Sparse signal representation

and

the tunable Q-factor wavelet transform

Ivan Selesnick

Polytechnic Institute of New York University

Brooklyn, New York

Introduction

Problem: Decomposition of a signal into the sum of two components:

1. Oscillatory (rhythmic, tonal) component
2. Transient (non-oscillatory) component

Outline

1. Signal resonance and Q-factors
2. Morphological component analysis (MCA)
3. Tunable Q-factor wavelet transform (TQWT)
4. Split augmented Lagrangian shrinkage algorithm (SALSA)
5. Examples

References (MDCT, etc)

S. N. Levine and J. O. Smith III. *A sines+transients+noise audio representation for data compression and time/pitch scale modications*. (1998)

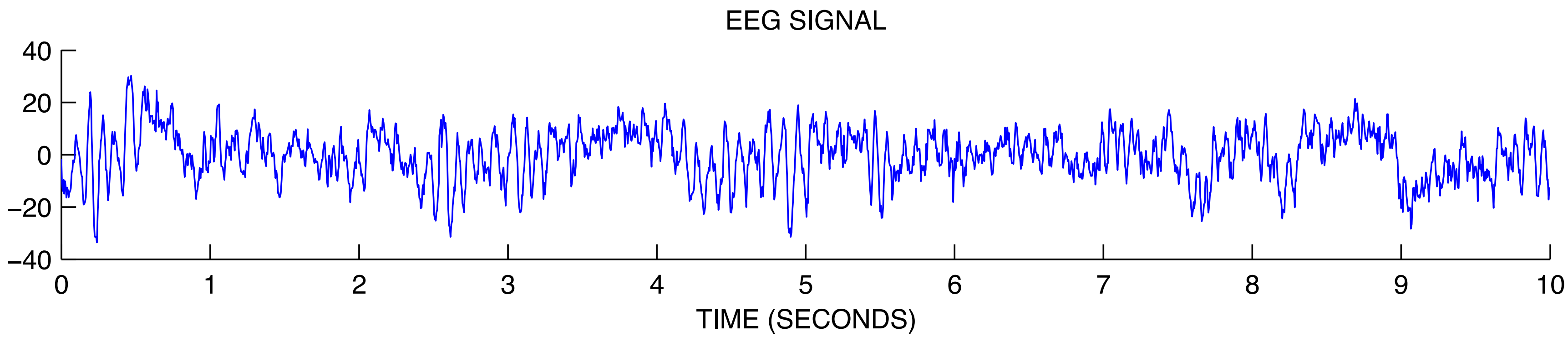
L. Daudet and B. Torrésani. *Hybrid representations for audiophonic signal encoding*. (2002)

S. Molla and B. Torrésani. *An hybrid audio coding scheme using hidden Markov models of waveforms*. (2005)

M. E. Davies and L. Daudet. *Sparse audio representations using the MCLT*. (2006)

Oscillatory (rhythmic) and Transient Components in EEG

Many measured signals have both an oscillatory and a non-oscillatory component.



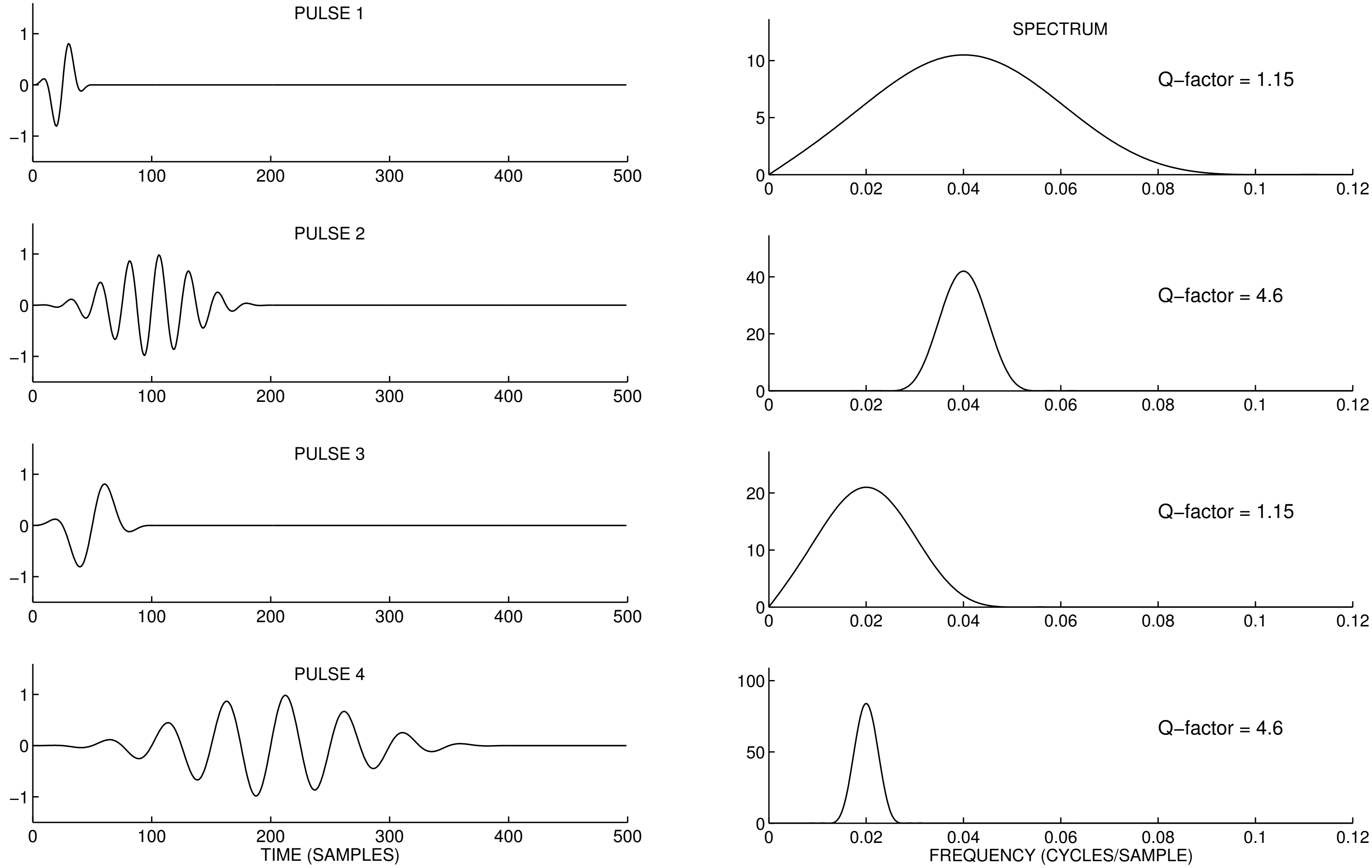
Rhythms of the EEG:

Delta	0 - 3 Hz
Theta	4 - 7 Hz
Alpha	8 - 12 Hz
Beta	12 - 30 Hz
Gamma	26 - 100 Hz

Transients in EEG due to:

1) unwanted measurement artifacts

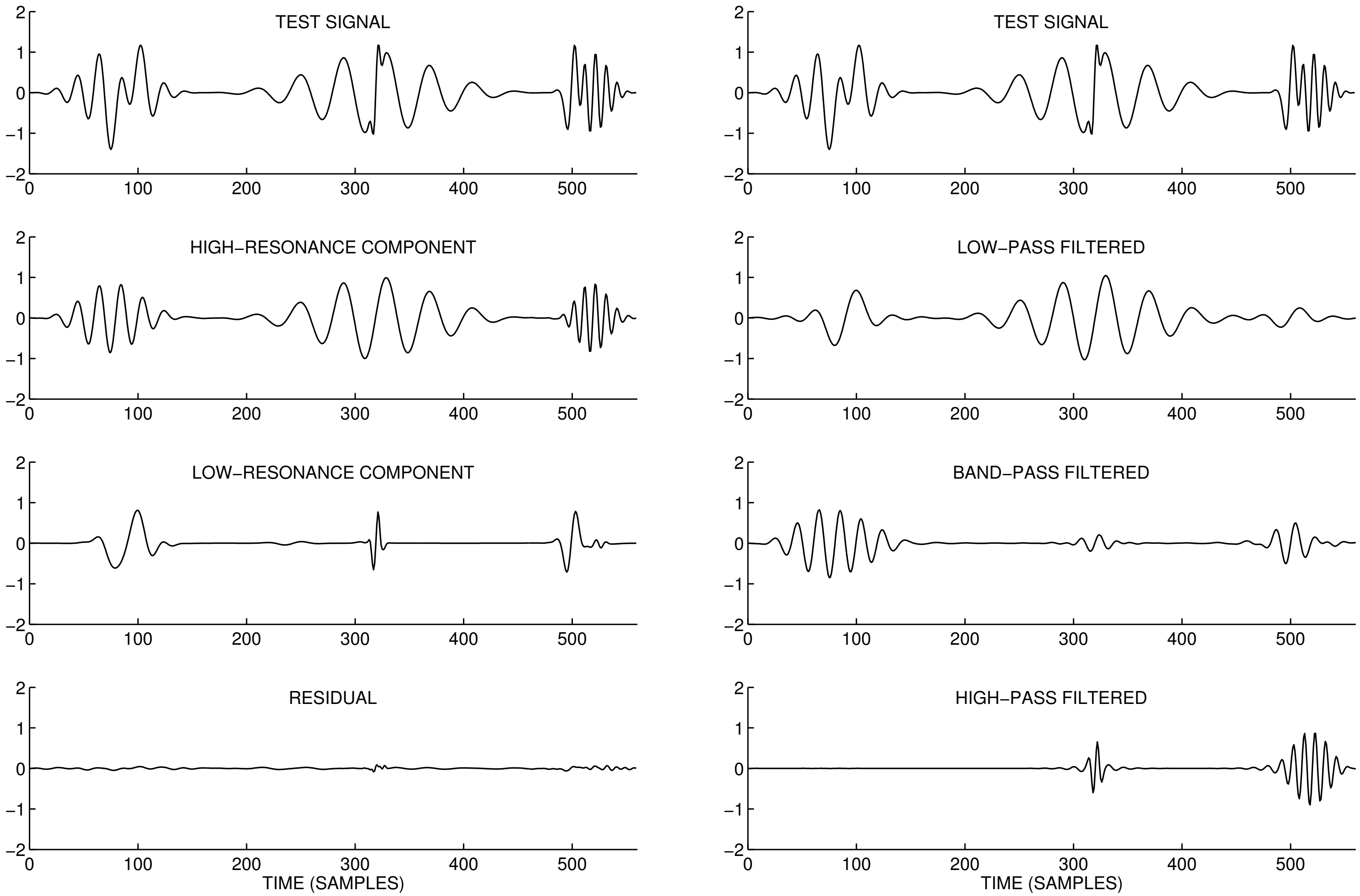
2) non-rhythmic brain activity (spikes, spindles, and vertex waves)



(a) Signals.

(b) Spectra.

Figure 1: The resonance of an isolated pulse can be quantified by its Q-factor, defined as the ratio of its center frequency to its bandwidth. Pulses 1 and 3, essentially a single cycle in duration, are low-resonance pulses. Pulses 2 and 4, whose oscillations are more sustained, are high-resonance pulses.



(a) Resonance-based decomposition.

(b) Frequency-based filtering.

Figure 2: Resonance- and frequency-based filtering. (a) Decomposition of a test signal into high- and low-resonance components. The high-resonance signal component is sparsely represented using a high Q-factor WT. Similarly, the low-resonance signal component is sparsely represented using a low Q-factor WT. (b) Decomposition of a test signal into low, mid, and high frequency components using LTI discrete-time filters.

Resonance-based signal decomposition must be nonlinear

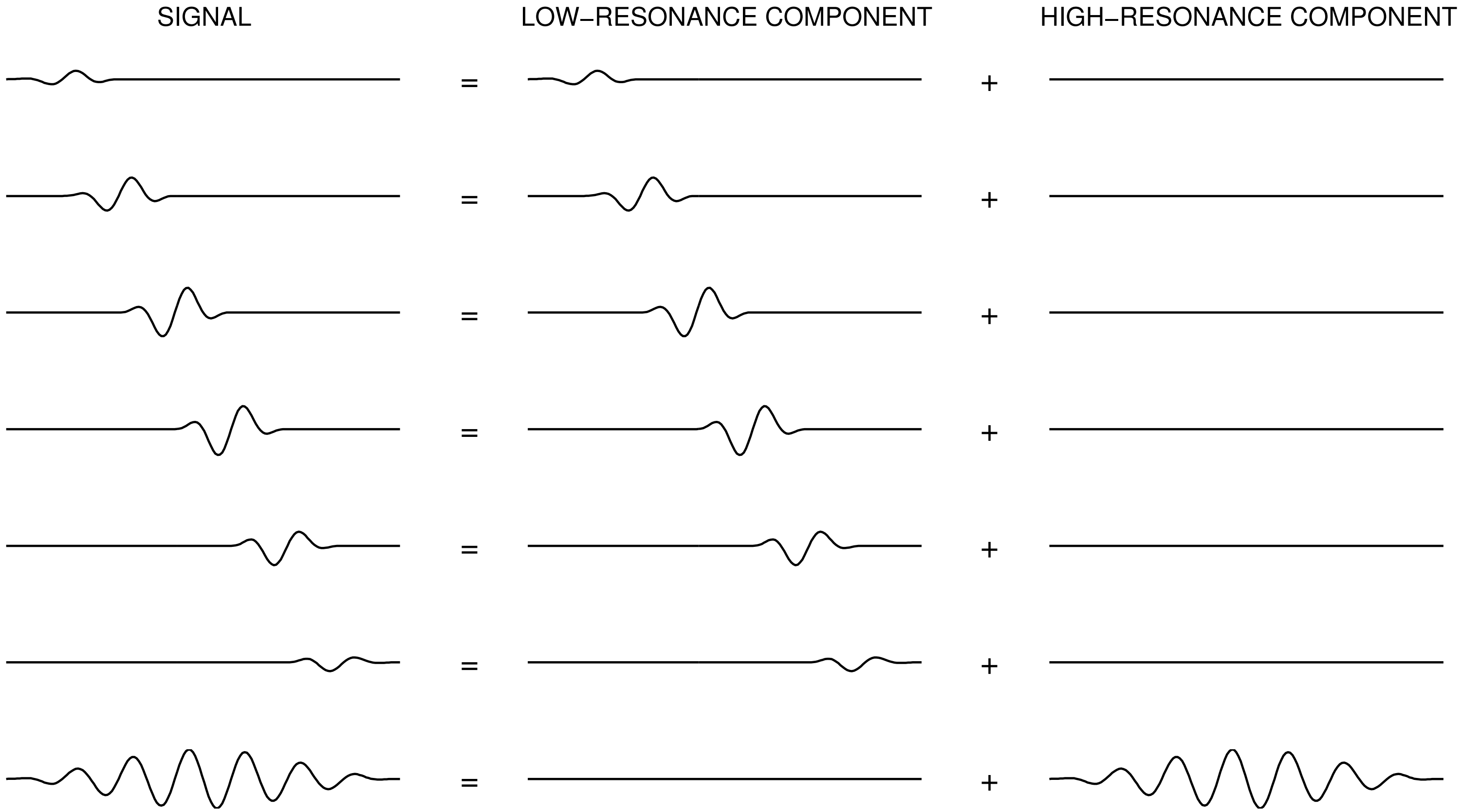


Figure 3: Resonance-based signal decomposition must be nonlinear: The signal in the bottom left panel is the sum of the signals above it; however, the low-resonance component of a sum is not the sum of the low-resonance components. The same is true for the high-resonance component. Neither the low- nor high-resonance components satisfy the superposition property.

$$F(\mathbf{s}_1 + \cdots + \mathbf{s}_6) \neq F(\mathbf{s}_1) + \cdots + F(\mathbf{s}_6)$$

Rational-dilation wavelet transform (RADWT)

Prior work on rational-dilation wavelet transforms addresses the critically-sampled case.

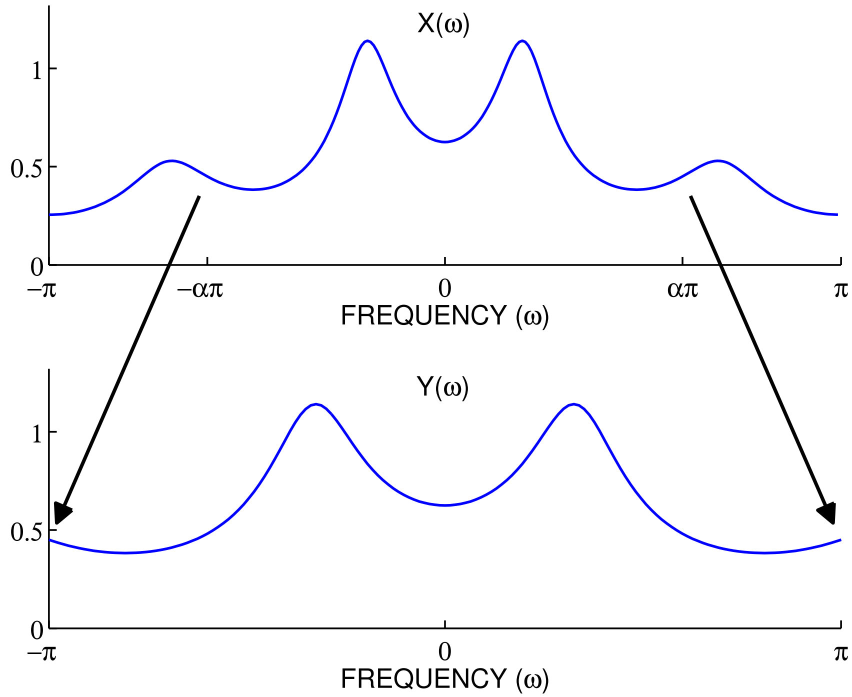
1. K. Nayebi, T. P. Barnwell III and M. J. T. Smith (1991)
2. P. Auscher (1992)
3. J. Kovacevic and M. Vetterli (1993)
4. T. Blu (1993, 1996, 1998)
5. A. Baussard, F. Nicolier and F. Truchetet (2004)
6. G. F. Choueiter and J. R. Glass (2007)

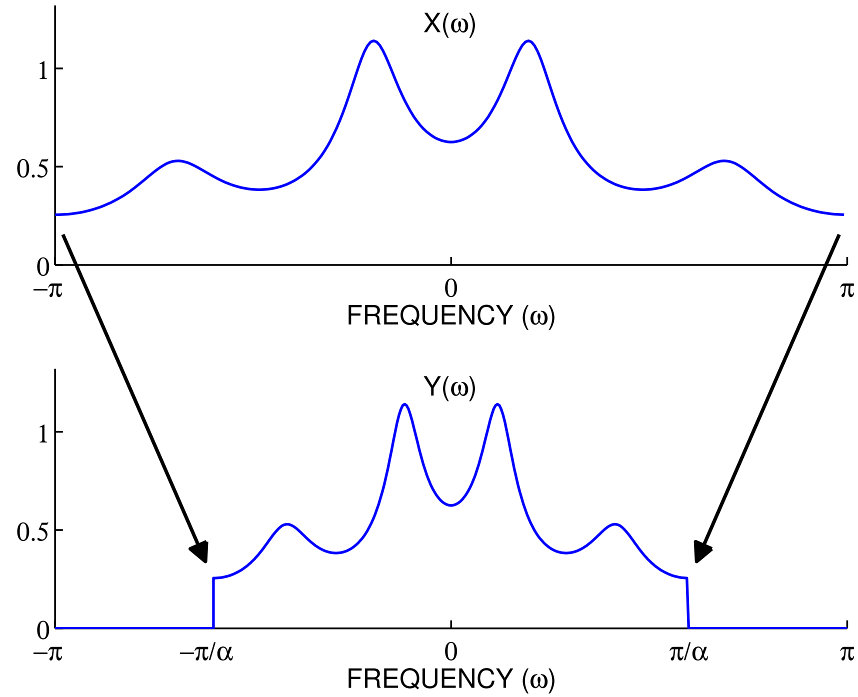
RADWT (2009) gives a solution for the overcomplete case.

