Oscillatory + Transient Signal Decomposition using Overcomplete Rational-Dilation Wavelet Transforms

Ivan W. Selesnick[a] and İlker Bayram[b]

[a]Polytechnic Institute of New York University, Brooklyn, New York, USA
[b]Biomedical Imaging Group, Swiss Federal Institute of Technology, Lausanne, Switzerland

ABSTRACT

This paper describes an approach for decomposing a signal into the sum of an oscillatory component and a transient component. The method uses a newly developed rational-dilation wavelet transform (WT), a self-inverting constant-Q transform with an adjustable Q-factor (quality-factor). We propose that the oscillatory component be modeled as signal that can be sparsely represented using a high Q-factor WT; likewise, we propose that the transient component be modeled as a piecewise smooth signal that can be sparsely represented using a low Q-factor WT. Because the low and high Q-factor wavelet transforms are highly distinct (having low coherence), morphological component analysis (MCA) successfully yields the desired decomposition of a signal into an oscillatory and non-oscillatory component. The method, being non-linear, is not constrained by the limits of conventional LTI filtering.

Keywords: wavelets, sparsity, morphological component analysis, constant Q, Q-factor

1. INTRODUCTION

Many natural signal are comprised of temporally overlapping oscillatory and transient components. For example, EEG signals are composed of the superposition of: 1) a rhythmic signal, and 2) a transient signal. The rhythmic components exist in several frequency bands. For example, 'alpha' is the rhythmic activity in the 8-12 Hz range, 'theta' is the 4-7 Hz range, etc. The transient component comes from unwanted measurement artifacts and from non-rhythmic brain activity (spikes, spindles, and vertex waves of varying amplitude and duration). Note that the rhythmic component may have both low and high frequencies. Likewise, the transient component is not well contained in any single frequency band. The rhythmic and transient components overlap in both time and frequency.

The alpha activity is normally extracted from an EEG signal by filtering the EEG signal with an appropriately designed band-pass filter. However, the alpha waveform obtained in this way may be corrupted by the presence of the transient component. Indeed, transients have energy at a broad range of frequencies including the pass-band of the band-pass filter used. Filtering a transient with a band-pass filter designed to pass 8-12 Hz will produce an output signal consisting of several oscillations having frequency 8-12 Hz. Therefore, when band-pass filtering an EEG to obtain alpha activity, a transient will produce a spurious short-duration alpha-like waveform in the output. The phenomenon can not be avoided as long as one uses linear time-invariant filters.

In this paper, we describe a method for the separation of the oscillatory and transient components of a signal. The proposed method uses morphological component analysis (MCA)[23,24,42–44] together with new wavelet (constant-Q) transforms.[1] The method is non-linear, so it is not constrained by the limits of LTI filters.

The proposed signal decomposition method may be useful as an EEG preprocessing step. After the separation is accomplished, the rhythmic component can be filtered (using a conventional LTI band-pass filter) to obtain an improved measurement of the alpha activity, for example. That in turn is useful for other basic procedures in EEG signal analysis, for example, tracking how the power in a specific frequency band varies as a function

Author Email: selesi@poly.edu, ilker.bayram@epfl.ch

of time. An improved measurement of the rhythmic component may also benefit the investigation of changes in synchrony between different parts of the brain as a subject performs mental tasks.

We note that notching out the dc component of an EEG signal using a dc-notch filter does not achieve the sought separation. Such filtering can be used to remove the base-line drift; however, spikes, step-like jumps, and other transients, will remain. For our problem — the decomposition of an signal into transient and rhythmic components — neither a frequency-domain nor a time-domain method alone is adequate. Even though the alpha activity (for example) is confined to a specific frequency band, the transients are not; therefore, conventional (LTI) filtering cannot solve the problem. Similarly, time-domain methods that rely on detecting and removing spikes (e.g. Ref. 36) according to a set of templates must account for a variation in transient morphology (shape); therefore, such methods are limited to a specific set of transients. In our approach, we model the transient component as a generic piecewise smooth signal.

The decomposition of a signal into transient and non-transient components has been addressed by several authors, especially for audio coding and processing.[18,19,25,34,35] In Ref. 34, it is proposed that sines, transients, and noise be coded separately for efficient audio compression. The separation of transient and tonal components using sparse signal representations with the wavelet transform and the modified discrete cosine transform (MDCT) is described in Refs. 18, 35 for audio compression. Note that the MDCT, used in Refs. 18, 35 to model the tonal part, is not a constant-Q transform; and that the decomposition method is different than MCA. In Ref. 19, an algorithm is described for the sparse representation of audio using overcomplete modulated complex lapped transform (MCLT) dictionaries. Ref. 17 introduces the molecular matching pursuit algorithm for the efficient decomposition of audio into a tonal part (as clusters of MDCT atoms), a transient part (clusters of wavelet atoms), and a residual. In Ref. 25, a fully Bayesian approach for audio signal representation is described, and an example given where a MDCT basis with long time resolution is used to model the tonal component and a MDCT with short time resolution is used to model the transient component. Ref. 32 introduces the time-frequency jigsaw puzzle and provides an example of its use for tonal/transient separation in audio. Distinct from these approaches, is the time-scale method of Ref. 6, which introduces an algorithm for separating transients that exploits the multi-scale characterization of transients.

1.1 Morphological Component Analysis

In this paper we will utilize morphological component analysis (MCA).[42] Given an observed signal $\mathbf{x}$ that is the superposition of two signals,

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2, \qquad \mathbf{x},\ \mathbf{x}_1,\ \mathbf{x}_2 \in \mathbb{R}^N,$$

MCA provides an approach to recover the components $\mathbf{x}_1$ and $\mathbf{x}_2$ individually. MCA utilizes a flexible and effective signal model: the sparsity of $\mathbf{x}_1$ with respect to one signal representation, and the sparsity of $\mathbf{x}_2$ with respect to a second, substantially different, signal representation. MCA relies on each of $\mathbf{x}_1$ and $\mathbf{x}_2$ having a sparse representation; otherwise, MCA is not applicable. The signal model is based on the specification of the two signal representations, and relies on the two representations providing sparse representations only for substantially different types of signals (signal morphologies).

If $\mathbf{x}_1$ is represented as a weighted sum of the columns of a matrix $\mathbf{S}_1$, and $\mathbf{x}_2$ is represented using $\mathbf{S}_2$ likewise, then the MCA approach, in one form (the *synthesis* form), asks that the objective function,

$$J(\mathbf{a}_1, \mathbf{a}_2) = \|\mathbf{x} - \mathbf{S}_1\mathbf{a}_1 - \mathbf{S}_2\mathbf{a}_2\|_2^2 + \lambda_1\,\mathrm{sparsity}(\mathbf{a}_1) + \lambda_2\,\mathrm{sparsity}(\mathbf{a}_2) \tag{1}$$

be minimized with respect to $\mathbf{a}_1$ and $\mathbf{a}_2$, where sparsity($\mathbf{u}$) denotes a measure of sparsity of the vector $\mathbf{u}$. Then the estimated components, $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$ are given by

$$\hat{\mathbf{x}}_1 = \mathbf{S}_1\mathbf{a}_1, \qquad \hat{\mathbf{x}}_2 = \mathbf{S}_2\mathbf{a}_2, \tag{2}$$

where ($\mathbf{a}_1$, $\mathbf{a}_2$) minimizes the objective function (1). The sparsity measure could be, for example, the ℓ_0-norm, in which case sparsity($\mathbf{u}$) = $\|\mathbf{u}\|_0$. Because the ℓ_0-norm is not convex, the ℓ_1 norm is often used instead; in which case sparsity($\mathbf{u}$) = $\|\mathbf{u}\|_1$. More general forms of MCA allow the sparsity measures for $\mathbf{x}_1$ and $\mathbf{x}_2$ in

(1) to be different from each other; additionally, the data fidelity term need not be an ℓ_2 norm. Examples in Refs. 23, 24, 42–44, illustrate the use of MCA with several transforms playing the role of $\mathbf{S}_1$ and $\mathbf{S}_2$: the curvelet transform, the block DCT, and wavelet transforms. In this case, $\mathbf{S}_1$ and $\mathbf{S}_2$ represent inverse transforms.

