



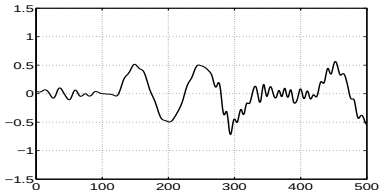
Original signal a_2



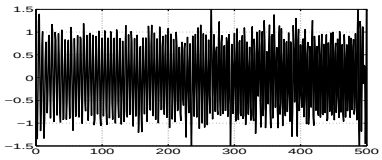
EMD d_1 (EQM = 0.249)



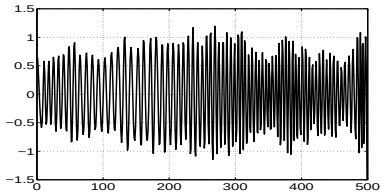
EMD d_2 (EQM = 0.316)



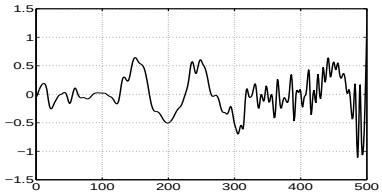
EMD a_2 (EQM = 0.054)



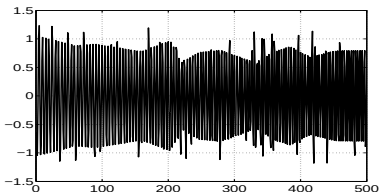
OMD d_1 (**EQM = 0.133**)



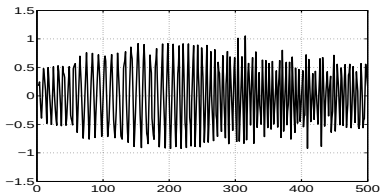
OMD d_2 (EQM = 0.202)



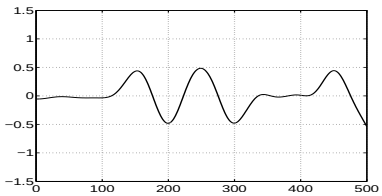
OMD a_2 (EQM = 0.474)



Seq-VMD d_1 (EQM = 0.143)

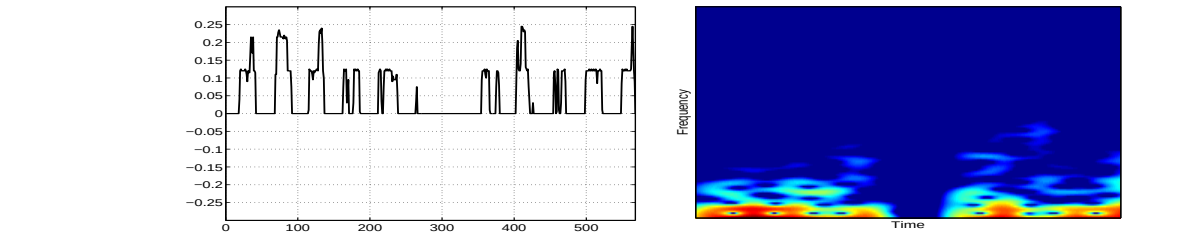


Seq-VMD d_2 (**EQM = 0.183**)

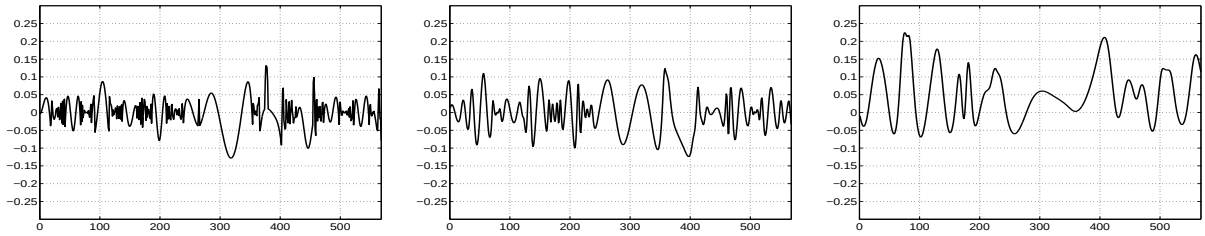


Seq-VMD a_2 (**EQM = 0.022**)

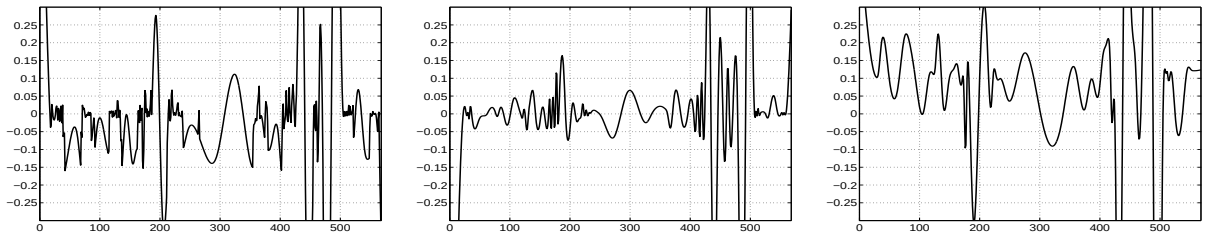
Figure 8: Decomposition of a 3 components signal.



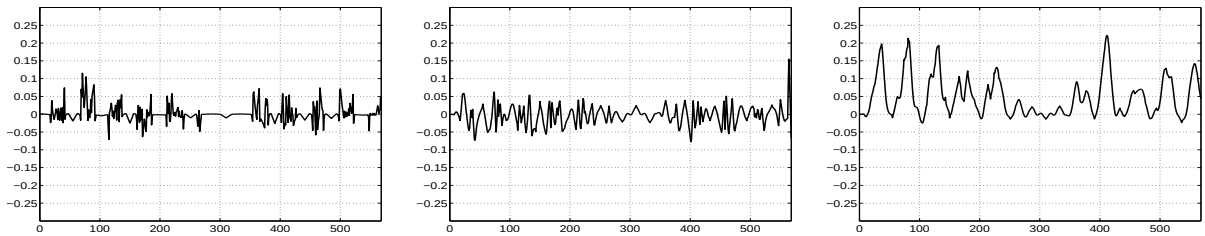
Original signal x (left) and Spectrogram (right)



EMD : d_1 (left), d_2 (middle), a_2 (right)



OMD : d_1 (left), d_2 (middle), a_2 (right)



Seq-VMD : d_1 (left), d_2 (middle), a_2 (right)

Figure 9: Real data decomposition.