Reference:

Bayram, Selesnick. *Frequency-domain design of overcomplete rational-dilation wavelet transforms.* IEEE Trans. on Signal Processing, 57, August 2009.

New: Tunable Q-factor wavelet transform — dilation need not be rational.



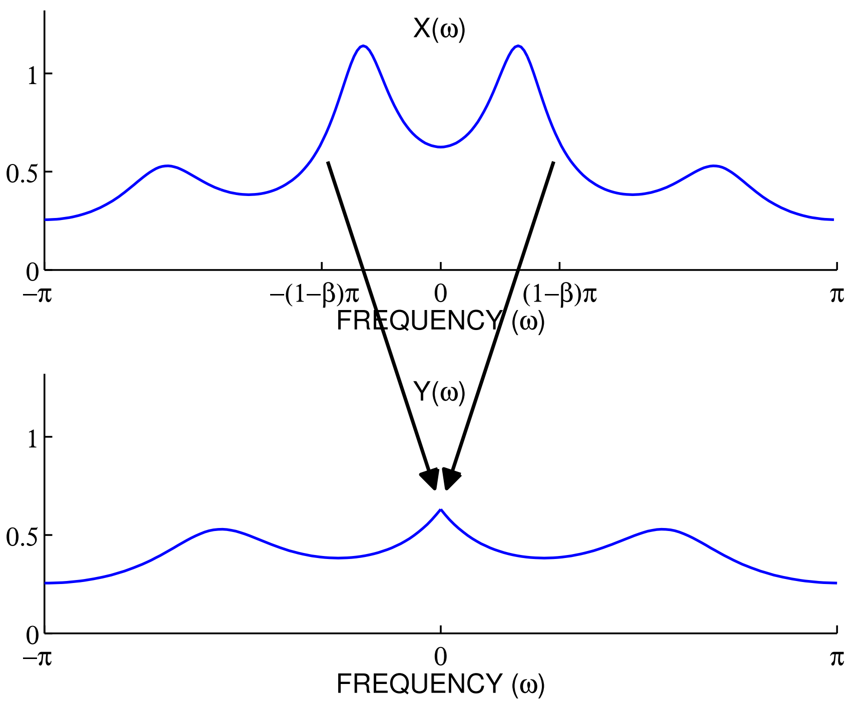


Low-pass scaling with $\alpha < 1$

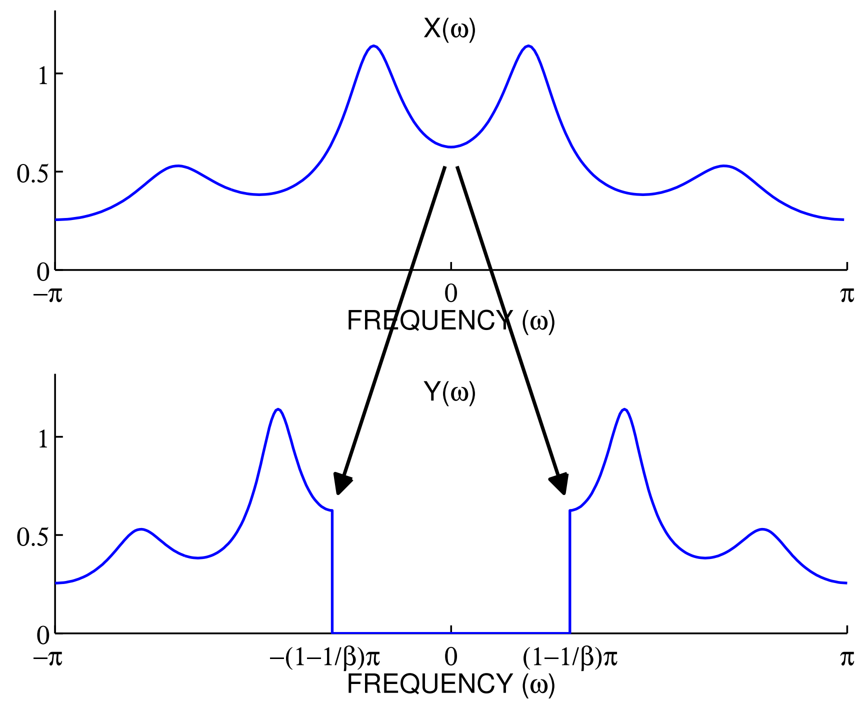
Low-pass scaling with $\alpha > 1$

High-pass Scaling

$$x(n) \longrightarrow \boxed{\text{HPS } \beta} \longrightarrow y(n)$$

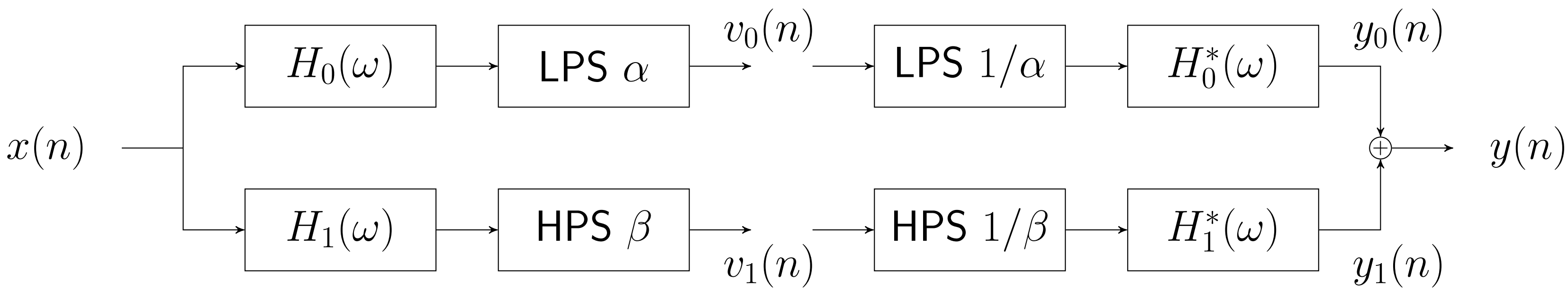


High-pass scaling with $\beta < 1$



High-pass scaling with $\beta > 1$

Tunable Q-factor wavelet transform (TQWT)



$$0 < \beta \le 1, \qquad 0 < \alpha < 1, \qquad \alpha + \beta > 1$$

$$Y_0(\omega) = \begin{cases} |H_0(\omega)|^2 \, X(\omega) & |\omega| \le \alpha\,\pi \\ 0 & \alpha\,\pi < |\omega| \le \pi \end{cases}$$

$$Y_1(\omega) = \begin{cases} 0 & |\omega| < (1-\beta)\,\pi \\ |H_1(\omega)|^2 \, X(\omega) & (1-\beta)\,\pi \le |\omega| \le \pi \end{cases}$$

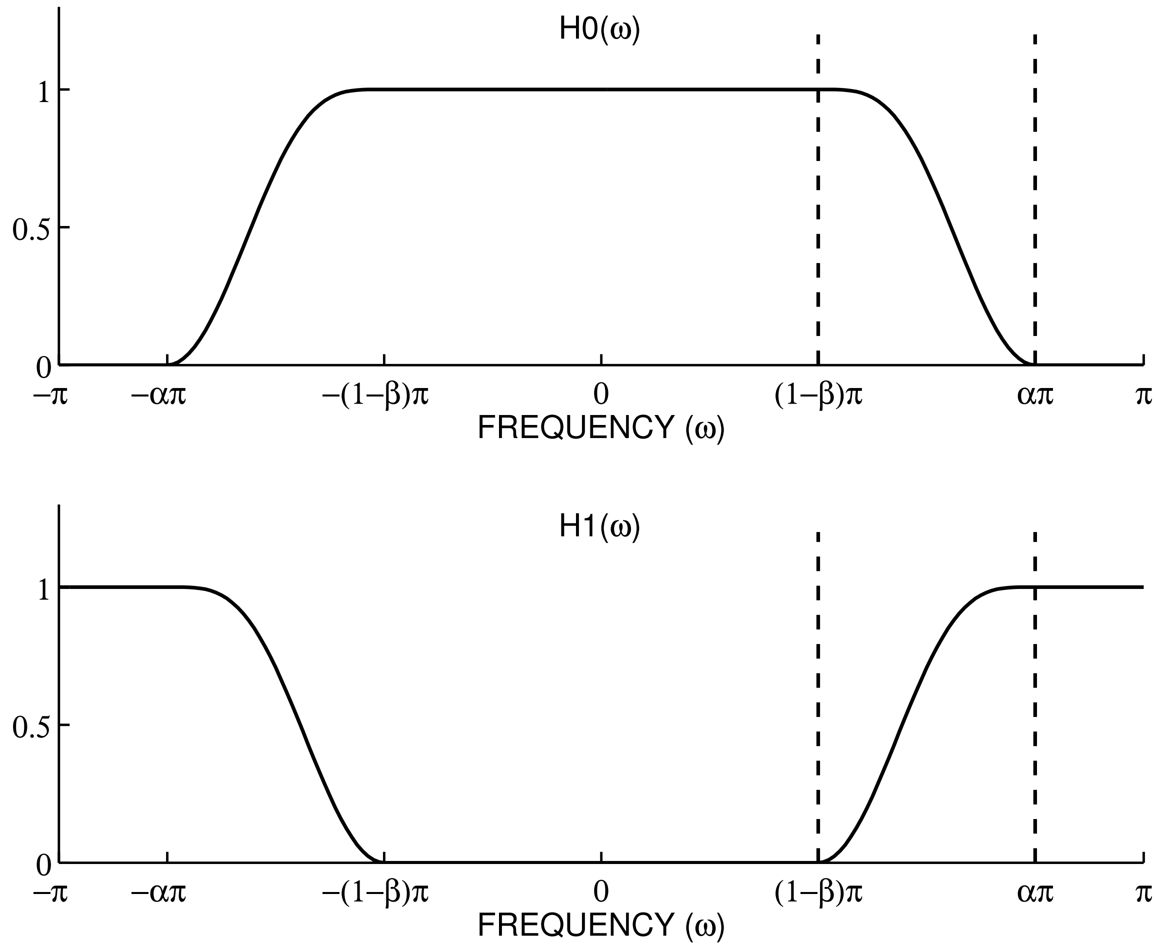
$$Y(\omega) = \begin{cases} |H_0(\omega)|^2 \, X(\omega) & |\omega| < (1-\beta)\,\pi \\ (|H_0(\omega)|^2 + |H_1(\omega)|^2)\, X(\omega) & (1-\beta)\,\pi \le |\omega| < \alpha\,\pi \\ |H_1(\omega)|^2 \, X(\omega) & \alpha\,\pi \le |\omega| \le \pi \end{cases}$$

For perfect reconstruction, the filters should satisfy

$$|H_0(\omega)| = 1, \qquad H_1(\omega) = 0, \qquad |\omega| \leq (1-\beta)\,\pi$$
$$H_0(\omega) = 0, \qquad |H_1(\omega)| = 1, \qquad \alpha\,\pi \leq |\omega| \leq \pi$$

The transition bands of $H_0(\omega)$ and of $H_1(\omega)$ must be chosen so that

$$|H_0(\omega)|^2 + |H_1(\omega)|^2 = 1 \qquad (1-\beta)\,\pi < |\omega| < \alpha\,\pi. \tag{1}$$



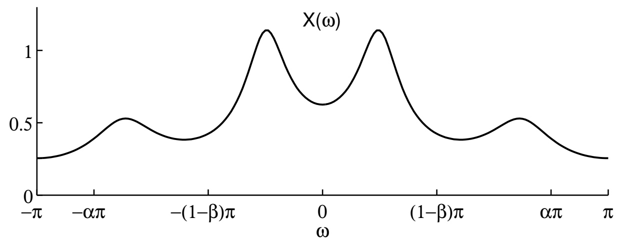
The transition band of $H_0(\omega)$ and $H_1(\omega)$ can be constructed using any power-complementary function, $\theta(\omega)$,

$$\theta^2(\omega) + \theta^2(\pi - \omega) = 1, \tag{2}$$

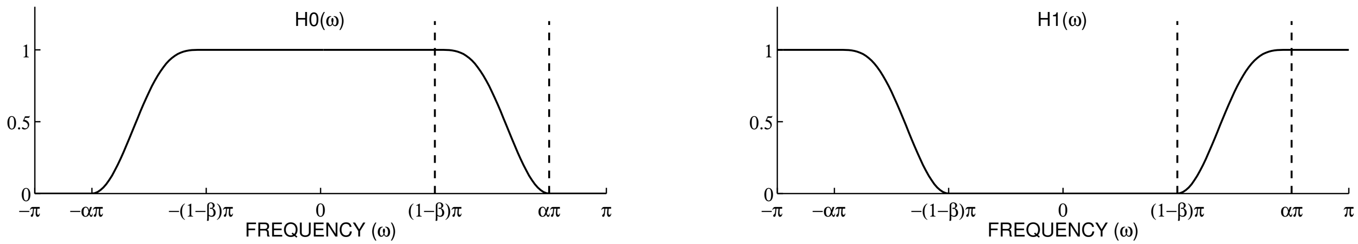
We use the Daubechies filter frequency response with two vanishing moments,

$$\theta(\omega) = \frac{1}{2}\left(1 + \cos(\omega)\right)\sqrt{2 - \cos(\omega)}, \qquad |\omega| \leq \pi. \tag{3}$$

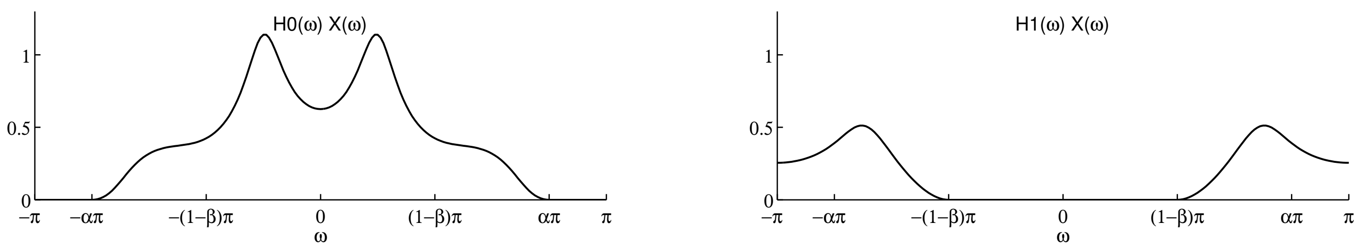
Scale and dilate $\theta(\omega)$ to obtain transition bands for $H_0(\omega)$ and $H_1(\omega)$.



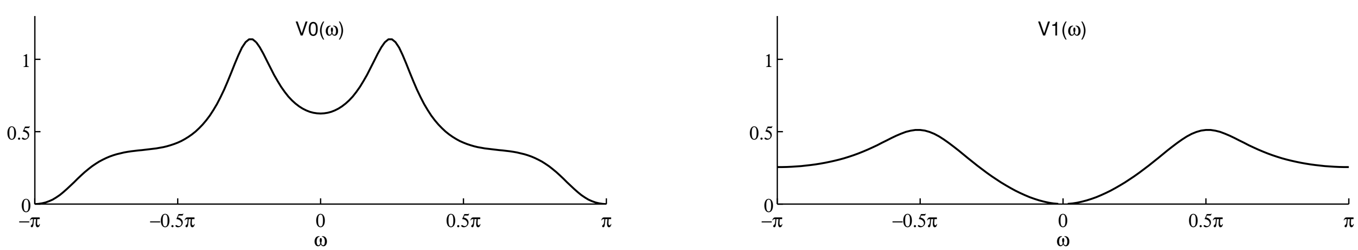
(a) Fourier transform of input signal, $X(\omega)$.



(b) Frequency responses $H_0(\omega)$ and $H_1(\omega)$.

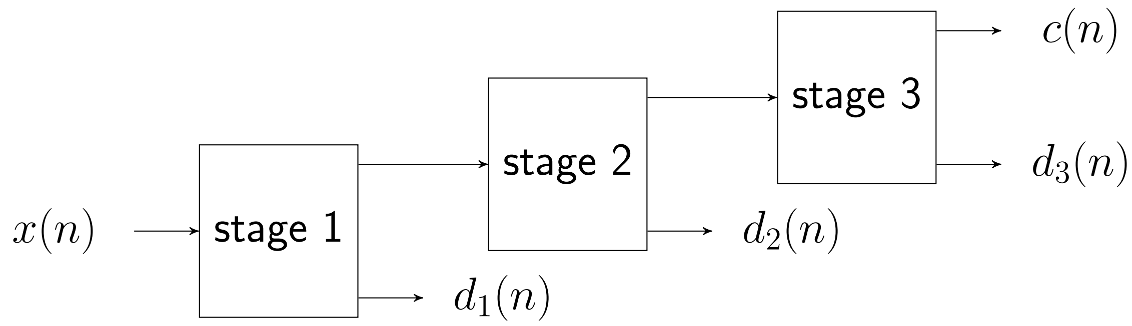


(c) Fourier transforms of input signal after filtering.



(d) Fourier transforms after scaling.