Another form of MCA uses a slightly different formulation. In this form (the *analysis* form), MCA asks that the objective function,

$$J(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{x} - \mathbf{x}_1 - \mathbf{x}_2\|_2^2 + \lambda_1 \,\text{sparsity}(\mathbf{A}_1\mathbf{x}_1) + \lambda_2 \,\text{sparsity}(\mathbf{A}_2\mathbf{x}_2) \tag{3}$$

be minimized with respect to $\mathbf{x}_1$ and $\mathbf{x}_2$, where again, sparsity($\mathbf{u}$) denotes a measure of sparsity of the vector $\mathbf{u}$. The matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ represent forward transforms.

The success of MCA relies on $\mathbf{x}_1$ and $\mathbf{x}_2$ having sparse representations (few non-zero coefficients) in their respective representations. Iterative algorithms for (approximately) solving the MCA optimization problem are given in Ref. 42. Appendix B describes simple algorithms for minimizing the two objective functions where sparsity is measured using the ℓ_1 norm.

In the decomposition of an EEG signal into rhythmic and transient components, $\mathbf{x}$ is the EEG signal, $\mathbf{x}_1$ is the rhythmic component, and $\mathbf{x}_2$ is the transient component. We need signal representations providing sparse representations of the rhythmic and transient components respectively. We propose that newly developed wavelet frames based on rational (non-dyadic) dilation factors[1] be used for both.

1.2 Constant-Q Transforms

The dyadic wavelet transform (WT) has become a well recognized tool in signal processing for image coding, noise reduction, deconvolution, and other applications. The success of the WT is due to its ability to sparsely represent piecewise smooth signals (for example, a scan line from a typical photographic image). However, the WT is used much less often for audio signals and other signals that, like audio, are quasi-periodic over short-time intervals. For such signals, the dyadic WT does not provide a sparse representation. Indeed, audio coding generally uses signal decompositions (the MDCT, etc) that have a much finer frequency resolution than the conventional dyadic WT.

However, the WT is attractive because it is a constant-Q transform where Q denotes 'quality factor'. The quality-factor (Q-factor) of a filter is the ratio of its center frequency to its bandwidth. The frequency resolution of a constant-Q transform varies with frequency. We suggest that the dyadic WT fails to provide a sparse representation of 'short-time quasi-periodic' signals (like speech) because of its very low Q-factor; that is, because it has very poor frequency resolution. Transforms like the MDCT or short-time Fourier transform, with uniform frequency bands, usually work better for such signals than the dyadic WT, but they are not constant-Q transforms. The low-frequency subbands have low quality factors, while the high-frequency subbands have high quality factors.

The design and implementation of an easily invertible discrete WT that allows high quality factors (equivalently, good frequency resolution) was one of our motivations for developing overcomplete wavelet transforms based on rational (non-dyadic) dilation factors.[1]

For the sparse representation of the transient component of a signal we propose to use a WT with a low quality factor (like the conventional dyadic WT) because the transient component can be modeled as a piecewise smooth signal. On the other hand, for the sparse representation of the oscillatory (rhythmic) component of a signal we propose to use a WT with a high quality factor (which must be a non-dyadic rational WT) because the rhythmic component is quasi-periodic over short-time intervals.

2. RATIONAL-DILATION WAVELET TRANSFORMS

The continuous wavelet transform (CWT) computes the inner products of a signal with $\{\sqrt{a}\,\psi(at - b)\}$ where a and b are allowed to change continuously.[14] As such, it maps a function on the real line to a function on the plane (the variables being a and b). Using the CWT, one can choose the wavelet function quite freely (it may have a low or high quality factor). For instance, all of the functions in Fig. 1 qualify as admissible wavelet functions for the CWT. Note that varying the scale parameter 'a' changes the frequency of $\sqrt{a}\,\psi(at - b)$, but

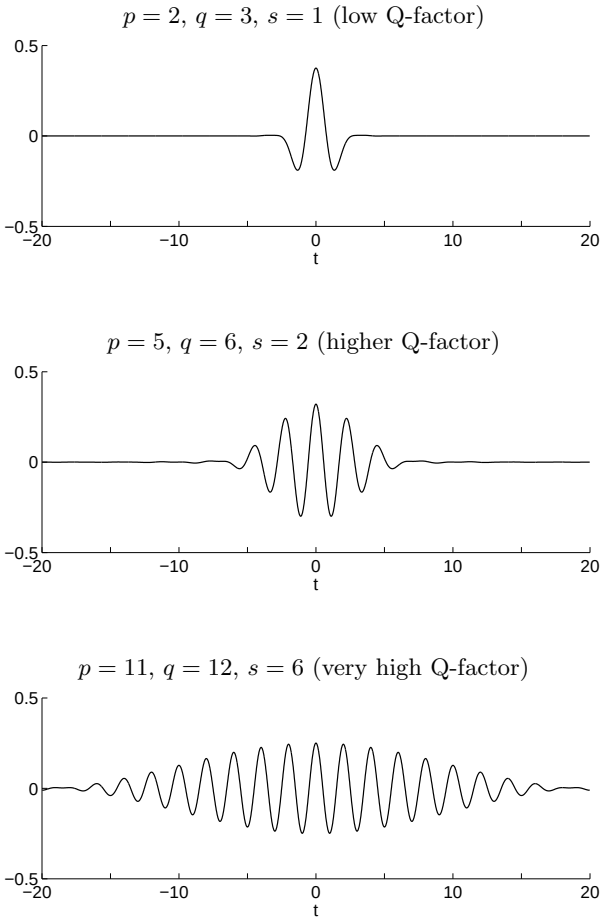

Figure 1. Wavelets of the rational-dilation wavelet transform (Ref. 1) for several parameter values.

preserves its shape and its Q-factor. The wavelet transform is therefore a constant-Q transform. However, the CWT is highly redundant and an implementation for discrete-time data is not easily inverted. (An efficient, stable inverse is important when performing MCA.)

Letting $\{a, b\}$ take the values $\{2^n, k\}_{n,k\in\mathbb{Z}}$ and using a valid function $\psi(t)$, one obtains the orthonormal (critically-sampled) dyadic wavelet basis. Because 'a' is only allowed to be powers of 2, the wavelet dilation from scale to scale is not always as gradual as might be desired. Moreover, the dyadic WT is highly shift-varying in the sense that a slight shift of the input causes the wavelet coefficients to vary substantially. Lastly, orthonormality and the restriction of 'a' to 2^n constrains the allowed wavelet functions to have low Q-factors.