Iterated Filters



Redundancy (total oversampling rate) for many stages:

$$r = \frac{\beta}{1-\alpha}$$

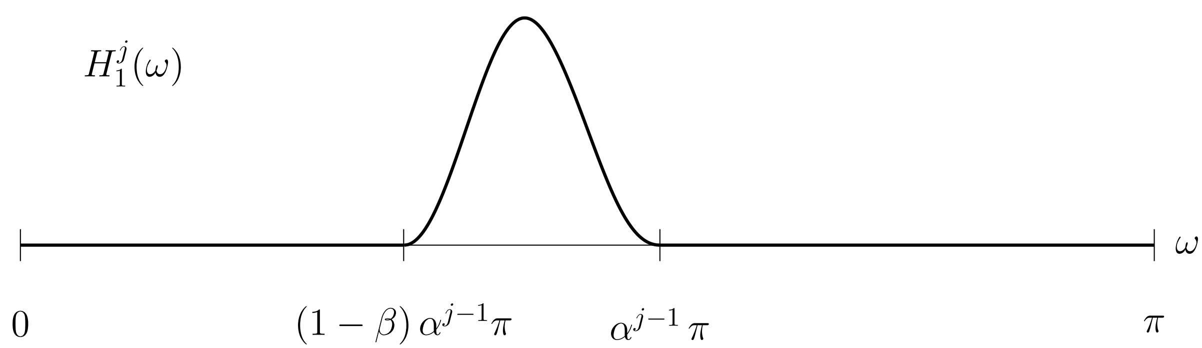
When $\alpha \le 1$, $\beta \le 1$, we have the system equivalence:

$$x(n) \longrightarrow \underbrace{\boxed{H_0(\omega)} \longrightarrow \boxed{\text{LPS } \alpha} \longrightarrow \cdots \longrightarrow \boxed{H_0(\omega)} \longrightarrow \boxed{\text{LPS } \alpha}}_{j-1 \text{ stages}} \longrightarrow \boxed{H_1(\omega)} \longrightarrow \boxed{\text{HPS } \beta} \longrightarrow$$

$$\equiv \quad \longrightarrow \boxed{H_1^{(j)}(\omega)} \longrightarrow \boxed{\text{LPS } \alpha^{j-1}} \longrightarrow \boxed{\text{HPS } \beta} \longrightarrow d_j(n)$$

The equivalent filter is

$$H_1^{(j)}(\omega) := \begin{cases} 0 & |\omega| < (1-\beta)\,\alpha^{j-1}\,\pi \\ H_1(\omega/\alpha^{j-1}) \displaystyle\prod_{m=0}^{j-2} H_0(\omega/\alpha^m) & (1-\beta)\,\alpha^{j-1}\,\pi \le |\omega| \le \alpha^{j-1}\,\pi \\ 0 & \alpha^{j-1}\,\pi < |\omega| \le \pi. \end{cases} \tag{4}$$



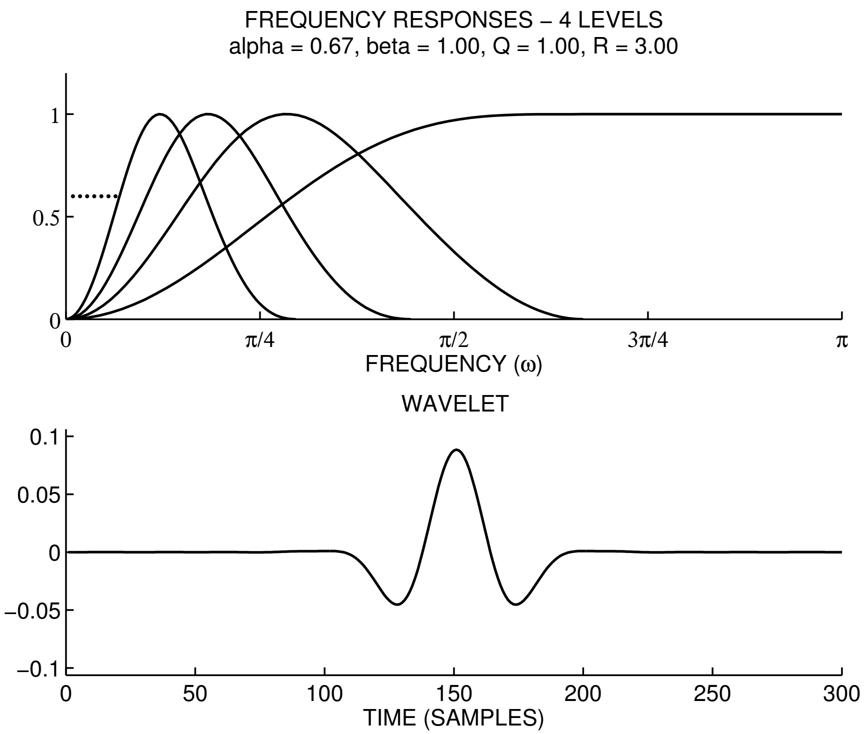
Q-factor:

$$Q = \frac{\omega_c}{\mathsf{BW}} = \frac{2-\beta}{\beta}$$

Scaling factors α, β:

$$\boxed{\beta = \frac{2}{Q+1}, \quad \alpha = 1 - \frac{\beta}{r}}$$

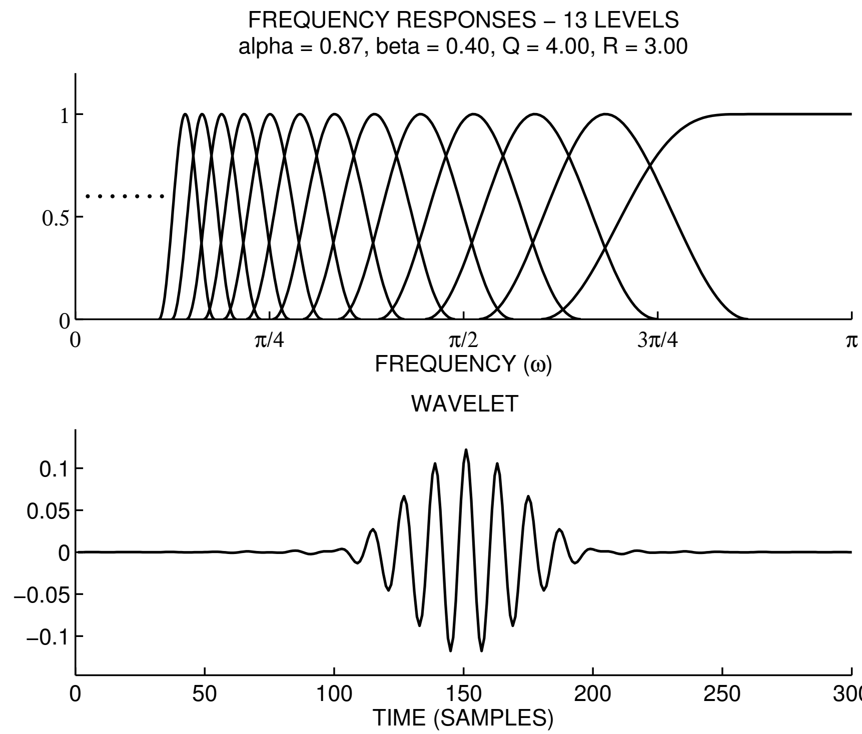
LOW Q-FACTOR WT



Q-factor $= 1$

Redundancy $= 3$

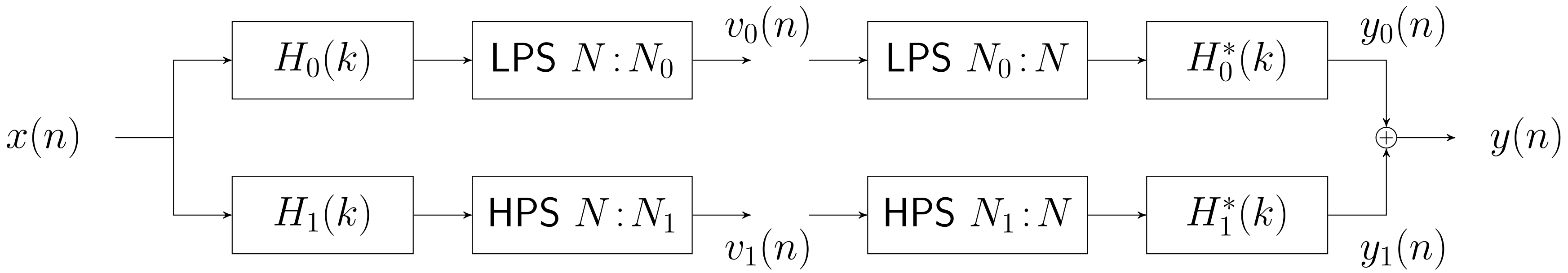
HIGH Q-FACTOR WT



Q-factor $= 4$

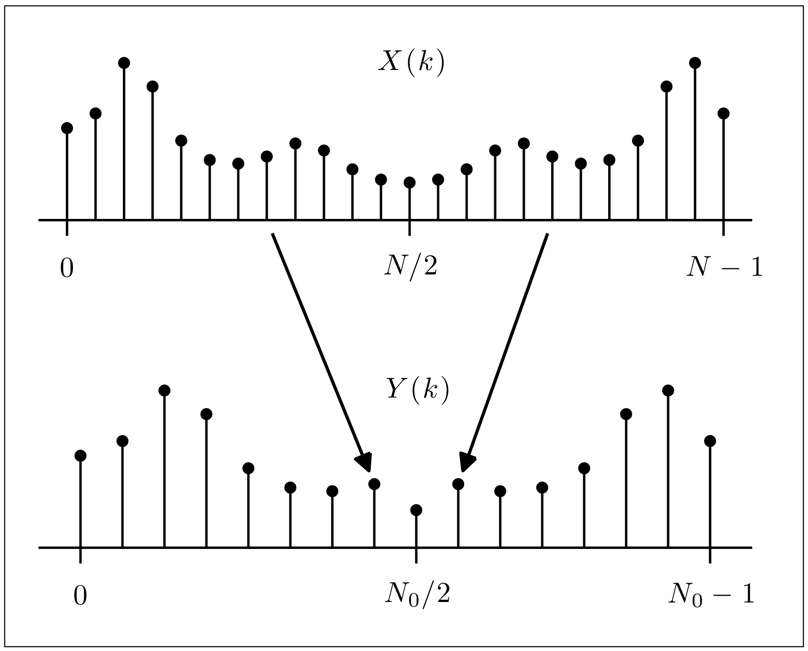
Redundancy $= 3$

The TWQT can be implemented for finite-length signals using the DFT (FFT)....

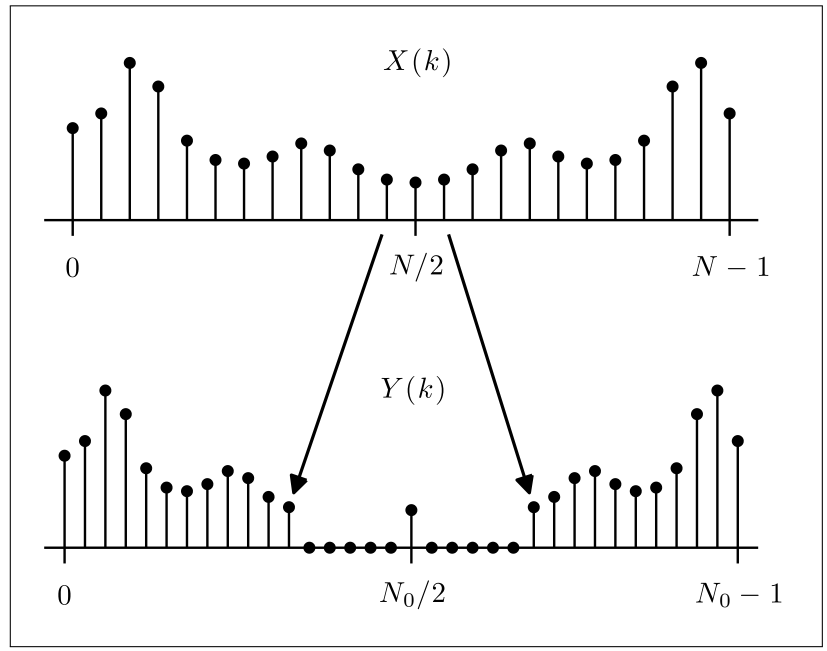


Low-pass scaling $N : N_0$

$$x(n) \longrightarrow \boxed{\text{LPS } N : N_0} \longrightarrow y(n)$$



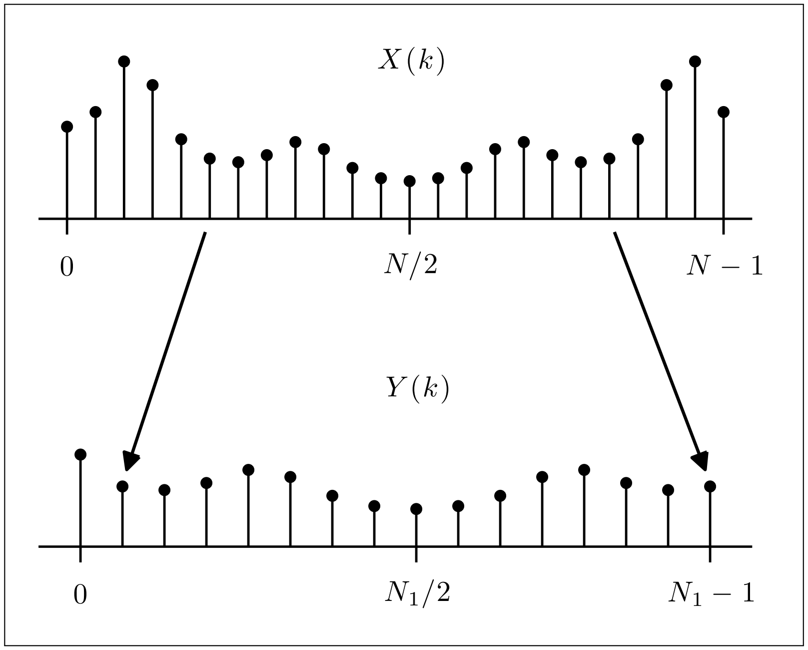
Low-pass scaling with $N_0 < N$

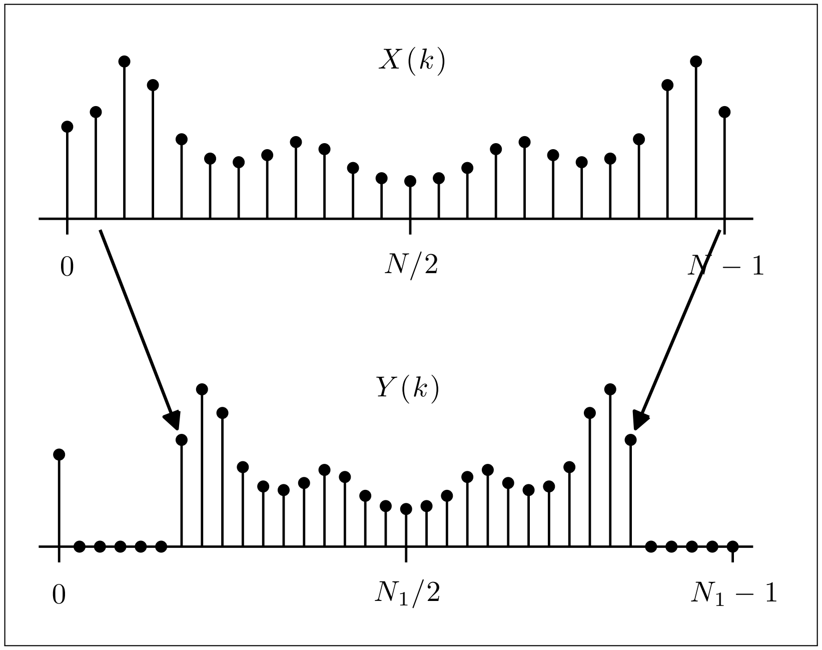


Low-pass scaling with $N_0 > N$.

High-pass scaling $N:N_1$

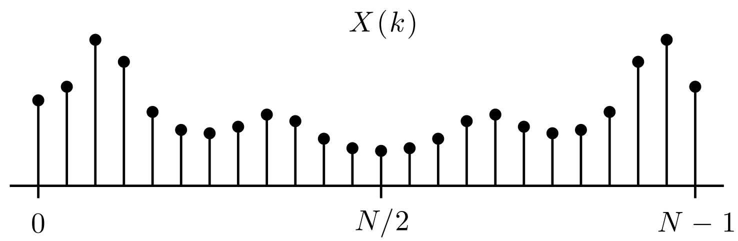
$$x(n) \longrightarrow \boxed{\text{HPS } N:N_1} \longrightarrow y(n)$$



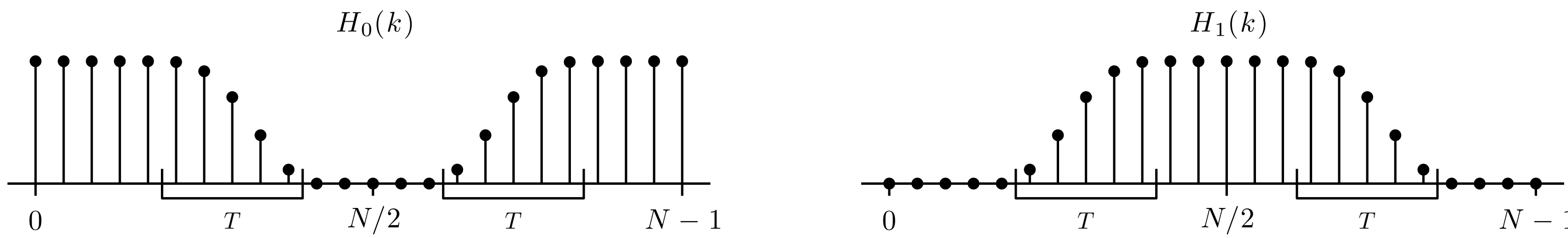


High-pass scaling with $N_1 < N$

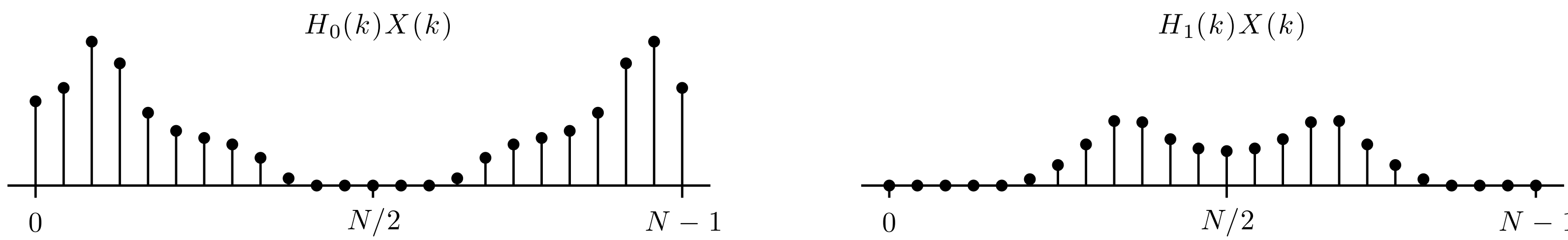
High-pass scaling with $N_1 > N$



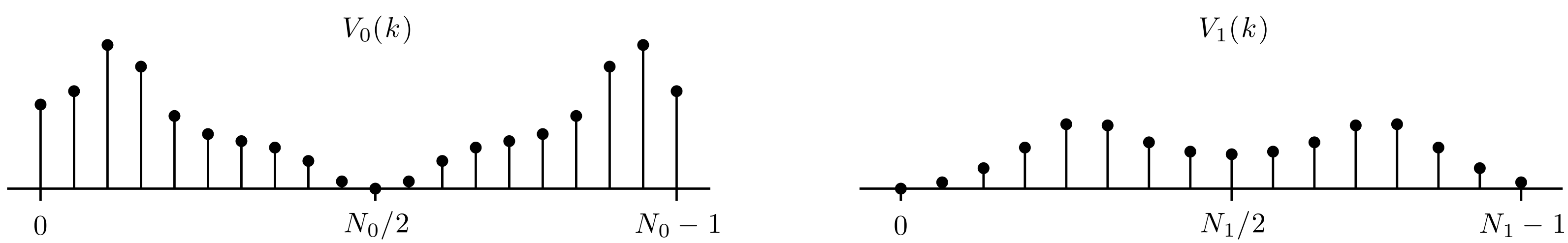
(a) DFT of input signal, $X(k)$.



(b) Filters $H_0(k)$ and $H_1(k)$. The transitions-bands are indicated by 'T'.

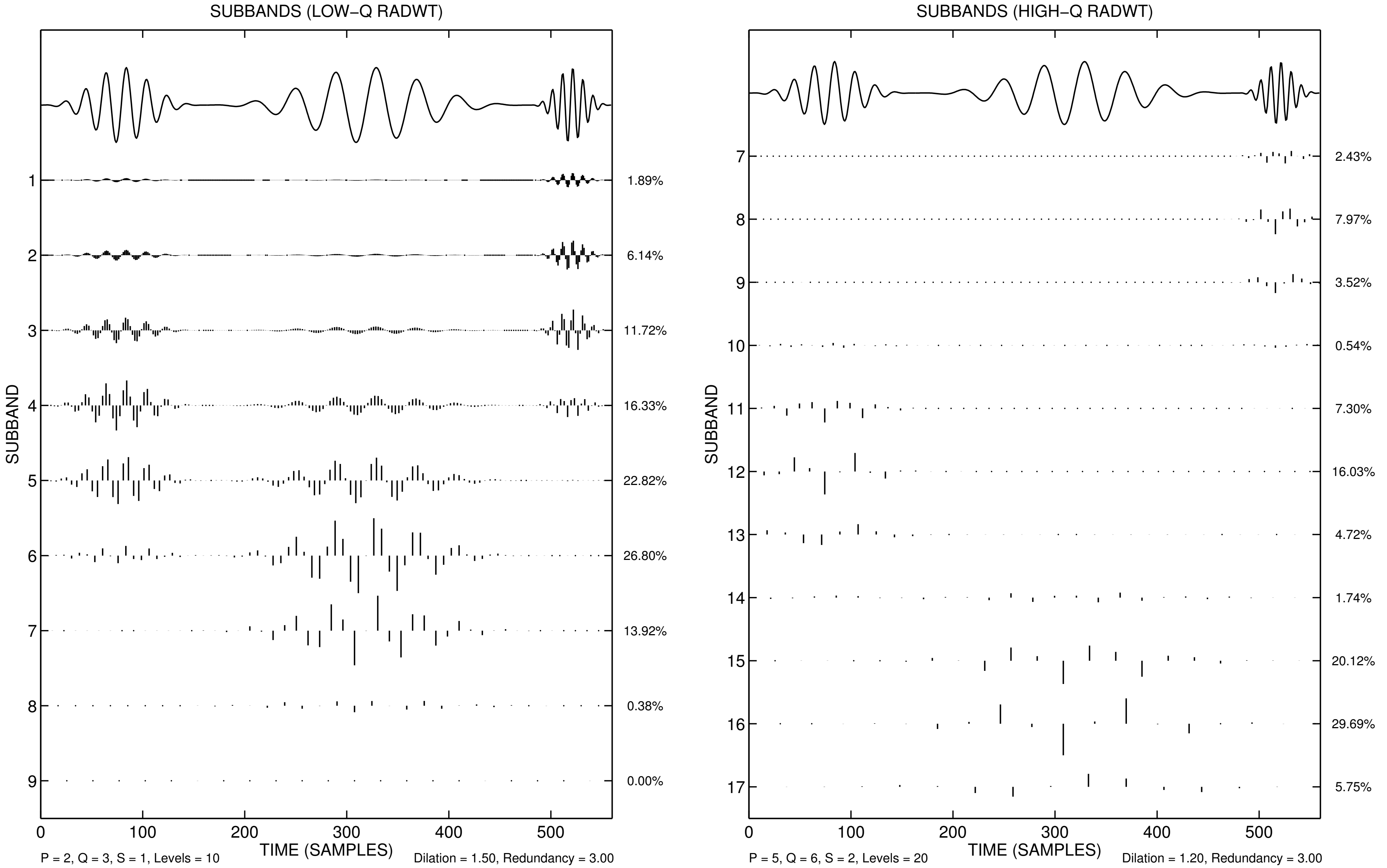


(c) DFT of input signal after filtering.

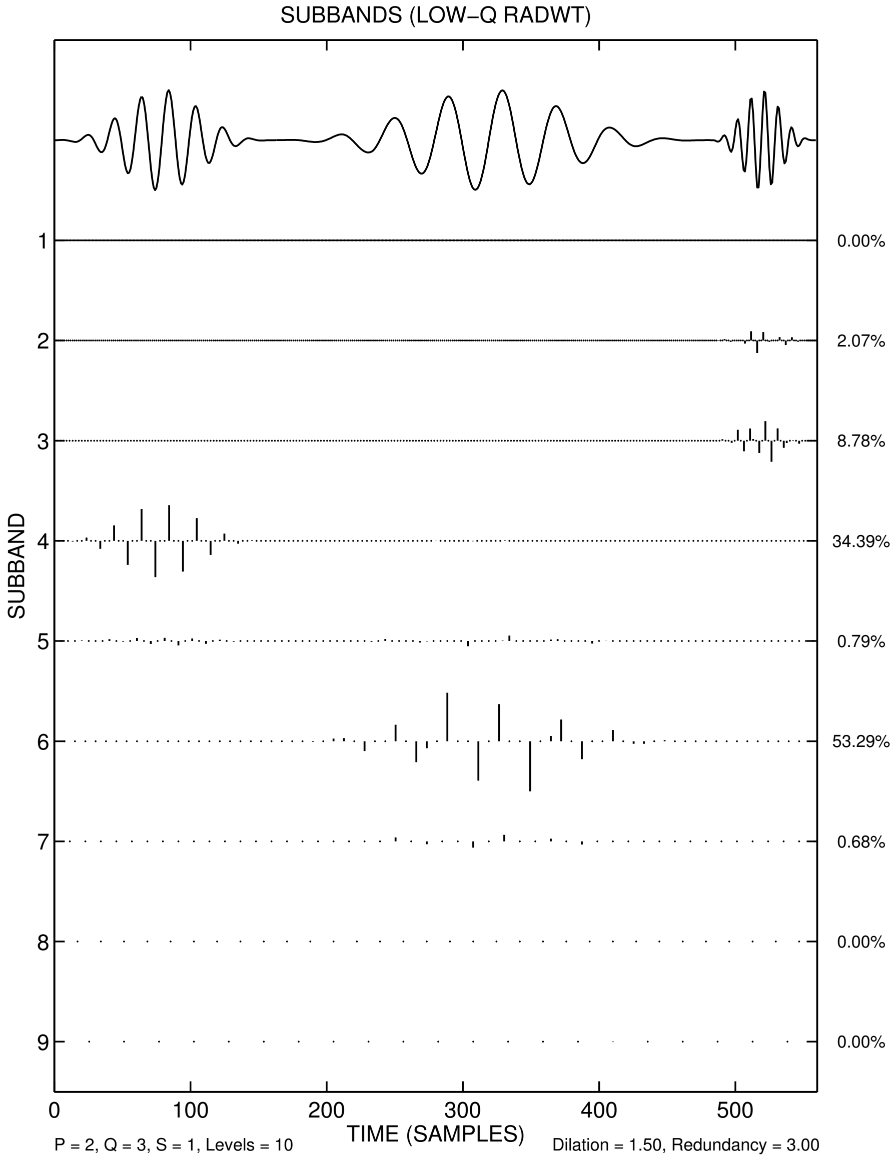


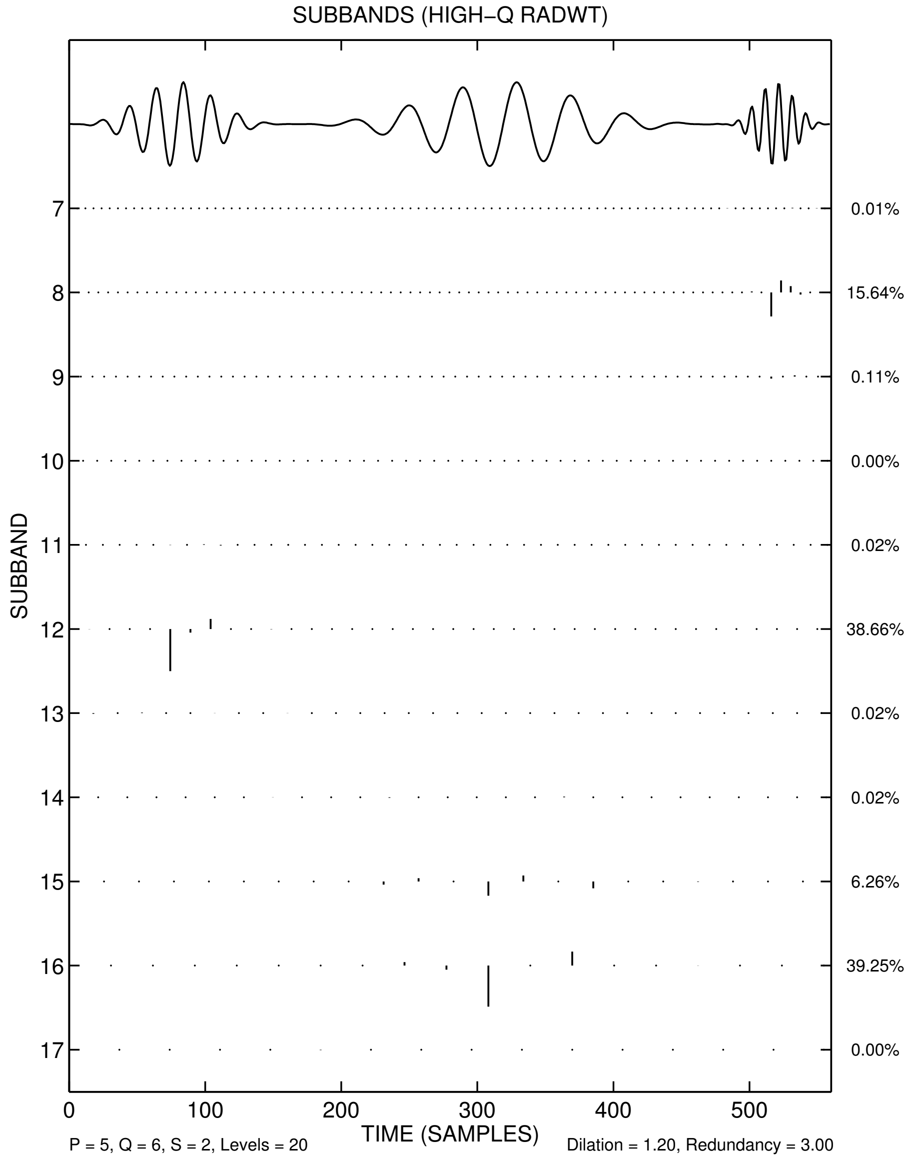
(d) DFT after scaling.

Example: Low Q-factor vs High Q-factor WT

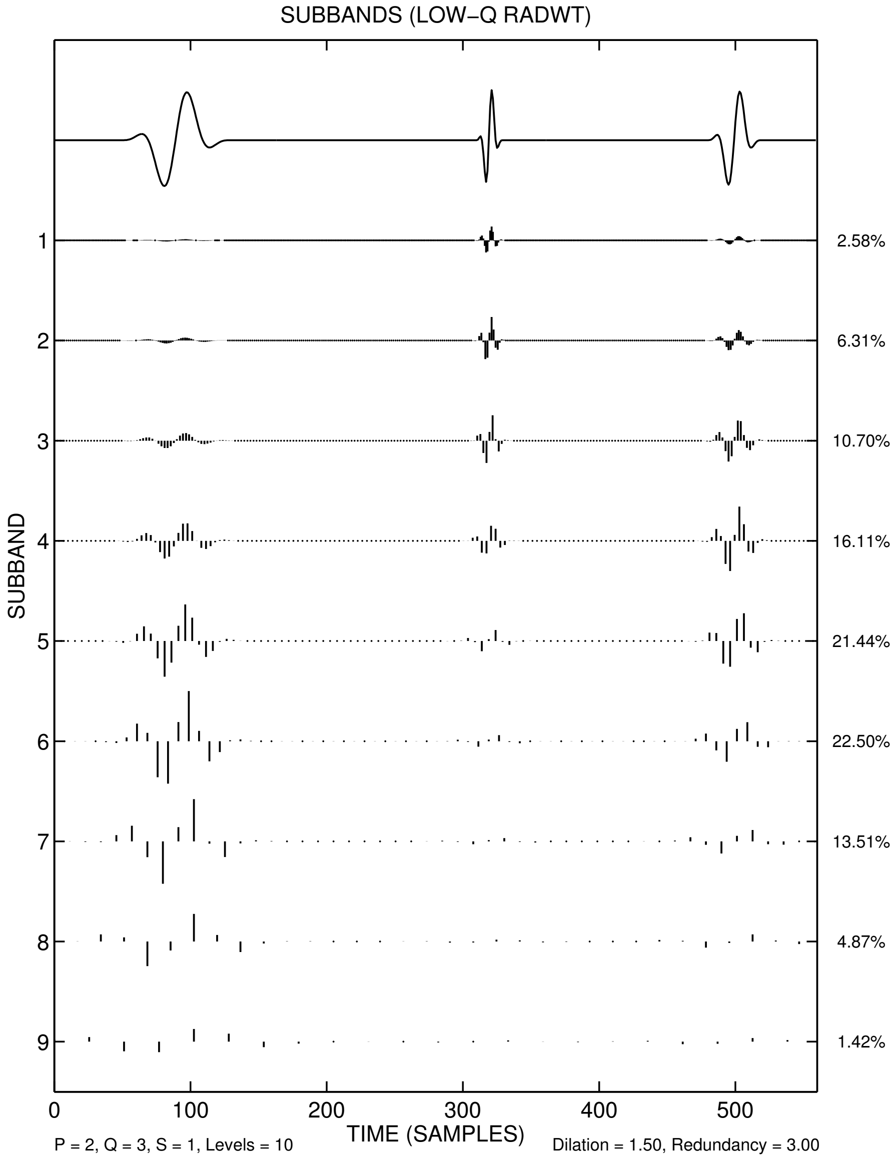


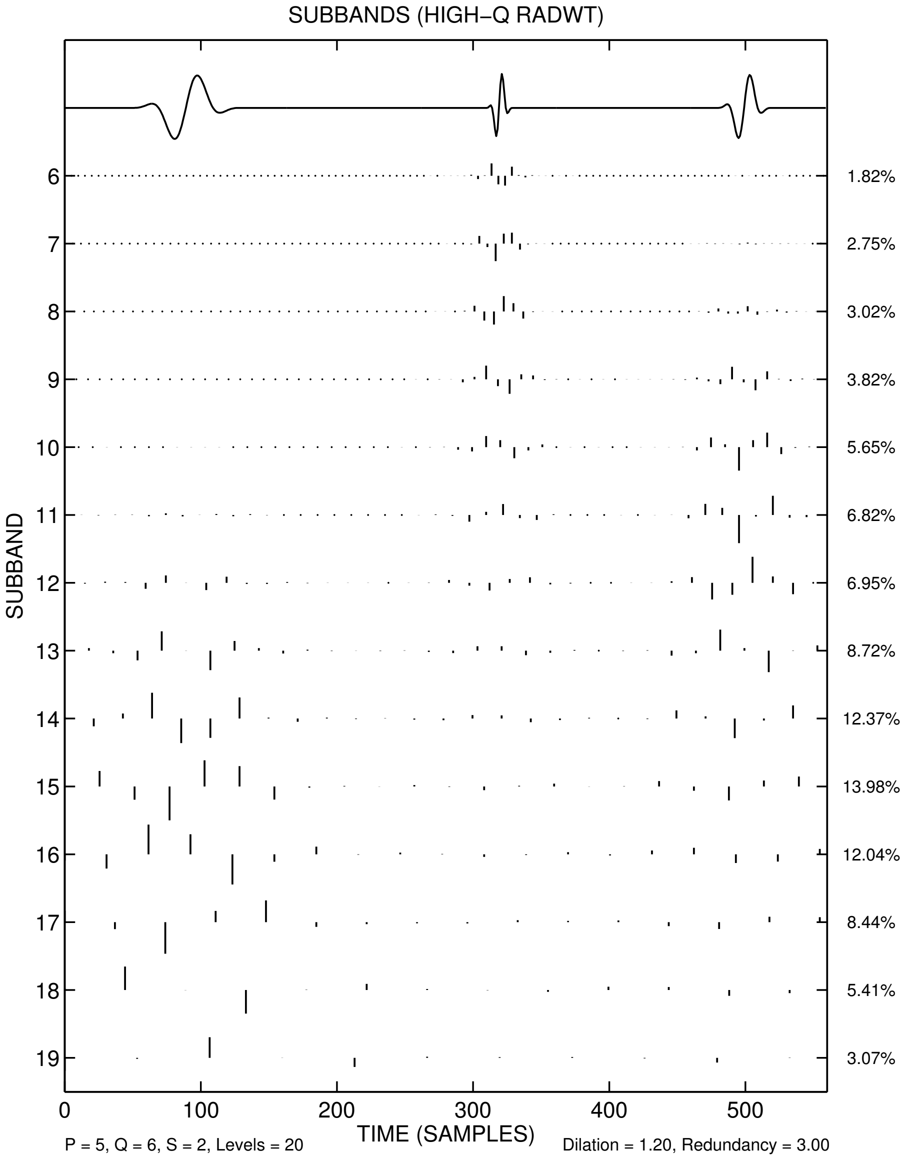
Low Q-factor vs High Q-factor WT after sparsification