To alleviate these issues, we have recently introduced a new family of overcomplete wavelet (constant-Q) transforms with a flexible range of quality factors and redundancy factors.[1] Specifically, they allow $\{a, b\}$ to take values from $\{q^n/p^n, spk/q\}_{n,k\in\mathbb{Z}}$ where p, q and s are integers. These overcomplete transforms (frames) are redundant by a factor of $\frac{1}{s}\frac{1}{q/p-1}$ and have several attractive properties. These transforms are: developed for discrete-time data, nearly shift-invariant, dilate the wavelet more gradually from scale to scale, permit the design of wavelets with both low and high quality factors as illustrated in Fig. 1, and are easily invertible (in fact, self-inverting). Because it has a simple, stable inverse (unlike the CWT) it can be easily used for MCA.

We use the new rational-dilation wavelet transform with low quality factor to model the transient component. The transient component has previously been modeled using the dyadic (low Q-factor) WT.[38] The new rational WT with a low Q-factor is similar, however, it allows for a more gradual dilation of the wavelet function from scale to scale. That is beneficial because some transients may be best matched to dilations of the wavelet $\psi(t)$ falling between the dyadic scales.[12,13]

In contrast to the transient signal, the rhythmic signal contains quasi-periodic structures of varying duration, calling for a transform with better frequency resolution (higher quality factor). The MDCT or STFT could be used; however, the analysis/synthesis functions of those transforms are all of the same duration — the low frequency analysis/synthesis functions may not have sufficient frequency resolution, while the high frequency analysis/synthesis may not have sufficient time resolution. For that reason, we utilize the new rational-dilaiton rational WT with higher Q value to model the rhythmic (oscillatory) component.

3. EXAMPLES

Example 1: To illustrate the decomposition of a signal into an oscillatory component and a transient component, we construct the test signal $x(n)$ illustrated in Fig. 2. The test signal consists of four pulses, non-overlapping in time, generated in MATLAB using the commands:

```
N1 = 80; x1 = cos(0.20*pi*(1:N1)) .* hamming(N1)';
N2 = 10; x2 = cos(0.50*pi*(1:N2)) .* hamming(N2)';
N3 = 20; x3 = cos(0.20*pi*(1:N3)) .* hamming(N3)';
N4 = 45; x4 = cos(0.42*pi*(1:N4)) .* hamming(N4)';
x = [x1 x2 x3 x4];
```

The first and last pulses are oscillatory. The second and third pulses we consider transient, in the sense that these pulses are less oscillatory than the first and last pulses. Our goal is to decompose the signal $x(n)$ into $x_1(n) + x_2(n)$ where $x_1(n)$ contains the oscillatory component (pulse 1 and 4), and where $x_2(n)$ contains the transient component (pulse 2 and 3). The decomposition can not be performed using low-pass and high-pass filters because pulses 1 and 3 are of the same frequency (0.2π) and because pulses 2 and 4 are of nearly the same frequency. However, using the rational-dilation wavelet transform with different Q-factors, the test signal can be decomposed as desired (approximately). Specifically, we use the rational-dilation wavelet transform with parameters ($p = 5$, $q = 6$, $s = 2$) to represent the oscillatory component; this is a 'high Q-factor' wavelet transform; it is denoted as $\mathbf{A}_1$. Similarly, we use the rational-dilation wavelet transform with parameters ($p = 2$, $q = 3$, $s = 1$) to represent the transient component; this is the 'low Q-factor wavelet transform; it is denoted as $\mathbf{A}_2$. The discrete-time wavelet functions are illustrated in Fig. 3. To perform MCA, we used formulation (3) with ℓ_1-norm regularization, namely (11). We set $\lambda_1 = 0.13$ and $\lambda_2 = 0.1$. To minimize (11) we used 400 iterations of algorithm (22)-(28) initialized with all-zero signals. The value of the objective function, $J(\mathbf{x}_1, \mathbf{x}_2)$ is illustrated in Fig. 2 for each iteration.

Example 2: Like Example 1, we construct a test signal consisting of two oscillatory pulses (at two frequencies) and two transient (non-oscillatory) pulses (also at two frequencies). However, in this example, the pulses overlap in time as illustrated in Fig. 4. The test signal is constructed in MATLAB with the command

```
x = [x1 x4] + [zeros(1,40)  x2 zeros(1,32) x3 zeros(1,23)];
```

where `x1`, `x2`, `x3`, and `x4` are as constructed in Example 1. We use the same two rational-dilation wavelet transforms as in Example 1 and the same λ_i. Fig. 4 illustrates the decomposition of the test signal into oscillatory and non-oscillatory components, obtained after 400 iterations of algorithm (22)-(28). Although the decomposition is not exact, it exhibits the sought behaviour.

Example 3 (EEG): Figure 5 illustrates the decomposition of a 10-second EEG signal into rhythmic and transient components using MCA and the new rational WT.[1] The signal is sampled at 200 samples per second. To represent the rhythmic component we use an 18-level rational WT with $p = 5$, $q = 6$, $s = 2$. For the transient component, we use the a 9-level rational WT with $p = 2$, $q = 3$, $s = 1$. The wavelets are illustrated in the first and second panels of Fig. 1. The obtained rhythmic component is quasi-periodic over short time durations. In addition, the obtained transient component is almost perfectly piecewise linear (the utilized wavelet has two vanishing moments). It turns out that for this example, the empirical mode decomposition (EMD)[29,31] can be used to produce a result very similar to that shown in Fig. 5*. Grouping the intrinsic mode functions (IMFs) 2 to 4 gives a rhythmic component, and grouping modes 5 to 10 gives a transient component.

*Personal communication: Patrick Flandrin, École Normale Supérieure de Lyon, France.