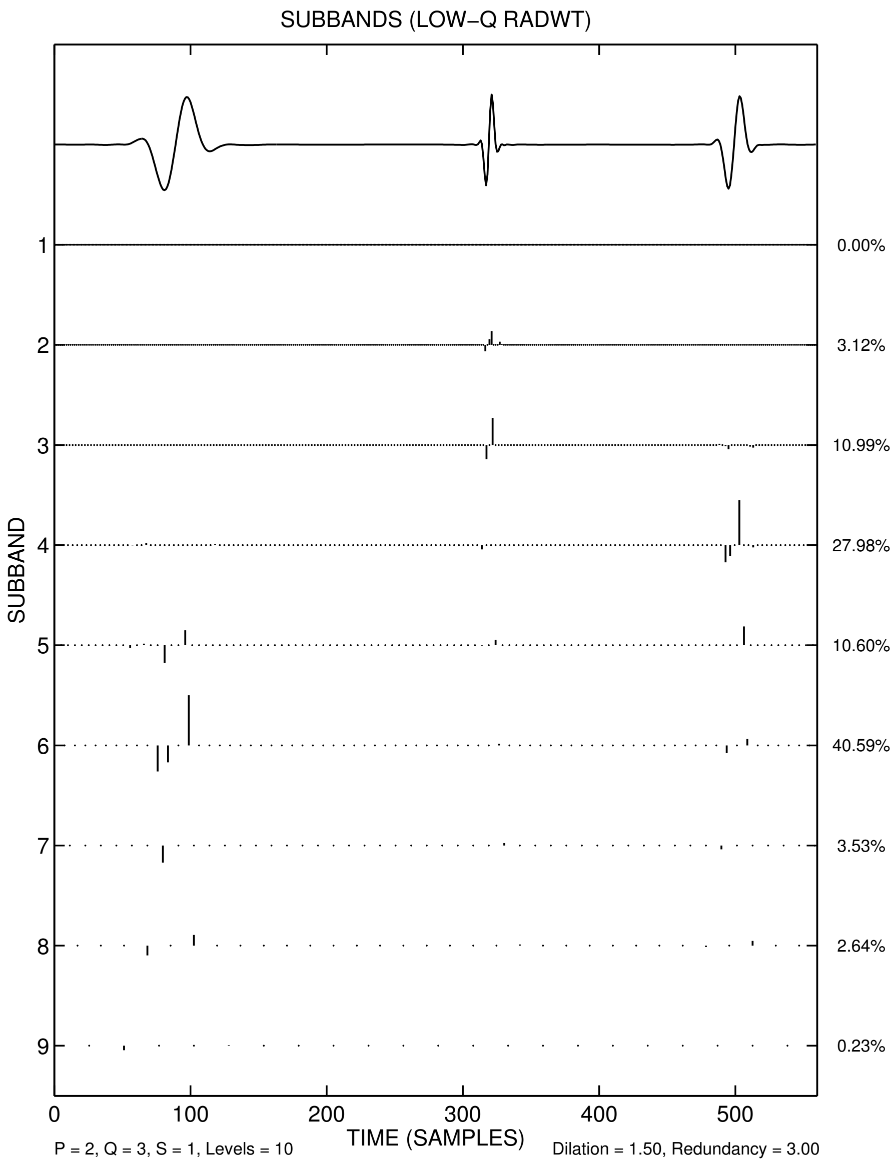


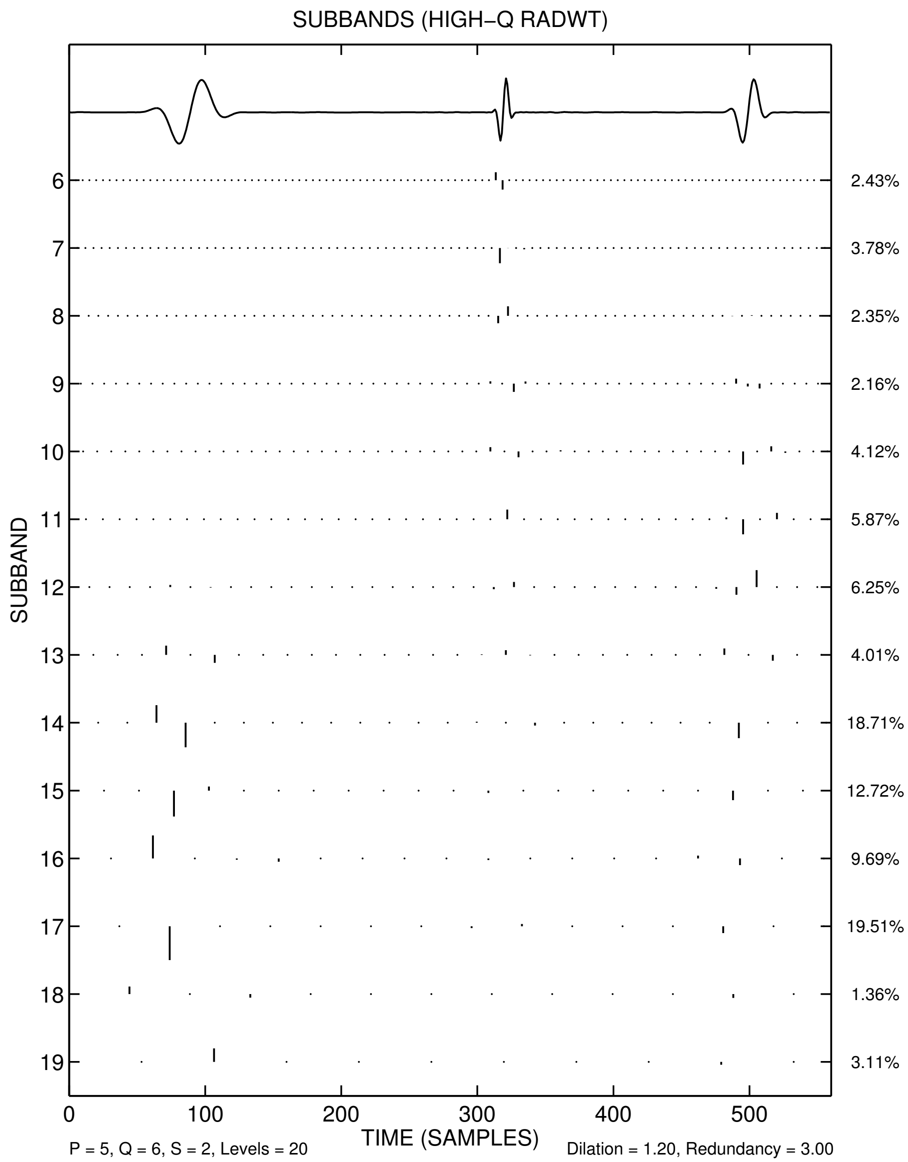
Low Q-factor vs High Q-factor WT





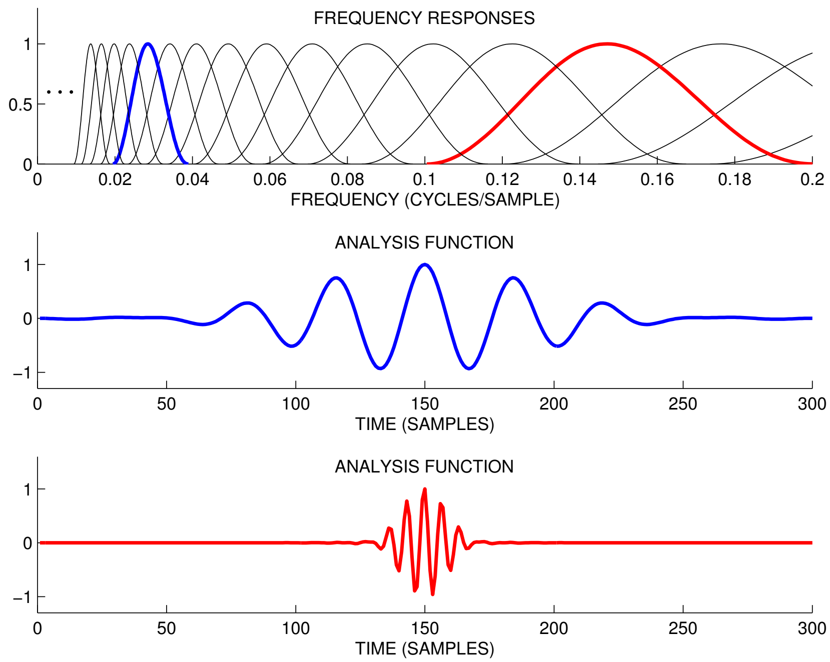
Low Q-factor vs High Q-factor WT after sparsification





Constant-Q vs Constant-BW

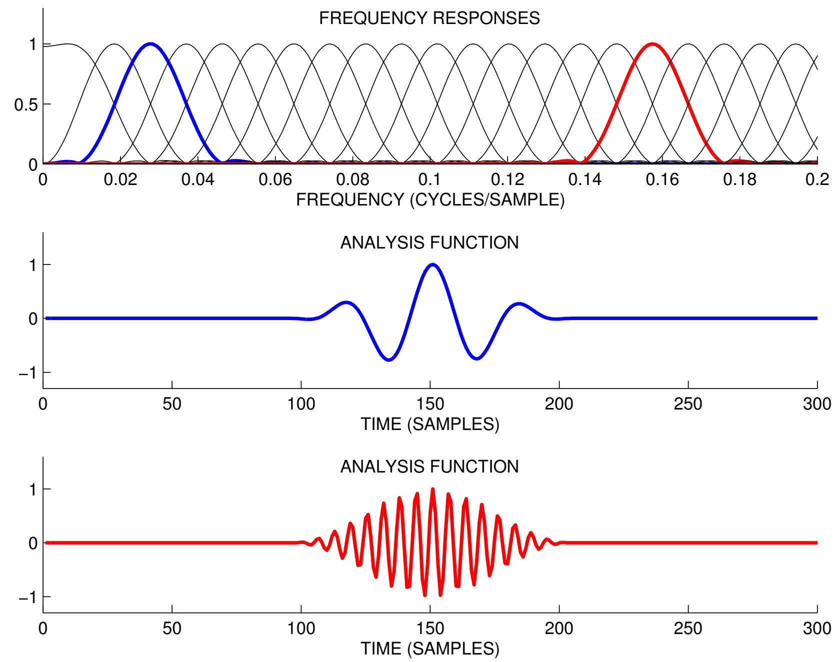
Constant-Q



fixed ‘resonance’

frequency-dependent temporal duration

Constant-BW



frequency-dependent ‘resonance’

fixed temporal duration

Tunable Q-factor wavelet transform (TQWT) — Run Times

Total execution time for forward/inverse N-point TQWT.

N	time (ms)	N	time (ms)
32	0.010	8192	4.938
64	0.025	16384	10.367
128	0.056	32768	21.935
256	0.120	65536	46.280
512	0.260	131072	96.659
1024	0.537	262144	203.580
2048	1.118	524288	448.498
4096	2.350	1048576	1014.719

- Run times measured on a 2010 base-model Apple MacBook Pro (2.4 GHz Intel Core 2 Duo).
- C implementation.

Summary:

1. Fully-discrete, modestly overcomplete
2. Exact perfect reconstruction ('self-inverting')
3. Adjustable Q-factor:

   Can attain higher Q-factors than (or same low Q-factor of) the dyadic WT.

   $\Longrightarrow$ Can achieve higher-frequency resolution needed for oscillatory signals.
4. Samples the time-frequency plane more densely in *both* time *and* frequency.

   $\Longrightarrow$ Exactly invertible, fully-discrete approximation of the continuous WT.
5. FFT-based implementation

Morphological Component Analysis (MCA)

Given an observed signal

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

the goal of MCA is to estimate/determine $\mathbf{x}_1$ and $\mathbf{x}_2$ individually. Assuming that $\mathbf{x}_1$ and $\mathbf{x}_2$ can be sparsely represented in bases (or frames) Φ_1 and Φ_2 respectively, they can be estimated by minimizing the objective function,

$$J(\mathbf{w}_1, \mathbf{w}_2) = \lambda_1 \|\mathbf{w}_1\|_1 + \lambda_2 \|\mathbf{w}_2\|_1$$

with respect to $\mathbf{w}_1$ and $\mathbf{w}_2$, subject to the constraint:

$$\Phi_1 \mathbf{w}_1 + \Phi_2 \mathbf{w}_2 = \mathbf{x}.$$

Then MCA provides the estimates

$$\hat{\mathbf{x}}_1 = \Phi_1 \mathbf{w}_1$$

and

$$\hat{\mathbf{x}}_2 = \Phi_2 \mathbf{w}_2.$$

Reference:

Starck, Elad, Donoho. *Image Decomposition via the Combination of Sparse Representations and a Variational Approach*, IEEE Trans. on Image Processing, Oct 2005.

Why not a quadratic cost function?

If a quadratic cost function is minimized,

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

subject to $\Phi_1 \mathbf{w}_1 + \Phi_2 \mathbf{w}_2 = \mathbf{x}$,

then, using $\Phi_1 \Phi_1^t = \Phi_2 \Phi_2^t = \mathbf{I}$, the $\mathbf{w}_1$ and $\mathbf{w}_2$ can be found in closed form:

$$\mathbf{w}_1 = \frac{\lambda_2^2}{\lambda_1^2 + \lambda_2^2} \, \Phi_1^t \, \mathbf{x},$$

$$\mathbf{w}_2 = \frac{\lambda_1^2}{\lambda_1^2 + \lambda_2^2} \, \Phi_2^t \, \mathbf{x}$$

and the estimated components, $\hat{\mathbf{x}}_1 = \Phi_1 \mathbf{w}_1$ and $\hat{\mathbf{x}}_2 = \Phi_2 \mathbf{w}_2$, are given by

$$\hat{\mathbf{x}}_1 = \frac{\lambda_2^2}{\lambda_1^2 + \lambda_2^2} \, \mathbf{x},$$

$$\hat{\mathbf{x}}_2 = \frac{\lambda_1^2}{\lambda_1^2 + \lambda_2^2} \, \mathbf{x}$$

Both $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$ are just scaled versions of $\mathbf{x}$.

$\Longrightarrow$ No separation at all!

MCA as a linear inverse problem

The constrained optimization problem

$$\min_{\mathbf{w}_1, \mathbf{w}_2} \quad \lambda_1 \|\mathbf{w}_1\|_1 + \lambda_2 \|\mathbf{w}_2\|_1$$

$$\text{subject to} \quad \Phi_1 \mathbf{w}_1 + \Phi_2 \mathbf{w}_2 = \mathbf{x}$$

can be written as

$$\min_{\mathbf{w}} \quad \|\boldsymbol{\lambda} \odot \mathbf{w}\|_1$$

$$\text{subject to} \quad \mathbf{H}\mathbf{w} = \mathbf{x}$$

where

$$\mathbf{H} = \begin{bmatrix} \Phi_1 & \Phi_2 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix}$$

and $\odot$ denotes point-by-point multiplication.

This is 'basis pursuit', an ℓ_1-regularized linear inverse problem ...

- Non-differentiable
- Convex

$\Longrightarrow$ We use a variant of SALSA (split augmented Lagrangian shrinkage algorithm).

Reference: Afonso, Bioucas-Dias, Figueiredo. *Fast Image Recovery Using Variable Splitting and Constrained Optimization.* IEEE Trans. on Image Processing, 2010.

SALSA for MCA (bais pursuit form)

Applying SALSA to the MCA problem yields the iterative algorithm:

$$\text{initialize: } \mu > 0, \mathbf{d}_i \tag{5}$$

$$\mathbf{u}_i \leftarrow \operatorname{soft}(\mathbf{w}_i + \mathbf{d}_i, 0.5\boldsymbol{\lambda}_i/\mu) - \mathbf{d}_i, \qquad i = 1, 2 \tag{6}$$

$$\mathbf{c} \leftarrow \mathbf{x} - \Phi_1\, \mathbf{u}_1 - \Phi_2\, \mathbf{u}_2 \tag{7}$$

$$\mathbf{d}_i \leftarrow \frac{1}{2}\Phi_i^t\, \mathbf{c} \qquad i = 1, 2 \tag{8}$$

$$\mathbf{w}_i \leftarrow \mathbf{d}_i + \mathbf{u}_i \qquad i = 1, 2 \tag{9}$$

$$\text{repeat} \tag{10}$$

where $\operatorname{soft}(x, T)$ is the soft-threshold rule with threshold T,

$$\operatorname{soft}(x, T) = x \max(0, 1 - T/|x|).$$

Note: no matrix inverses; only forward and inverse transforms.

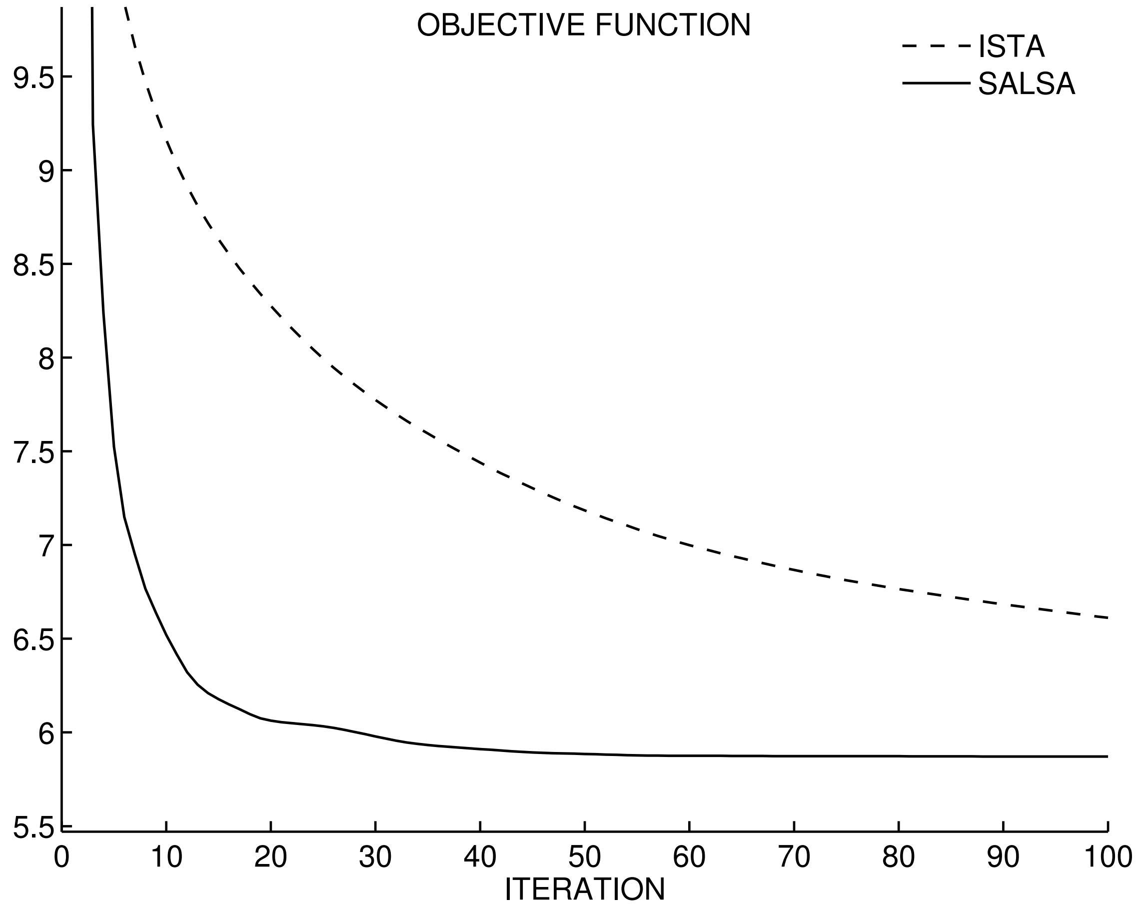


Figure 4: Reduction of objective function during the first 100 iterations. SALSA converges faster than ISTA.