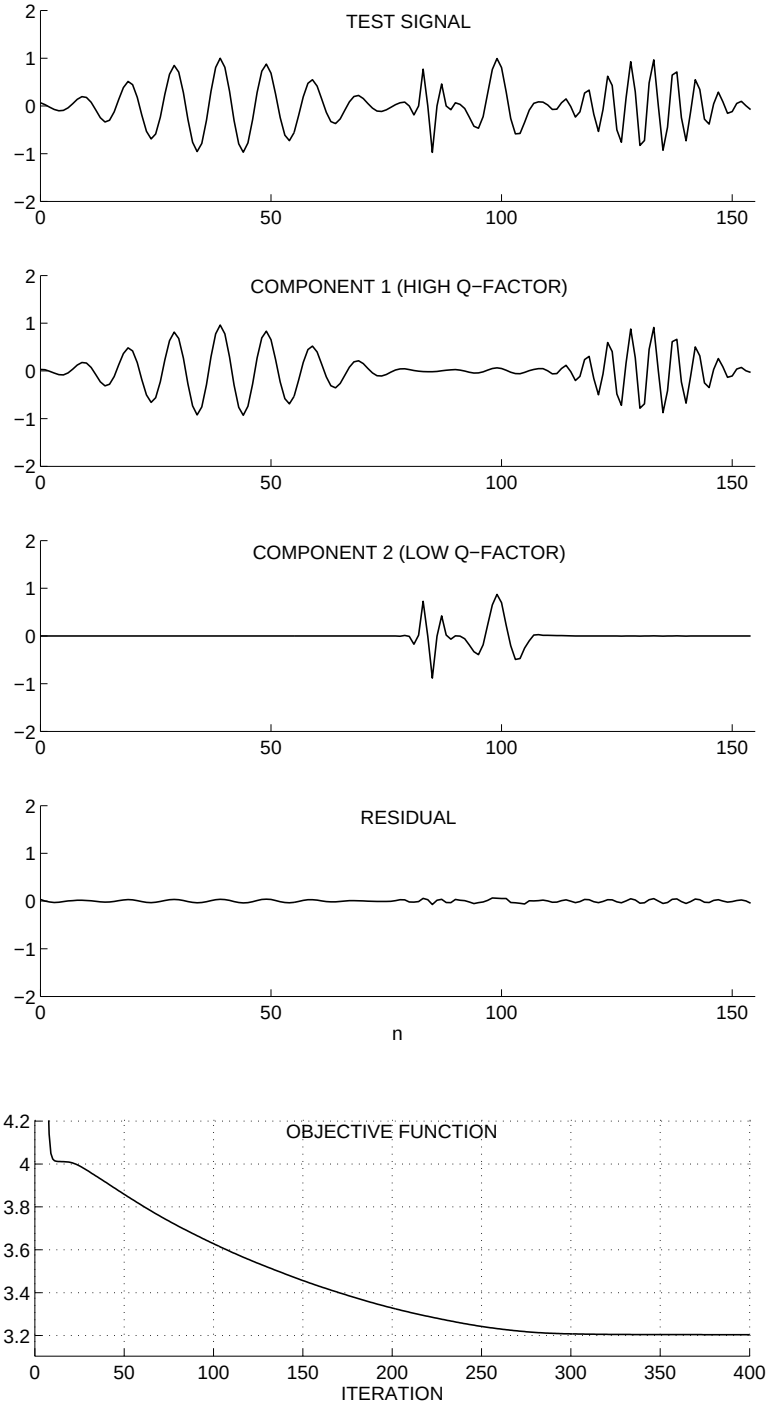

Figure 2. Example 1: Decomposition of a test signal into an oscillatory (high Q-factor) component and a transient (low Q-factor) component.

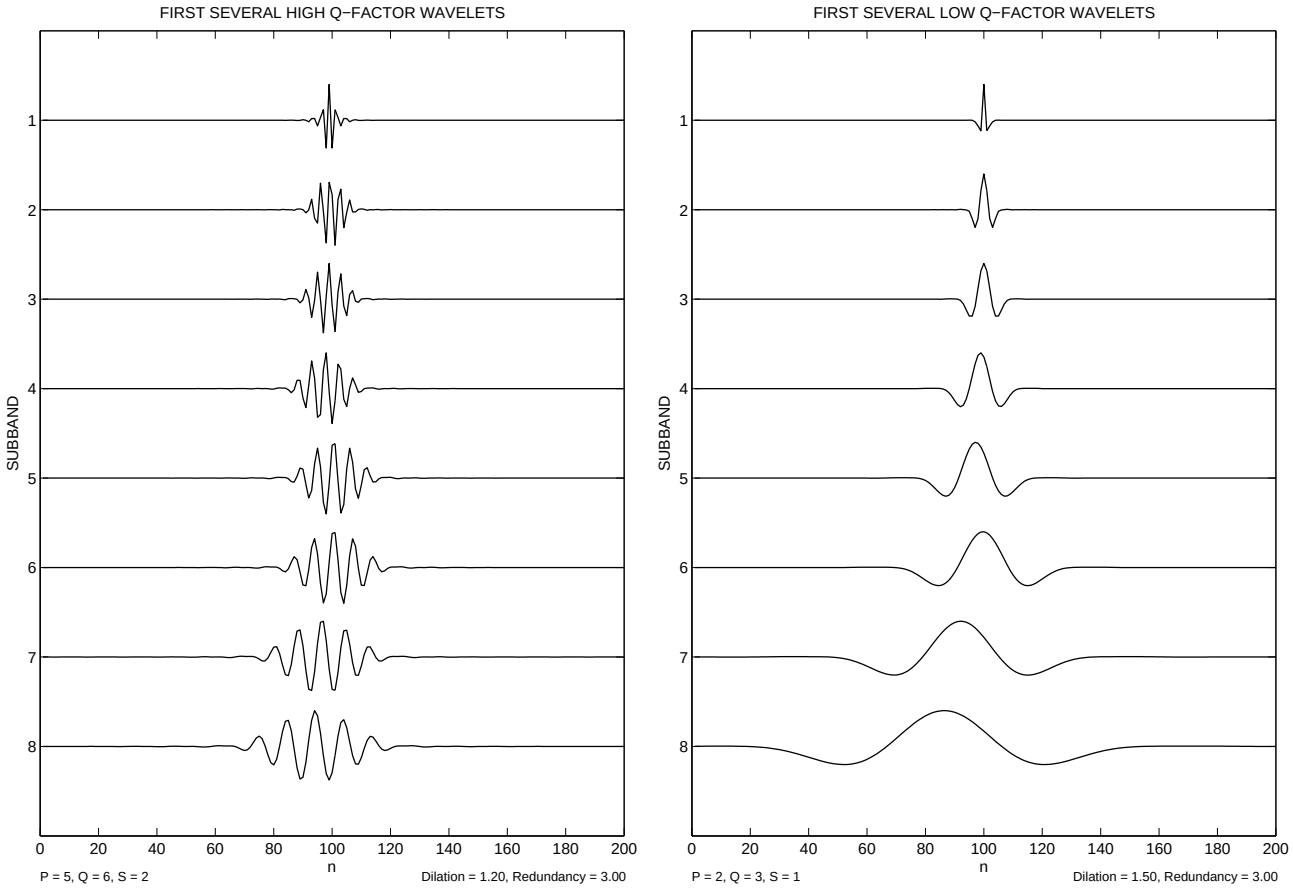

Figure 3. Wavelets at first several scales as implemented by the fully-discrete rational-dilation wavelet transform. High Q-factor and low Q-factor wavelets are obtained with different transform parameters; see Ref. 1.

4. CONCLUSION

In this paper we model the transient component of a signal (for example, EEG) as a piecewise smooth signal which can be sparsely represented using a low Q-factor wavelet transform. We model the oscillatory (rhythmic) component as a signal that can be sparsely represented using a high Q-factor wavelet transform. We then propose that the transient and oscillatory components of a signal can be separated using these two representations in conjunction with morphological component analysis (MCA). This non-linear separation procedure may be used as a preprocessing step to reduce transients in a signal for more accurate spectral estimation.

5. ACKNOWLEDGMENT

IWS thanks Hans von Gizycki and Ivan Bodis-Wollner of SUNY-Downstate Medical Center, Brooklyn, NY, for many stimulating discussions on EEG signal analysis/interpretation.

APPENDIX A. TWO ℓ_1-NORM REGULARIZATION PROBLEMS

In this Appendix, we summarize two simple algorithms for ℓ_1-norm regularized signal restoration. Namely, Section A.1 considers deconvolution with an ℓ_1 sparsity prior; and Section A.2 considers denoising with a general linear ℓ_1 analysis prior. These two algorithms will be used in turn in Appendix B to describe algorithms for performing MCA using the ℓ_1-norm.

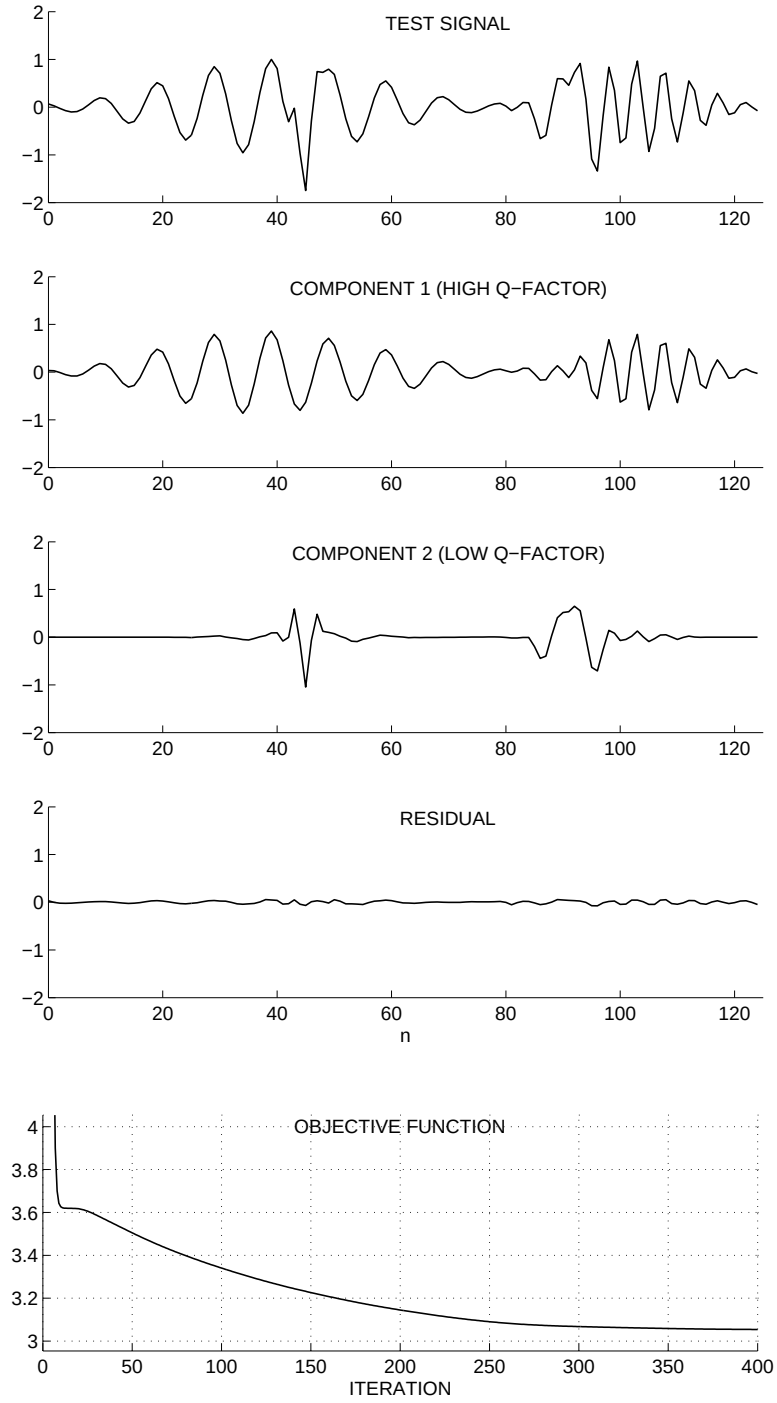

Figure 4. Example 2: Decomposition of a test signal into an oscillatory (high Q-factor) component and a transient (low Q-factor) component. The transient and oscillatory components overlap.

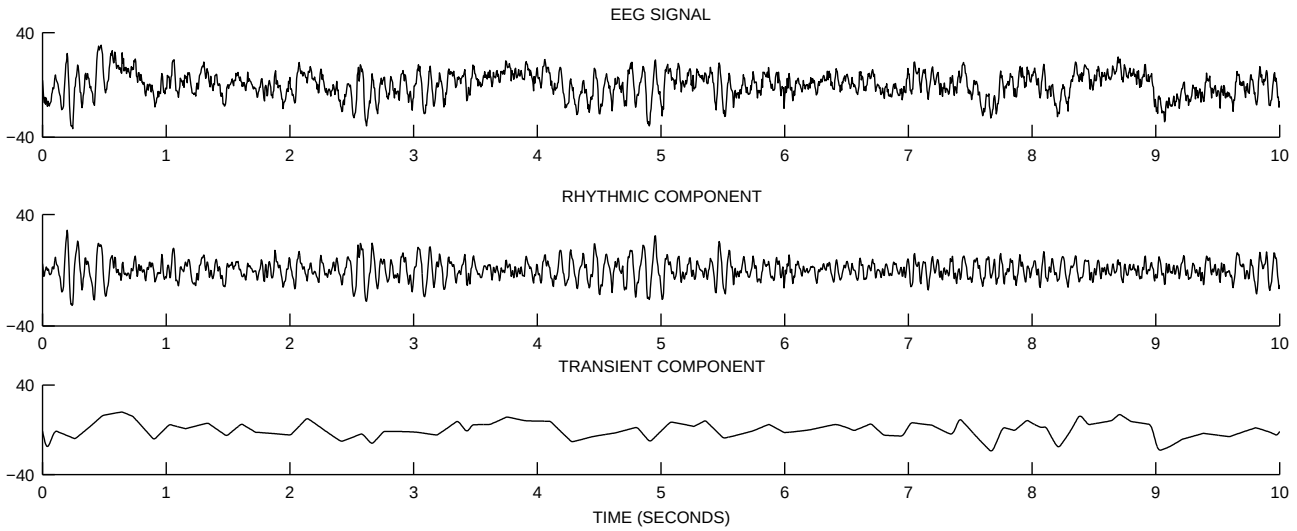

Figure 5. Example 3: Decomposition of an EEG signal into a sum of rhythmic and transient components using the proposed MCA/wavelet method.

A.1 ℓ_1-norm regularized deconvolution

An algorithm for minimizing the objective function

$$J(\mathbf{w}) = \|\mathbf{y} - \mathbf{H}\mathbf{w}\|_2^2 + \lambda\|\mathbf{w}\|_1 \tag{4}$$

is given by iterative thresholding:

$$\mathbf{w}^{(i+1)} = \mathrm{soft}\Big(\mathbf{w}^{(i)} + \frac{1}{\alpha}\mathbf{H}^t\,(\mathbf{y} - \mathbf{H}\mathbf{w}^{(i)}),\ \lambda/(2\alpha)\Big) \tag{5}$$

where the superscript i is the iteration index, and $\mathrm{soft}(x, T)$ is the soft-threshold function with threshold T,

$$\mathrm{soft}(x, T) := \mathrm{sign}(x)\max(0, |x| - T) \tag{6}$$

In (5), the soft-threshold function is applied element-wise. Setting $\alpha \geq \mathrm{maxeig}(\mathbf{H}^t\mathbf{H})$ ensures monotone convergence to the minimizer of (4).

Algorithm (5) is known as the *iterated soft-thresholding* (IST) algorithm or as the *thresholded-Landweber* (TL) algorithm.[15,16,26,27] For some problems this algorithm converges slowly. Numerous algorithms have been proposed having faster convergence properties.[3,5,9,21,28,33,46]

A.2 ℓ_1-norm regularized denoising using an analysis prior

An algorithm for minimizing the objective function

$$J(\mathbf{x}) = \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda\|\mathbf{A}\mathbf{x}\|_1. \tag{7}$$

is given by the algorithm:

$$\mathbf{z}^{(i+1)} = \Big(c\,\mathbf{z}^{(i)} + \mathbf{A}(\mathbf{b}^{(i)} - \mathbf{A}^t\,\mathbf{z}^{(i)})\Big)\,./\Big(\frac{2}{\lambda}\,|\mathbf{A}\mathbf{x}^{(i)}| + c\Big) \tag{8}$$

$$\mathbf{x}^{(i+1)} = \mathbf{y}^{(i)} - \mathbf{A}^t\,\mathbf{z}^{(i+1)} \tag{9}$$

where the superscript i is the iteration index. The operations ./ and $|\cdot|$ in (9) are element-wise division and element-wise absolute value. To ensure convergence, we can set $c > \mathrm{maxeig}(\mathbf{A}\mathbf{A}^t)$.