Example: Resonance-selective nonlinear band-pass filtering

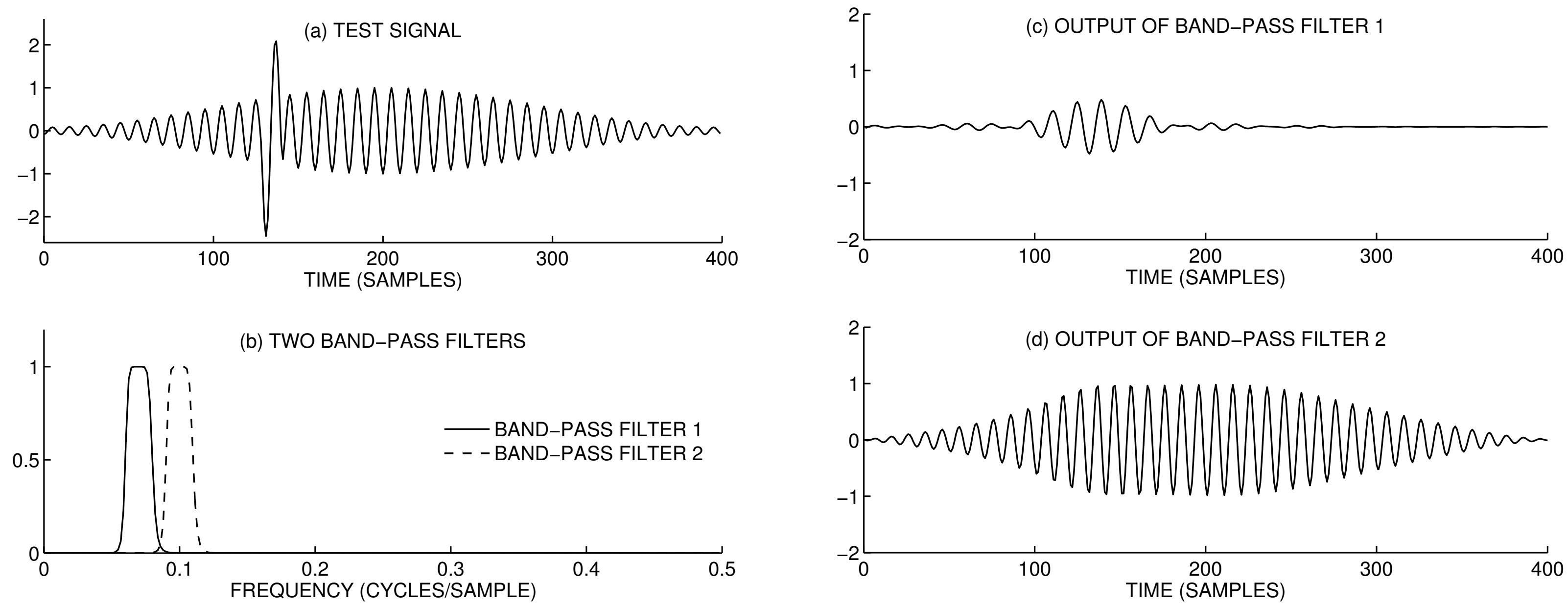


Figure 5: LTI band-pass filtering. The test signal (a) consists of a sinusoidal pulse of frequency 0.1 cycles/sample and a transient. Band-pass filters 1 and 2 in (b) are tuned to the frequencies 0.07 and 0.10 cycles/second respectively. The output signals, obtained by filtering the test signal with each of the two band-pass filters, are shown in (c) and (d). The output of band-pass filter 1, illustrated in (c), contains oscillations due to the transient in the test signal. Moreover, the transient oscillations in (c) have a frequency of 0.07 Hz even though the test signal (a) contains no sustained oscillatory behavior at this frequency.

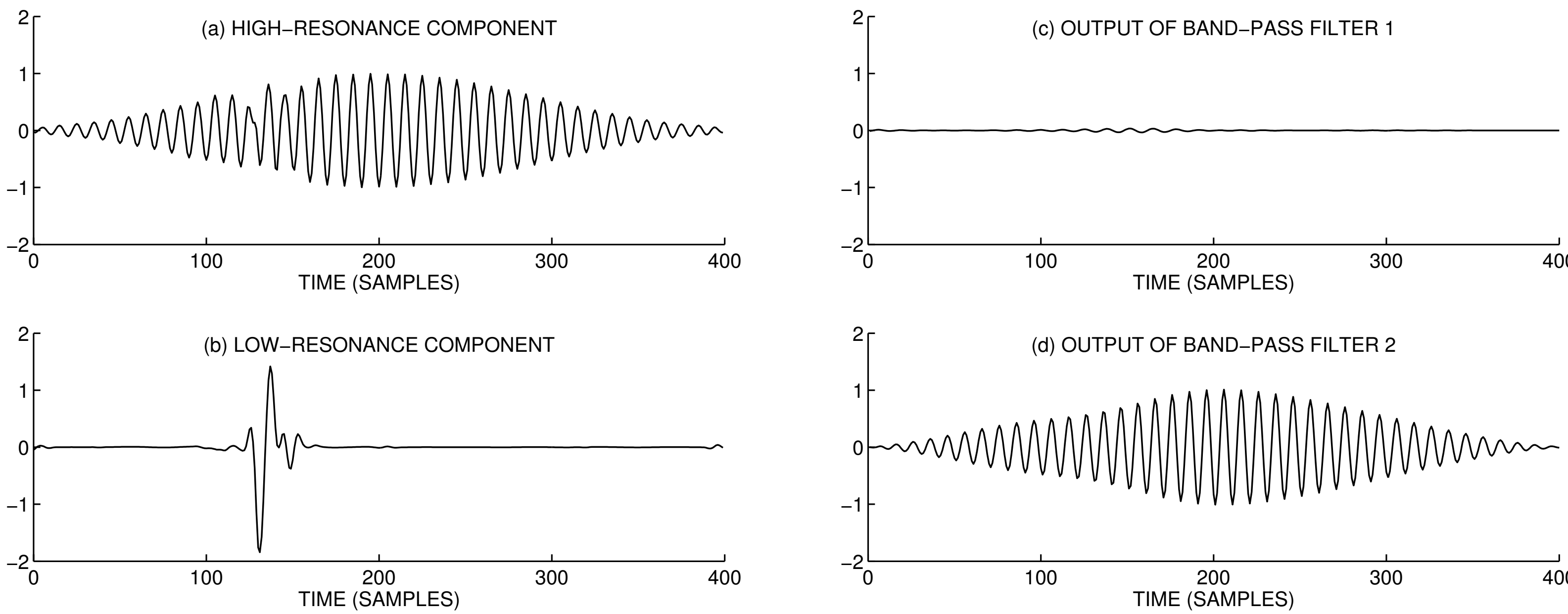


Figure 6: Resonance-based decomposition and band-pass filtering. When resonance-based analysis method is applied to the test signal in Fig. 5a, it yields the high- and low-resonance components illustrated in (a) and (b). The output signals, obtained by filtering the high-resonance component (a) with each of the two band-pass filters shown in Fig. 5b, are illustrated in (c) and (d). The transient oscillations in (c) are substantially reduced compared to Fig. 5c.

Example: Resonance-based decomposition of speech

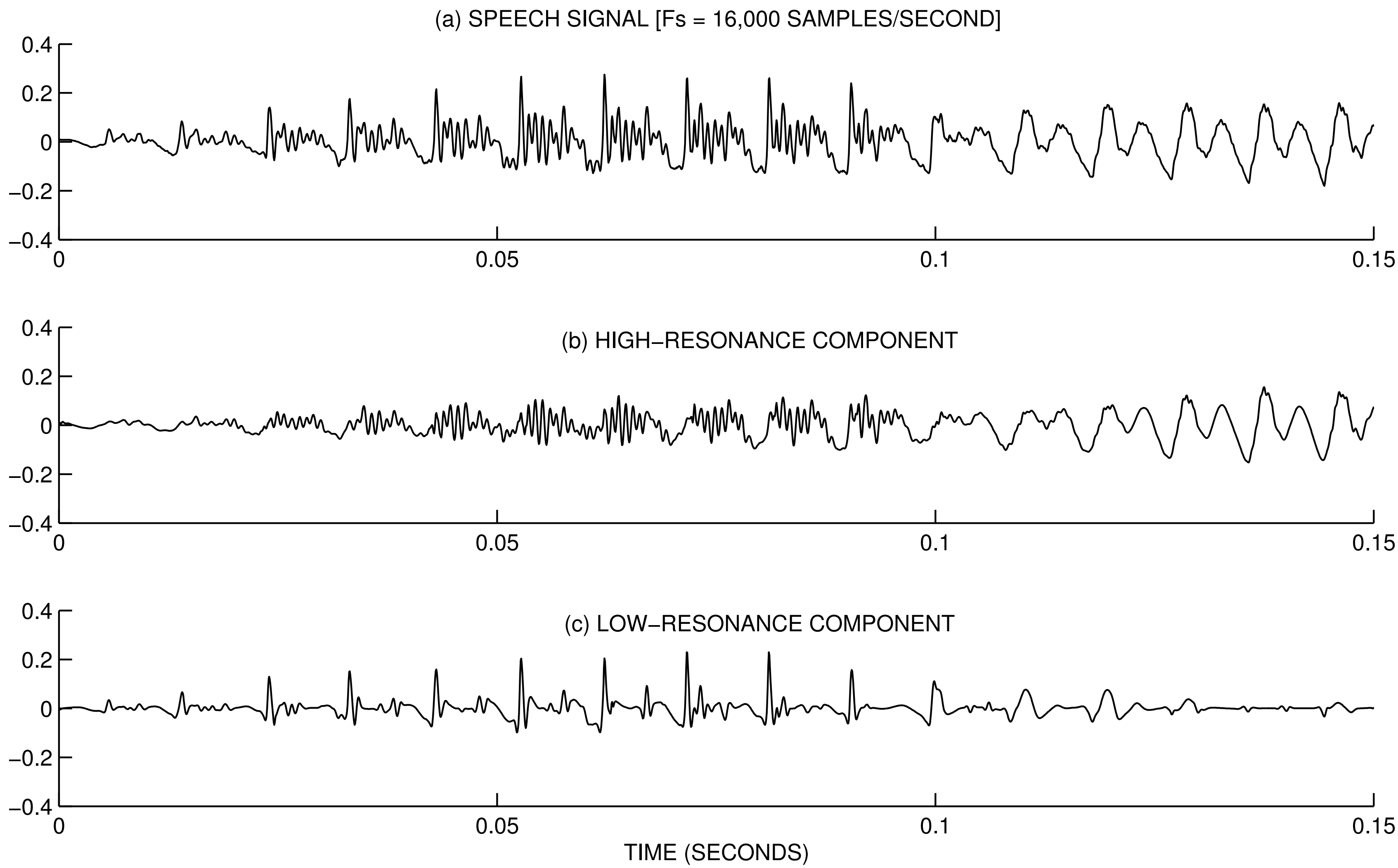


Figure 7: Decomposition of a speech signal ("I'm") into high- and low-resonance components. The high-resonance component (b) contains the sustained oscillations present in the speech signal, while the low-resonance component (c) contains non-oscillatory transients. (The residual is not shown.)

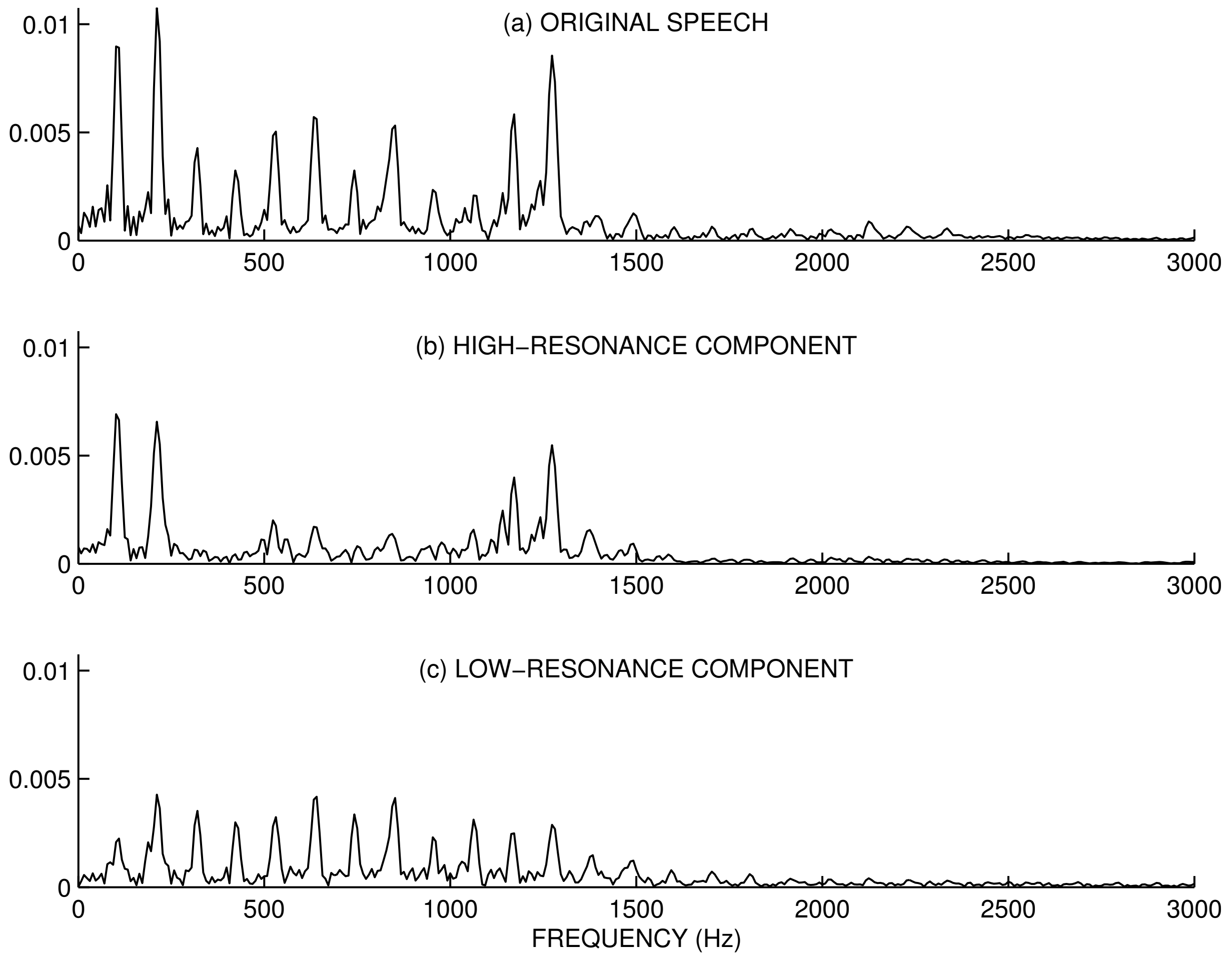


Figure 8: Frequency spectra of the speech signal in Fig. 7 and of the extracted high- and low-resonance components. The spectra are computed using the 50 msec segment from 0.05 to 0.10 seconds. The energy of each resonance component is widely distributed in frequency and their frequency-spectra overlap.

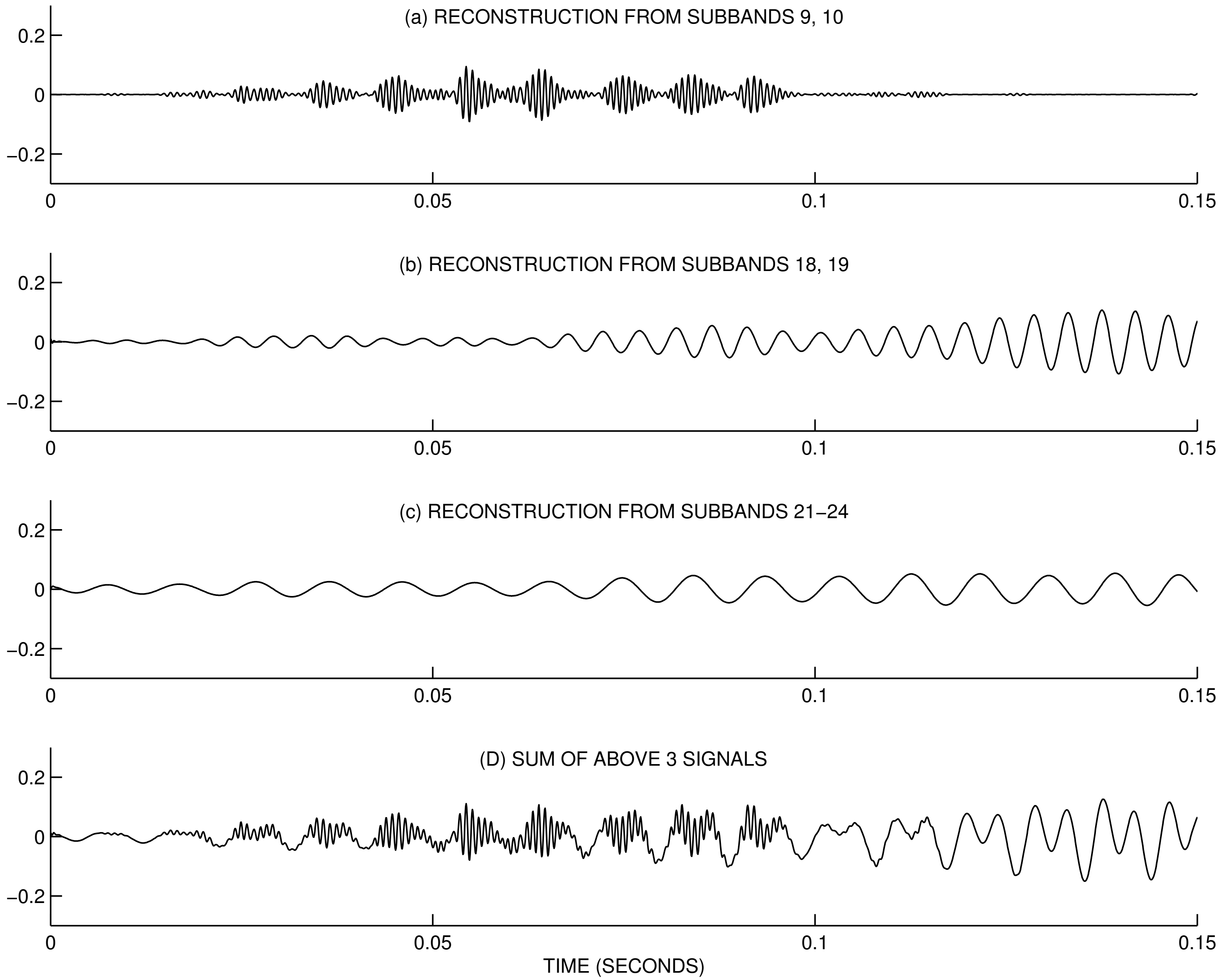
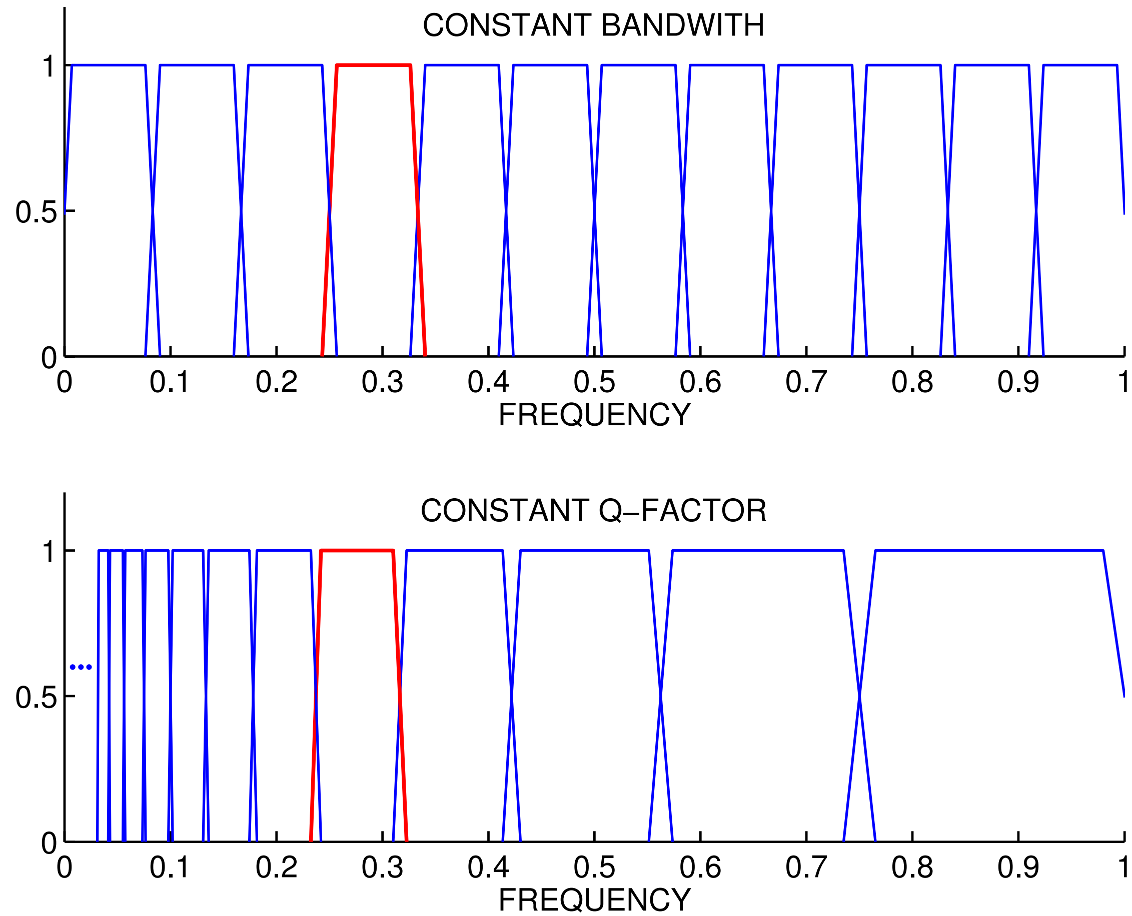