Algorithm (8)-(9) is a special case of Chambole's algorithm.[4,10] This algorithm can have slow asymptotic convergence and the development of faster algorithms for minimizing objective functions of this type is an active area of research.[2,8,30,37,40,45,47]

APPENDIX B. ALGORITHMS FOR MCA BASED ON THE ℓ_1 NORM

In the morphological component analysis (MCA) approach, the sought signal $\mathbf{x}$ is represented as a sum of two (or more) signals, $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$. The observed signal, $\mathbf{y}$, is given by $\mathbf{y} = \mathbf{x} + \text{noise}$. The goal is to recover both $\mathbf{x}_1$ and $\mathbf{x}_2$. The idea of MCA is that, if $\mathbf{x}_1$ and $\mathbf{x}_2$ are sparsely represented in distinct domains, then they can be (approximately) recovered by solving a suitable sparsity-regularlized linear inverse problem. MCA is developed in Refs. 23, 24, 42–44. As emphasized in Ref. 43, two approaches to MCA are (1) synthesis-prior formulation, and (2) the analysis-prior formulation.

MCA using an *synthesis* prior, asks that the objective function,

$$J(\mathbf{w}_1, \mathbf{w}_2) = \|\mathbf{y} - \mathbf{S}_1\mathbf{w}_1 - \mathbf{S}_2\mathbf{w}_2\|_2^2 + \lambda_1\|\mathbf{w}_1\|_1 + \lambda_2\|\mathbf{w}_2\|_1, \tag{10}$$

be minimized with respect to $\mathbf{w}_1$ and $\mathbf{w}_2$. Then $\mathbf{x}_1 = \mathbf{S}_1\mathbf{w}_1$ and $\mathbf{x}_2 = \mathbf{S}_2\mathbf{w}_2$.

MCA using an *analysis* prior, asks that the objective function,

$$J(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{y} - \mathbf{x}_1 - \mathbf{x}_2\|_2^2 + \lambda_1\|\mathbf{A}_1\mathbf{x}_1\|_1 + \lambda_2\|\mathbf{A}_2\mathbf{x}_2\|_1. \tag{11}$$

be minimized with respect to $\mathbf{x}_1$ and $\mathbf{x}_2$.

The distinctions between synthesis and analysis formulations of the sparsity-based approach for signal restoration has been discussed by several authors, see Refs. 7, 8, 11, 22, 39.

B.1 MCA with Synthesis Prior

In this section, we describe the derivation of a simple algorithm for the minimization of (10). The derivation uses the IST iteration in Appendix A.1. Note that (10) can be written as

$$J(\mathbf{w}_1, \mathbf{w}_2) = \|\mathbf{y} - \mathbf{H}\mathbf{w}\|_2^2 + \lambda_1\|\mathbf{w}_1\|_1 + \lambda_2\|\mathbf{w}_2\|_1$$

where

$$\mathbf{H} = \begin{bmatrix}\mathbf{S}_1 & \mathbf{S}_2\end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix}\mathbf{w}_1 \\ \mathbf{w}_2\end{bmatrix}.$$

Using majorization-minimization approach,[26] an iterative algorithm for finding the minimizer is:

$$\mathbf{b}^i = \mathbf{w}^i + \frac{1}{\alpha}\mathbf{H}^t(\mathbf{y} - \mathbf{H}\mathbf{w}^i)$$

$$\mathbf{w}^{i+1} = \underset{\mathbf{w}}{\operatorname{argmin}} \left[\alpha\|\mathbf{b}^i - \mathbf{w}\|_2^2 + \lambda_1\|\mathbf{w}_1\|_1 + \lambda_2\|\mathbf{w}_2\|_1\right]$$

where $\alpha > \text{maxeig}(\mathbf{H}^t\mathbf{H})$. Note that we can write $\mathbf{b}^i$ as

$$\mathbf{b}^i = \begin{bmatrix}\mathbf{w}_1^i \\ \mathbf{w}_2^i\end{bmatrix} + \frac{1}{\alpha}\begin{bmatrix}\mathbf{S}_1^t \\ \mathbf{S}_2^t\end{bmatrix}\left(\mathbf{y} - \begin{bmatrix}\mathbf{S}_1 & \mathbf{S}_2\end{bmatrix}\begin{bmatrix}\mathbf{w}_1^i \\ \mathbf{w}_2^i\end{bmatrix}\right) =: \begin{bmatrix}\mathbf{b}_1^i \\ \mathbf{b}_2^i\end{bmatrix}.$$

Therefore, using the IST algorithm (5) for each of $\mathbf{w}_1$ and $\mathbf{w}_2$, an iterative algorithm to minimize $J(\mathbf{w}_1, \mathbf{w}_2)$ is:

$$\mathbf{r}^i = \mathbf{y} - \mathbf{S}_1\mathbf{w}_1^i - \mathbf{S}_2\mathbf{w}_2^i \tag{12}$$

$$\mathbf{b}_1^i = \mathbf{w}_1^i + \frac{1}{\alpha}\mathbf{S}_1^t\,\mathbf{r}^i \tag{13}$$

$$\mathbf{b}_2^i = \mathbf{w}_2^i + \frac{1}{\alpha}\mathbf{S}_2^t\,\mathbf{r}^i \tag{14}$$

$$\mathbf{w}_1^{i+1} = \text{soft}(\mathbf{b}_1^i,\, \lambda_1/(2\alpha)) \tag{15}$$

$$\mathbf{w}_2^{i+1} = \text{soft}(\mathbf{b}_2^i,\, \lambda_2/(2\alpha)). \tag{16}$$

B.2 MCA using Analysis Prior

In this section, we describe the derivation of a simple algorithm for the minimization of (11). The derivation uses the iteration in Appendix A.2. Note that (11) can be written as

$$J(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \lambda_1 \|\mathbf{A}_1\mathbf{x}_1\|_1 + \lambda_2 \|\mathbf{A}_2\mathbf{x}_2\|_1$$

where

$$\mathbf{H} = \begin{bmatrix}\mathbf{I} & \mathbf{I}\end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{bmatrix}.$$

Using the majorization-minimization approach,[26] an iterative algorithm for finding the minimizer is:

$$\mathbf{b}^i = \mathbf{x}^i + \frac{1}{\alpha}\mathbf{H}^t(\mathbf{y} - \mathbf{H}\mathbf{x}^i)$$
$$\mathbf{x}^{i+1} = \underset{\mathbf{x}}{\operatorname{argmin}} \left[\alpha\|\mathbf{b}^i - \mathbf{x}\|_2^2 + \lambda_1\|\mathbf{A}_1\mathbf{x}_1\|_1 + \lambda_2\|\mathbf{A}_2\mathbf{x}_2\|_1\right]$$

where $\alpha > \text{maxeig}(\mathbf{H}^t\mathbf{H}) = 2$. Equivalently, we can write

$$\begin{bmatrix}\mathbf{b}_1^i \\ \mathbf{b}_2^i\end{bmatrix} = \begin{bmatrix}\mathbf{x}_1^i \\ \mathbf{x}_2^i\end{bmatrix} + \frac{1}{\alpha}\begin{bmatrix}\mathbf{I} \\ \mathbf{I}\end{bmatrix}\left(\mathbf{y} - \begin{bmatrix}\mathbf{I} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{bmatrix}\right)$$
$$\mathbf{x}^{i+1} = \underset{\mathbf{x}}{\operatorname{argmin}} \left[\alpha\|\mathbf{b}_1^i - \mathbf{x}_1\|_2^2 + \alpha\|\mathbf{b}_2^i - \mathbf{x}_2\|_2^2 + \lambda_1\|\mathbf{A}_1\mathbf{x}_1\|_1 + \lambda_2\|\mathbf{A}_2\mathbf{x}_2\|_1\right].$$

Note $\mathbf{x}_1$ and $\mathbf{x}_2$ are uncoupled in the minimization problem, so this can be rewritten as:

$$\mathbf{r}^i = \mathbf{y} - \mathbf{x}_1^i - \mathbf{x}_2^i \tag{17}$$
$$\mathbf{b}_1^i = \mathbf{x}_1^i + \frac{1}{\alpha}\mathbf{r}^i \tag{18}$$
$$\mathbf{b}_2^i = \mathbf{x}_2^i + \frac{1}{\alpha}\mathbf{r}^i \tag{19}$$
$$\mathbf{x}_1^{i+1} = \underset{\mathbf{x}_1}{\operatorname{argmin}} \left[\alpha\|\mathbf{b}_1^i - \mathbf{x}_1\|_2^2 + \lambda_1\|\mathbf{A}_1\mathbf{x}_1\|_1\right] \tag{20}$$
$$\mathbf{x}_2^{i+1} = \underset{\mathbf{x}_2}{\operatorname{argmin}} \left[\alpha\|\mathbf{b}_2^i - \mathbf{x}_2\|_2^2 + \lambda_2\|\mathbf{A}_2\mathbf{x}_2\|_1\right]. \tag{21}$$

Using a single step of algorithm (8)-(9) suggests the iteration:

$$\mathbf{r}^i = \mathbf{y} - \mathbf{x}_1^i - \mathbf{x}_2^i \tag{22}$$
$$\mathbf{b}_1^i = \mathbf{x}_1^i + \frac{1}{\alpha}\mathbf{r}^i \tag{23}$$
$$\mathbf{b}_2^i = \mathbf{x}_2^i + \frac{1}{\alpha}\mathbf{r}^i \tag{24}$$
$$\mathbf{z}_1^{i+1} = \left(c_1\,\mathbf{z}_1^i + \mathbf{A}_1\mathbf{b}_1^i - \mathbf{A}_1\mathbf{A}_1^t\,\mathbf{z}_1^i\right)./\left(\frac{2\alpha}{\lambda_1}\,|\mathbf{A}_1\mathbf{x}_1^i| + c_1\right) \tag{25}$$
$$\mathbf{x}_1^{i+1} = \mathbf{b}_1^i - \mathbf{A}_1^t\,\mathbf{z}_1^{i+1} \tag{26}$$
$$\mathbf{z}_2^{i+1} = \left(c_2\,\mathbf{z}_2^i + \mathbf{A}_2\mathbf{b}_2^i - \mathbf{A}_2\mathbf{A}_2^t\,\mathbf{z}_2^i\right)./\left(\frac{2\alpha}{\lambda_2}\,|\mathbf{A}_2\mathbf{x}_2^i| + c_2\right) \tag{27}$$
$$\mathbf{x}_2^{i+1} = \mathbf{b}_2^i - \mathbf{A}_2^t\,\mathbf{z}_2^{i+1} \tag{28}$$

where $c_1 > \text{maxeig}(\mathbf{A}_1\mathbf{A}_1^t)$ and $c_2 > \text{maxeig}(\mathbf{A}_2\mathbf{A}_2^t)$.

An example of MCA is illustrated in Fig. 6. The test signal is the sum of the 'blocks' signal,[20] a sparse signal, and independent white Gaussian noise. Because $\mathbf{x}_1$ is piecewise constant, we choose $\mathbf{A}_1$ to be the first difference operator (so that $\|\mathbf{A}_1\mathbf{x}\|_1$ is the total variation[41] of $\mathbf{x}$). Because $\mathbf{x}_2$ is itself sparse, we choose $\mathbf{A}_2$ to be the identity matrix. The algorithm (22)-(28) was used to perform simultaneous denoising and decomposition. The components are illustrated in the figure.

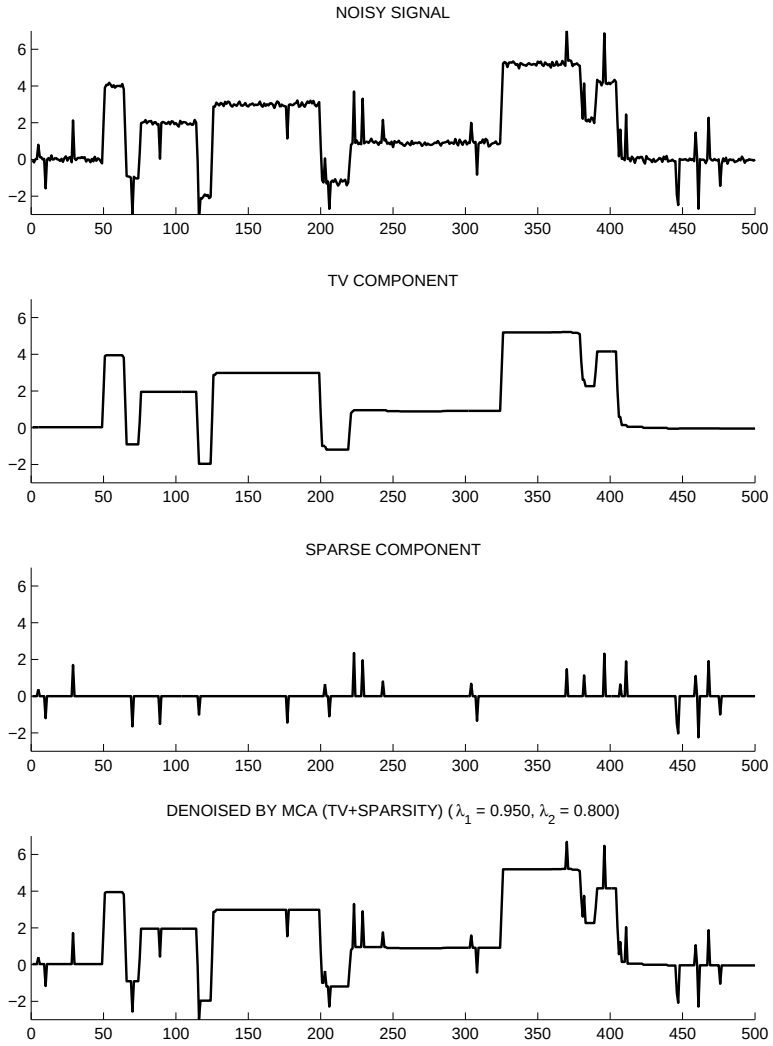

Figure 6. MCA Example: the decomposition of a noisy signal (piecewise constant plus sparse) using the iteration (22)-(28).

REFERENCES

[1] İ. Bayram and I. W. Selesnick. Frequency-domain design of overcomplete rational-dilation wavelet transforms. *IEEE Trans. on Signal Processing*, 57(8):2957–2972, August 2009.

[2] A. Beck and M. Teboulle. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. Preprint, retrieved from http://iew3.technion.ac.il/Home/Users/becka.html, 2009.

[3] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imag. Sci.*, 2(1):183–202, 2009.

[4] J. Bect, L. Blanc-Féaud, G. Aubert, and A. Chambolle. A l^1-unified variational framework for image restoration. In T. Pajdla and J. Matas, editors, *European Conference on Computer Vision, Lecture Notes in Computer Sciences*, volume 3024, pages 1–13, 2004.