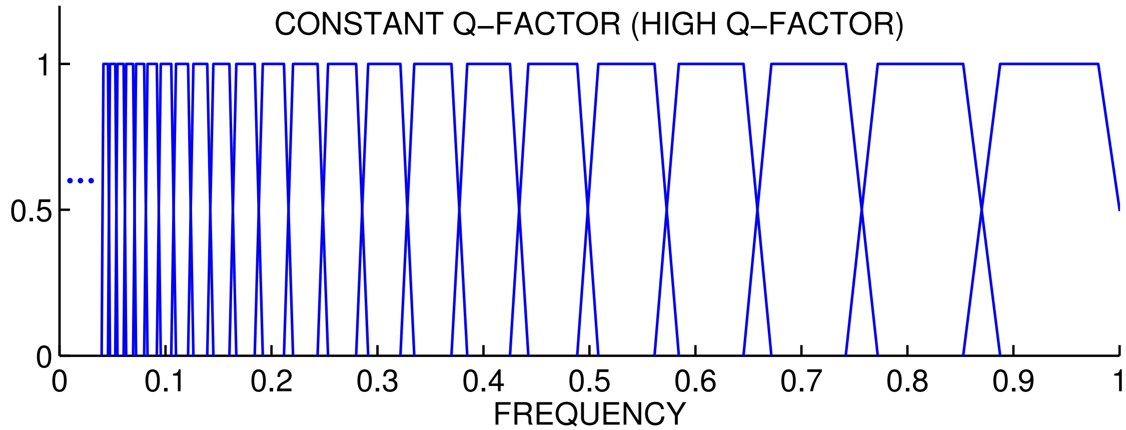
Figure 9: Frequency decomposition of high-resonance component in Fig. 7. Reconstructing the high-resonance component from a few subbands of the high Q-factor WT at a time, yields an efficient AM/FM decomposition.

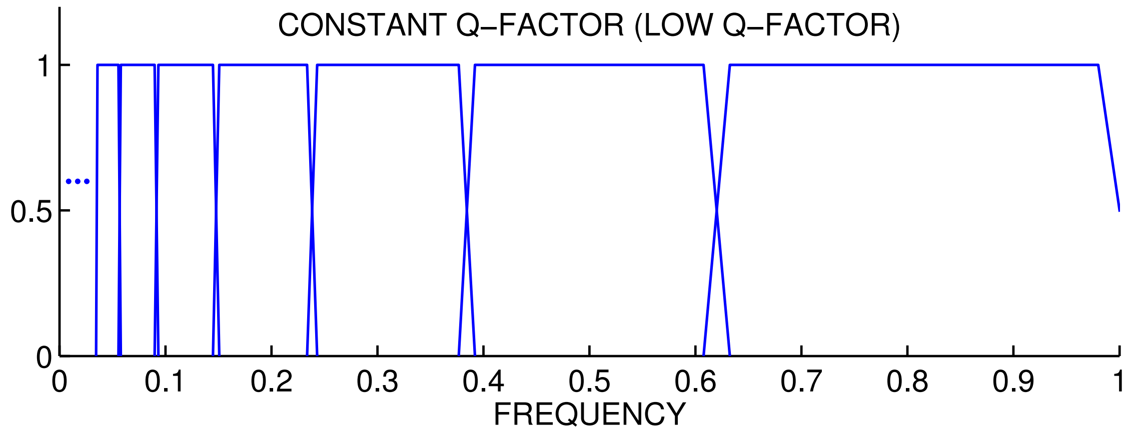
Constant bandwidth + Constant Q-factor



A constant bandwidth and a constant Q-factor decomposition can have high coherence due to some analysis functions, from each decomposition, having similar frequency support. This can degrade the results of MCA in principle.

Constant Q-factor: High Q-factor + Low Q-factor





Two constant Q-factor decompositions with markedly different Q-factors will have low coherence because no analysis functions from the two decompositions will have similar frequency support. This is beneficial for the operation of MCA.

Small coherence between low and high Q-factor WTs

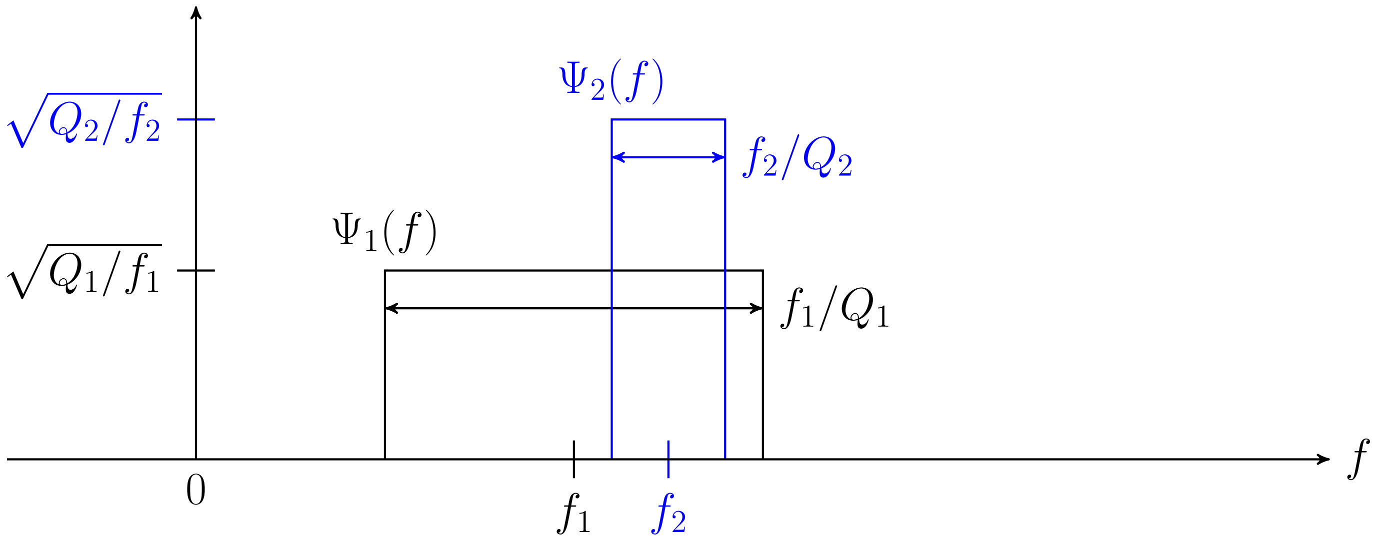


Figure 10: For reliable resonance-based decomposition, the inner product between the low-Q and high-Q wavelets should be small for all dilations and translations. The computation of the maximum inner product is simplified by assuming the wavelets are ideal band-pass functions and expressing the inner product in the frequency domain.

The inner products can be defined in the frequency domain,

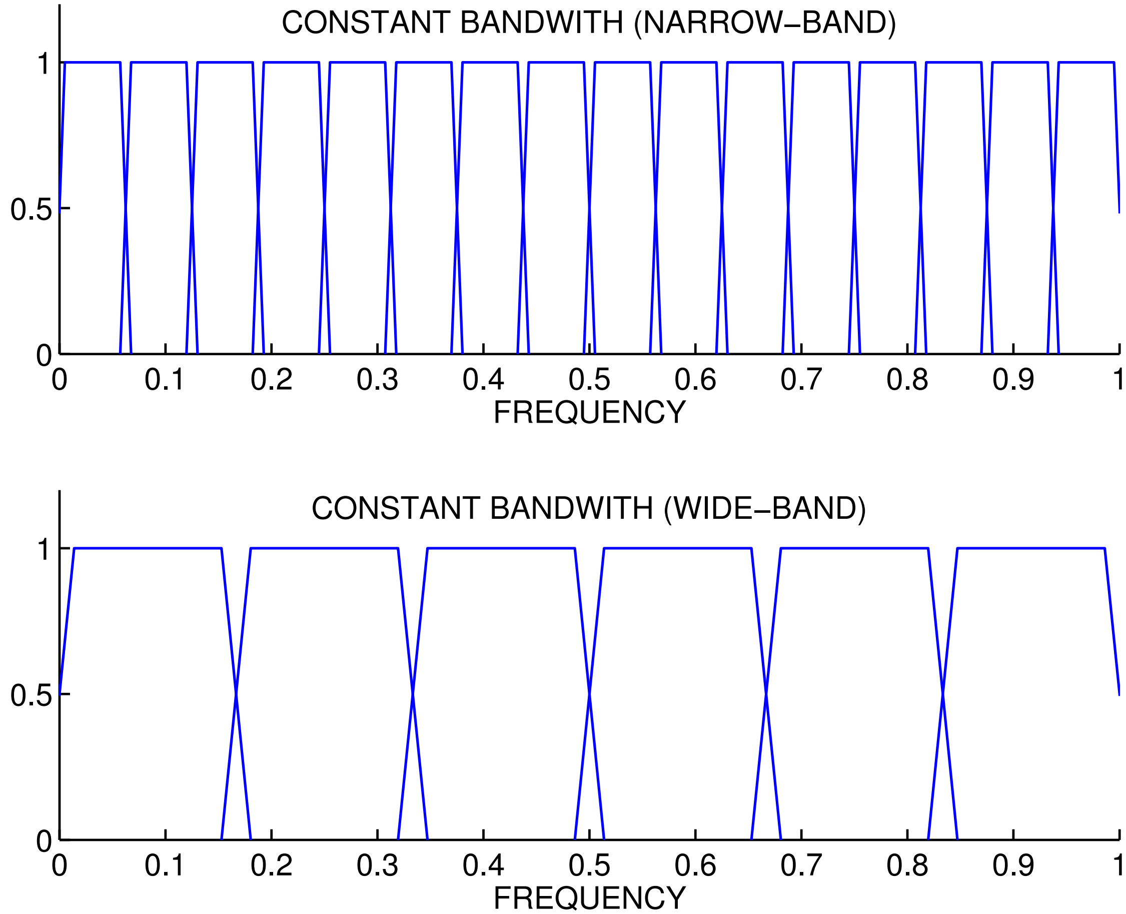
$$\rho(f_1, f_2) := \int \Psi_1(f)\Psi_2(f)\, df,$$

as a function of their center frequencies (equivalently, dilation).

The maximum value of the inner product, $\rho(f_1, f_2)$, occurs when $f_2 = f_1(2 + 1/Q_1)/(2 + 1/Q_2)$ and is given by

$$\rho_{\mathsf{max}} = \sqrt{\frac{Q_1 + 1/2}{Q_2 + 1/2}}, \quad Q_2 > Q_1. \tag{11}$$

Constant Bandwidth: Narrow-band + Wide-band



Two constant bandwidth decompositions with markedly different bandwidths will also have low coherence and are therefore also suitable transform for MCA-based signal decomposition. This gives a bandwidth-based decomposition, rather than a resonance-based decomposition.

Conclusion: Resonance-based signal decomposition

Low Q-factor WT used for sparse representation of the *transient* component.

High Q-factor WT used for sparse representation of the *oscillatory* (rhythmic) component.

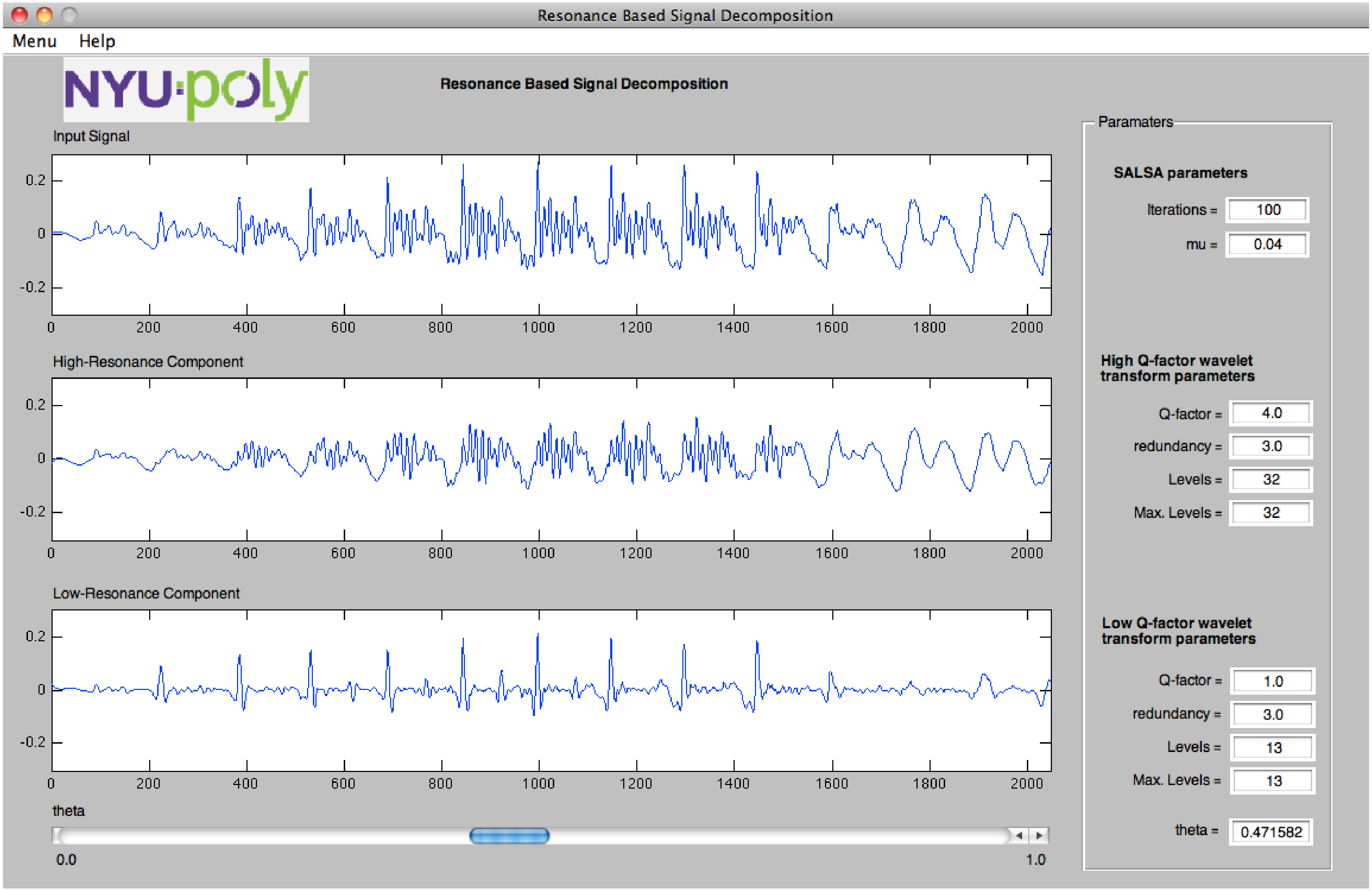
Morphological component analysis (MCA) used to separate the two signal components.

- Oscillatory component not necessarily high-pass — contains both low and high frequencies.

- Transient component not necessarily a low-pass signal — contains sharp bumps and jumps.

- Software available (Matlab software on web, C code by request).
- `http://eeweb.poly.edu/iselesni/TQWT/`

Graphical User Interface: Facilitates interactive selection and tuning of parameters.