[5] J. M. Bioucas-Dias and M. A. T. Figueiredo. A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Trans. on Image Processing*, 16(12):2992–3004, December 2007.

[6] V. Bruni and D. Vitulano. Transients detection in the time-scale domain. In *Proc. 3rd International Conference on Image and Signal Processing (ICISP)*, pages 254–262, Berlin, Heidelberg, 2008. Springer-Verlag. LNCS 5099.

[7] J.-F. Cai, S. Osher, and Z. Shen. Linearized Bregman iterations for frame-based image deblurring. *SIAM J. Imag. Sci.*, 2(1):226–252, 2009.

[8] J.-F. Cai, S. Osher, and Z. Shen. Split Bregman methods and frame based image restoration. Preprint, retrieved from http://www.math.nus.edu.sg/ matzuows/publist.html, 2009.

[9] E. Candes and J. Romberg. L1-magic: Recovery of sparse signals. *http://www.acm.caltech.edu/l1magic/*, 2005.

[10] A. Chambolle. An algorithm for total variation minimization and applications. *J. of Math. Imaging and Vision*, 20:89–97, 2004.

[11] S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.*, 20(1):33–61, 1998.

[12] I. Clark, R. Biscay, M. Echeverría, and T. Virués. Multiresolution decomposition of non-stationary EEG signals: A preliminary study. *Comput. Biol. Med.*, 25(4):373–382, 1995.

[13] C. E. D'Attellis, S. I. Isaacson, and R. O. Sirne. Detection of epileptic events in electroencephalograms using wavelet analysis. *Annals of Biomedical Engineering*, 25:286–293, 1997.

[14] I. Daubechies. *Ten Lectures On Wavelets*. SIAM, 1992.

[15] I. Daubechies, M. Defriese, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math*, LVII:1413–1457, 2004.

[16] I. Daubechies, G. Teschke, and L. Vese. On some iterative concepts for image restoration. In P. W. Hawkes, editor, *Advances in Imaging and Electron Physics*, volume 150, pages 2–51. Elsevier, 2008.

[17] L. Daudet. Sparse and structured decompositions of signals with the molecular matching pursuit. *IEEE Trans. on Audio, Speech, and Lang. Proc.*, 14(5):1808–1816, September 2006.

[18] L. Daudet and B. Torrésani. Hybrid representations for audiophonic signal encoding. *Signal Processing*, 82(11):1595–1617, November 2002.

[19] M. E. Davies and L. Daudet. Sparse audio representations using the MCLT. *Signal Processing*, 86(3):457–470, March 2006.

[20] D. Donoho, A. Maleki, and M. Shahram. Wavelab 850. `http://www-stat.stanford.edu/~wavelab/`.

[21] M. Elad, B. Matalon, and M. Zibulevsky. Coordinate and subspace optimization methods for linear least squares with non-quadratic regularization. *J. of Appl. and Comp. Harm. Analysis*, 23:346–367, 2007.

[22] M. Elad, P. Milanfar, and R. Rubinstein. Analysis versus synthesis in signal priors. *Inverse Problems*, 23(3):947–968, 2007.

[23] M. Elad, J. Starck, P. Querre, and D. Donoho. Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). *J. of Appl. and Comp. Harm. Analysis*, 19(3):340–358, November 2005.

[24] M. J. Fadili, J.-L. Starck, and F. Murtagh. Inpainting and zooming using sparse representations. *The Computer Journal*, 52(1):64–79, 2009. Published online on July 29, 2007.

[25] C. Févotte, L. Daudet, S. J. Godsill, and B. Torrésani. Sparse regression with structured priors: Application to audio denoising. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, volume 3, 2006.

[26] M. Figueiredo, J. Bioucas-Dias, and R. Nowak. Majorization-minimization algorithms for wavelet-based image restoration. *IEEE Trans. on Image Processing*, 16(12):2980–2991, December 2007.

[27] M. Figueiredo and R. Nowak. An EM algorithm for wavelet-based image restoration. *IEEE Trans. on Image Processing*, 12(8):906–916, August 2003.

[28] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE J. Sel. Top. Signal Process.*, 1(4):586–598, December 2007.

[29] P. Flandrin and P. Gonçalvès. Empirical mode decompositions as data-driven wavelet-like expansions. *Int. J. Wavelets Multires. Inf. Process.*, 2(4):477496, 2004.

[30] T. Goldstein and S. Osher. The split Bregman method for L1-regularized problems. *SIAM J. Imag. Sci.*, 2(2):323–343, 2009.

[31] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. Roy. Soc. Lon. A*, 454(1971):903–995, March 8, 1998.

[32] F. Jaillet and B. Torrésani. Time-frequency jigsaw puzzle: adaptive multiwindow and multilayered Gabor representations. *Int. J. Wavelets Multires. Inf. Process.*, 5(2):293–316, March 2007.

[33] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky. An interior-point method for large-scale l1-regularized least squares. *IEEE. J. Sel. Top. Signal Processing*, 1(4):606–617, December 2007.

[34] S. N. Levine and J. O. Smith III. A sines+transients+noise audio representation for data compression and time/pitch scale modications. In *Proceedings of the 105th Audio Engineering Society Convention*, 1998.

[35] S. Molla and B. Torrésani. An hybrid audio coding scheme using hidden Markov models of waveforms. *J. of Appl. and Comp. Harm. Analysis*, 18(2):137–166, 2005.

[36] S. Nishida, M. Nakamura, A. Ikeda, and H. Shibasaki. Signal separation of background EEG and spike using morphological filter. *Medical Engineering and Physics*, 21:601–608, 1999.

[37] J. Oliveira, J. Bioucas-Dias, and M. A. T. Figueiredo. Adaptive total variation image deblurring: A majorization-minimization approach. *Signal Processing*, 89(9):1683–1693, September 2009.

[38] L.-S. Pon, M. Sun, and R. J. Sclabassi. Adaptive separation of background activity and transient phenomenon in epileptic EEG using mathematical morphology and wavelet transforms. In *Proc. IEEE Annual EMBS Conf.*, 2000.

[39] J. Portilla and L. Mancera. L0-based sparse approximation: two alternative methods and some applications. In *Proceedings of SPIE*, volume 6701 (Wavelets XII), 2007.

[40] P. Rodríguez and B. Wohlberg. Efficient minimization method for a generalized total variation functional. *IEEE Trans. on Image Processing*, 18(2):322–332, February 2009.

[41] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[42] J.-L. Starck, M. Elad, and D. Donoho. Redundant multiscale transforms and their application for morphological component analysis. *Advances in Imaging and Electron Physics*, 132:287–348, 2004.

[43] J.-L. Starck, M. Elad, and D. Donoho. Image decomposition via the combination of sparse representation and a variational approach. *IEEE Trans. on Image Processing*, 14(10), 2005.

[44] J.-L. Starck, Y. Moudden, J. Bobina, M. Elad, and D.L. Donoho. Morphological component analysis. In *Proceedings of SPIE*, volume 5914 (Wavelets XI), 2005.

[45] Y. Wang, J. Yang, W. Yin, and Y. Zhang. A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. on Imaging Sciences*, 1(3):248–272, 2008.

[46] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo. Sparse reconstruction by separable approximation. *IEEE Trans. on Signal Processing*, 57(7):2479–2493, July 2009.

[47] M. Zhu, S. J. Wright, and T. F. Chan. Duality-based algorithms for total-variation-regularized image restoration. *J. Comp. Optimization and Applications*, 2008.