Group Sparsity with Fully Overlapping Groups

Po-Yu Chen

Basis pursuit denoising (no group structure):

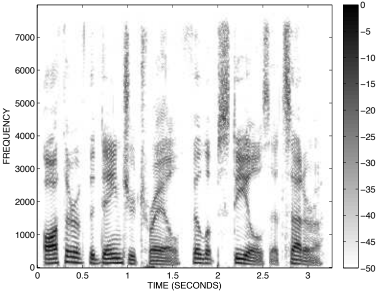
$$\underset{c}{\operatorname{argmin}} \; \|\mathbf{y} - \Phi\mathbf{c}\|_2^2 + \lambda\|\mathbf{c}\|_1$$

Generalized basis pursuit denoising (overlapping group sparsity):

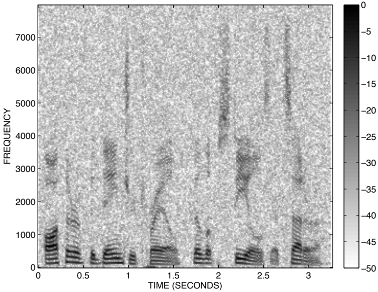
$$\underset{c}{\operatorname{argmin}} \; \|\mathbf{y} - \Phi\mathbf{c}\|_2^2 + \lambda \sum_{i=0}^{N-3} \sqrt{|c(i)|^2 + |c(i+1)|^2 + |c(i+2)|^2}$$

(group size 3)

Example: Speech Enhancement



Noise-free signal spectrogram

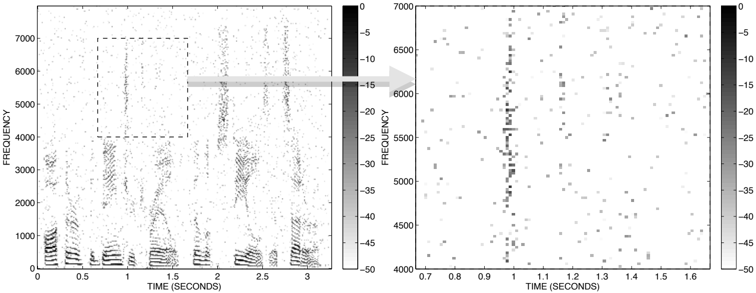


Noisy signal spectrogram

Example: Speech Enhancement

$(K_1, K_2) = (1, 1)$

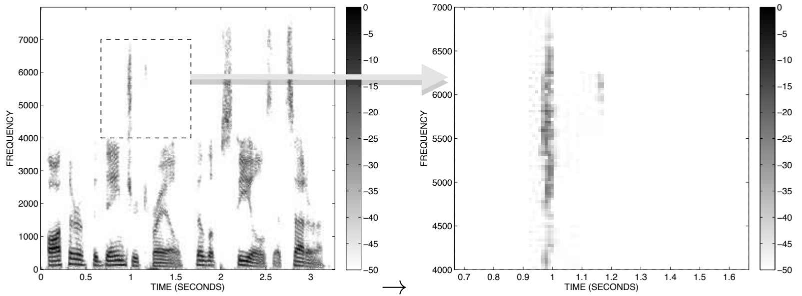
Basis pursuit denoising using ℓ_1 norm, i.e. no group structure.



The spectrogram exhibits many spurious noise spikes which produces 'musical noise'

Example: Speech Enhancement

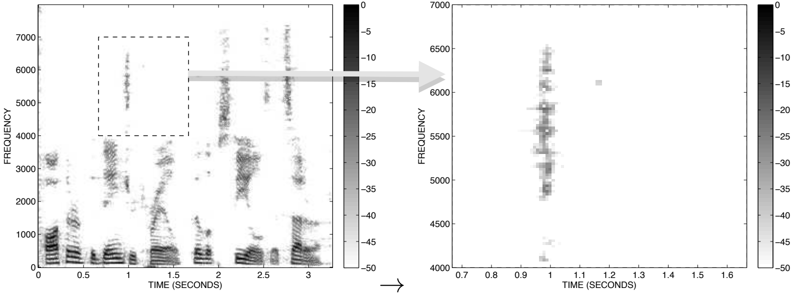
$$(K_1,\ K_2) = (2,\ 8)$$



The spectrogram does not exhibit spurious noise spikes. The denoised signal is free of ‘musical noise’ artifact

Example: Speech Enhancement

$(K_1,\ K_2) = (8,\ 2)$



The spectrogram does not exhibit spurious noise spikes. The denoised signal is free of ‘musical noise’ artifact

Quasi-Polynomial Approximation via Sparse Derivatives

Xiaoran Ning

$$\mathbf{y} = \mathbf{x} + \text{noise}$$

Total variation denoising:

$$\underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{D}\mathbf{x}\|_1$$

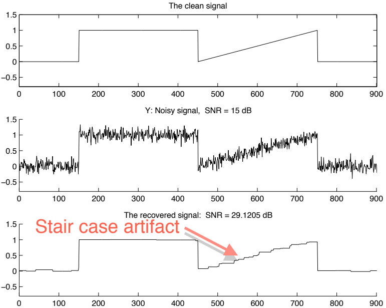
First order difference (derivative)

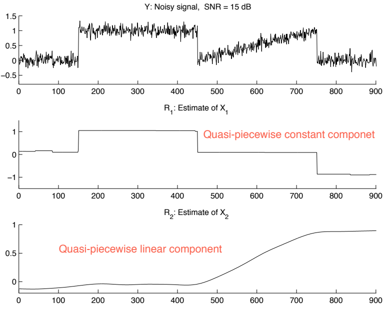
Dual-component polynomial denoising:

$$\underset{\mathbf{x}_1, \mathbf{x}_2}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{x}_1 - \mathbf{x}_2\|_2^2 + \lambda_1 \|\mathbf{D}\mathbf{x}_1\|_1 + \lambda_2 \|\mathbf{D}^2\mathbf{x}_2\|_1$$

Second order derivative

Total Variation Denoising





The recovered signal = $R_1 + R_2$, SNR = 31.2347 dB

To compare the two methods

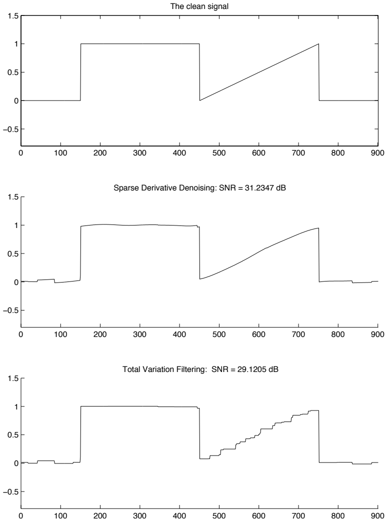


Figure: Sparse Derivative Denoising and TV Filtering.

More slides...

Comparison: EMD & TQWT

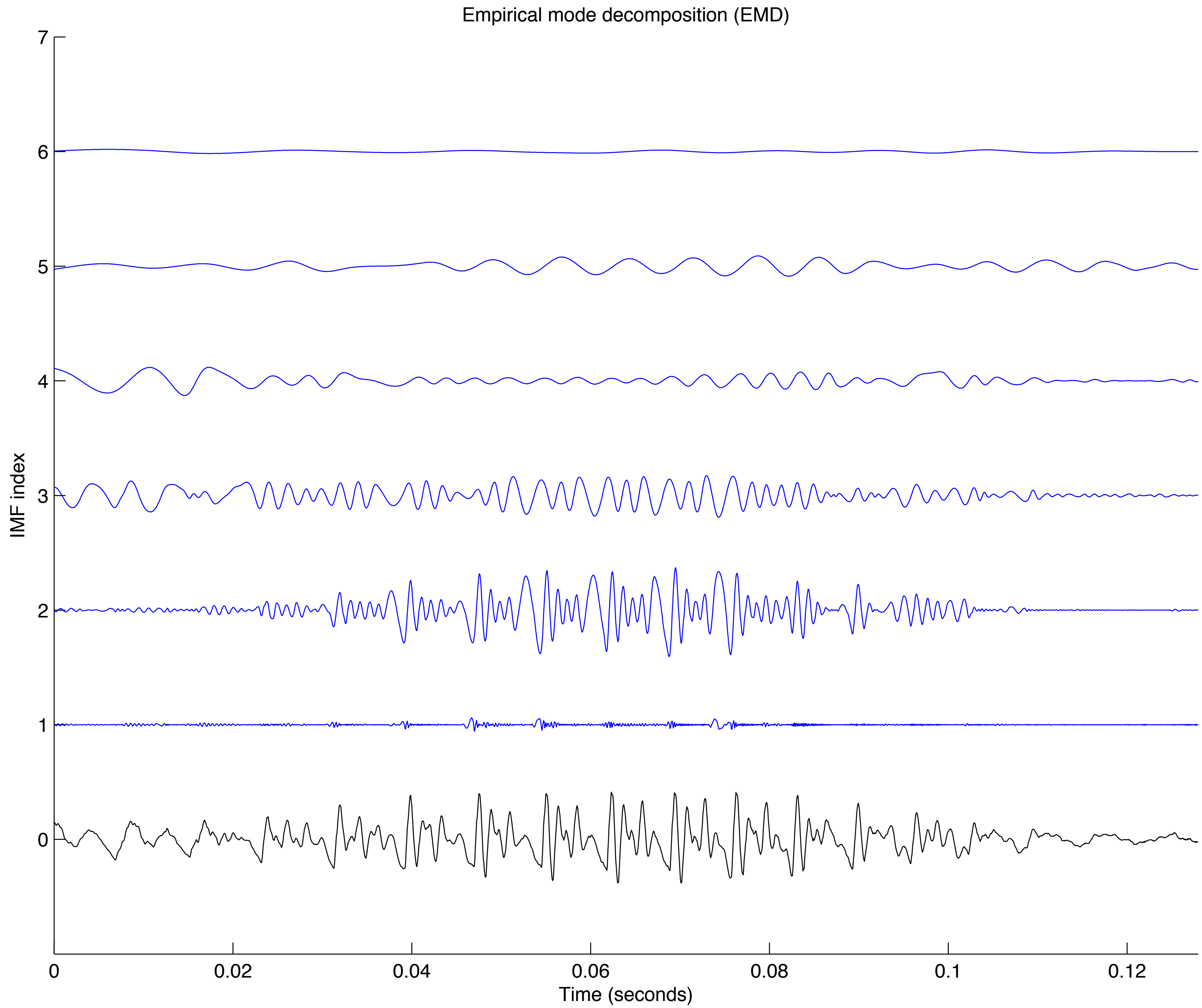
We compare

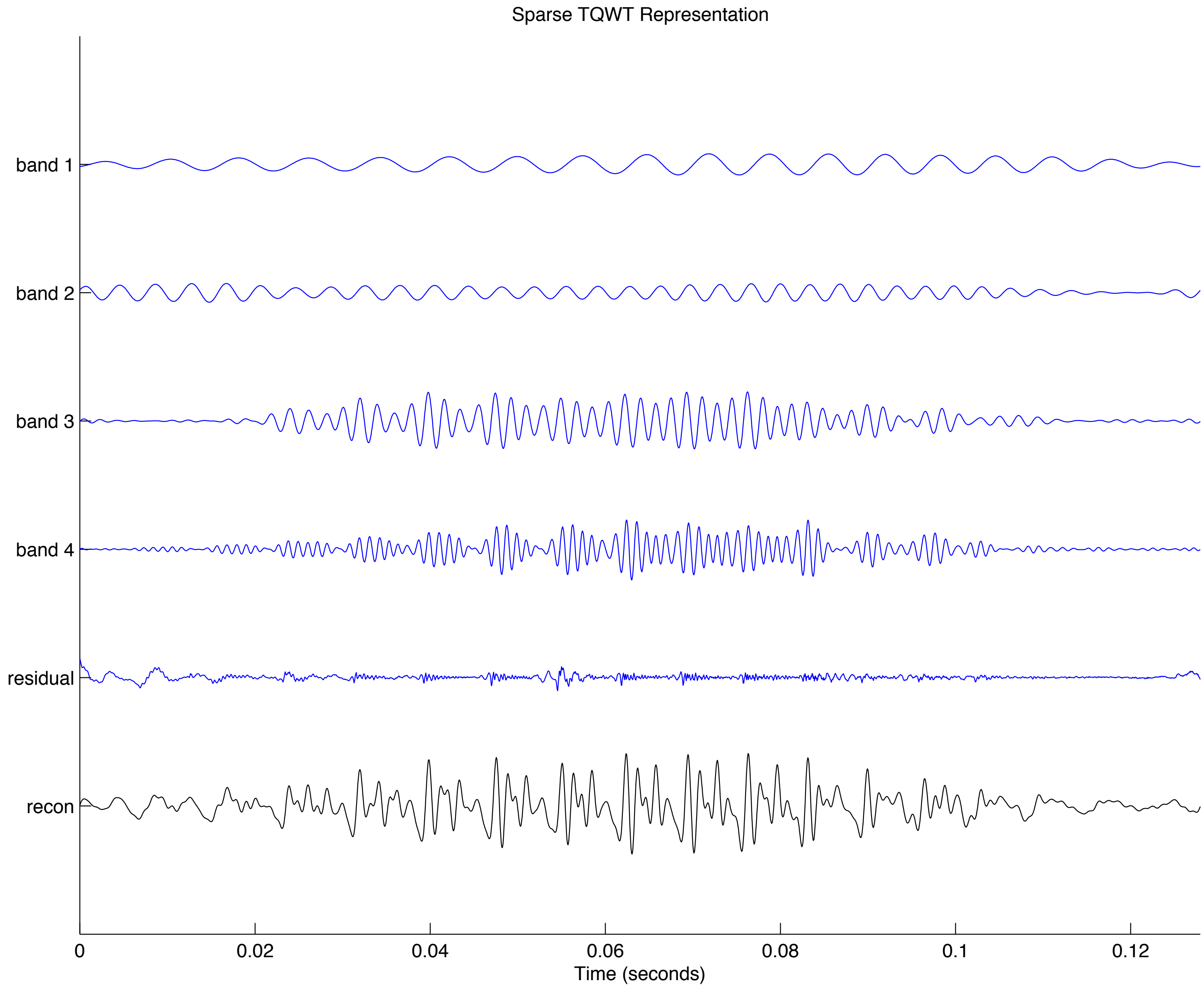
- Empirical Mode Decomposition (EMD)

- Sparse TQWT Representation

A goal of EMD is to decompose a multicomponent signal into several narrow-band components (intrinsic mode functions).

For some real signals, a sparse TQWT representation leads to a more reasonable decomposition than EMD (next two slides).


Empirical mode decomposition (EMD)
IMF index
7
6
5
4
3
2
1
0
0
0.02
0.04
0.06
0.08
0.1
0.12
Time (seconds)



Sparse TQWT Representation
band 1
band 2
band 3
band 4
residual
recon
0
0.02
0.04
0.06
0.08
0.1
0.12
Time (seconds)

Morphological Component Analysis (MCA) — Noisy signal case

In the noisy case, we should not ask for exact equality.

Given an observed signal

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{n}, \quad \text{with} \quad \mathbf{x},\ \mathbf{x}_1,\ \mathbf{x}_2 \in \mathbb{R}^N,$$

where $\mathbf{n}$ is noise, the components $\mathbf{x}_1$ and $\mathbf{x}_2$ can be estimated by minimizing the objective function,

$$J(\mathbf{w}_1, \mathbf{w}_2) = \|\mathbf{x} - \Phi_1\mathbf{w}_1 - \Phi_2\mathbf{w}_2\|_2^2 + \lambda_1\|\mathbf{w}_1\|_1 + \lambda_2\|\mathbf{w}_2\|_1$$

with respect to $\mathbf{w}_1$ and $\mathbf{w}_2$. Then MCA provides the estimates

$$\hat{\mathbf{x}}_1 = \Phi_1\mathbf{w}_1$$

and

$$\hat{\mathbf{x}}_2 = \Phi_2\mathbf{w}_2.$$

Reference:

Starck, Elad, Donoho. *Image Decomposition via the Combination of Sparse Representations and a Variational Approach*, IEEE Trans. on Image Processing, Oct 2005.

Why not a quadratic cost function?

If the ℓ_2-norm is used for the penalty term,

$$J(\mathbf{w}_1, \mathbf{w}_2) = \|\mathbf{x} - \Phi_1\mathbf{w}_1 - \Phi_2\mathbf{w}_2\|_2^2 + \lambda_1\|\mathbf{w}_1\|_2^2 + \lambda_2\|\mathbf{w}_2\|_2^2,$$

then, using $\Phi_1\Phi_1^t = \Phi_2\Phi_2^t = \mathbf{I}$, the minimizing $\mathbf{w}_1$ and $\mathbf{w}_2$ can be found in closed form:

$$\mathbf{w}_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_1\,\lambda_2}\,\Phi_1^t\,\mathbf{x}$$

$$\mathbf{w}_2 = \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_1\,\lambda_2}\,\Phi_2^t\,\mathbf{x}$$

and the estimated components, $\hat{\mathbf{x}}_1 = \Phi_1\mathbf{w}_1$ and $\hat{\mathbf{x}}_2 = \Phi_2\mathbf{w}_2$, are given by

$$\hat{\mathbf{x}}_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_1\,\lambda_2}\,\mathbf{x}$$

$$\hat{\mathbf{x}}_2 = \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_1\,\lambda_2}\,\mathbf{x}$$

Both $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$ are just scaled versions of $\mathbf{x}$.

$\Longrightarrow$ No separation at all!

MCA as a linear inverse problem

The objective function

$$J(\mathbf{w}_1, \mathbf{w}_2) = \|\mathbf{x} - \Phi_1 \mathbf{w}_1 - \Phi_2 \mathbf{w}_2\|_2^2 + \lambda_1 \|\mathbf{w}_1\|_1 + \lambda_2 \|\mathbf{w}_2\|_1$$

can be written as

$$J(\mathbf{w}) = \|\mathbf{x} - \mathbf{H}\mathbf{w}\|_2^2 + \|\boldsymbol{\lambda} \odot \mathbf{w}\|_1$$

where

$$\mathbf{H} = \begin{bmatrix} \Phi_1 & \Phi_2 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix}$$

and $\odot$ denotes point-by-point multiplication.

An ℓ_1-regularized linear inverse problem ...

- Non-differentiable
- Convex

$\Longrightarrow$ Use *Iterative Soft Thresholding Algorithm* (ISTA) or another algorithm to minimize $J(\mathbf{w})$.

Other algorithms include SparSA, TwIST, FISTA, SALSA, etc.

Split augmented Lagrangian shrinkage algorithm (SALSA)

SALSA is an algorithm for minimizing

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

SALSA is based on the minimization of

$$\min_{\mathbf{u}} \quad f_1(\mathbf{u}) + f_2(\mathbf{u}) \tag{12}$$

by the alternating split augmented Lagrangian algorithm:

$$\mathbf{u}^{(k+1)} = \arg\min_{\mathbf{u}} f_1(\mathbf{u}) + \mu\|\mathbf{u} - \mathbf{v}^{(k)} - \mathbf{d}^{(k)}\|_2^2 \tag{13}$$

$$\mathbf{v}^{(k+1)} = \arg\min_{\mathbf{v}} f_2(\mathbf{v}) + \mu\|\mathbf{u}^{(k+1)} - \mathbf{v} - \mathbf{d}^{(k)}\|_2^2 \tag{14}$$

$$\mathbf{d}^{(k+1)} = \mathbf{d}^{(k)} - \mathbf{u}^{(k+1)} + \mathbf{v}^{(k+1)} \tag{15}$$

Reference:

Afonso, Bioucas-Dias, Figueiredo.

Fast Image Recovery Using Variable Splitting and Constrained Optimization.

IEEE Trans. on Image Processing, 2010.

Applying SALSA to the MCA problem yields the iterative algorithm:

$$\text{initialize: } \mu > 0, \mathbf{d} \tag{16}$$

$$\mathbf{u}_i \leftarrow \mathrm{soft}(\mathbf{w}_i + \mathbf{d}_i, 0.5\boldsymbol{\lambda}_i/\mu) - \mathbf{d}_i, \qquad i = 1, 2 \tag{17}$$

$$\mathbf{c} \leftarrow \mathbf{x} - \Phi_1 \mathbf{u}_1 - \Phi_2 \mathbf{u}_2 \tag{18}$$

$$\mathbf{d}_i \leftarrow \frac{1}{\mu + 2} \Phi_i^t \mathbf{c} \qquad i = 1, 2 \tag{19}$$

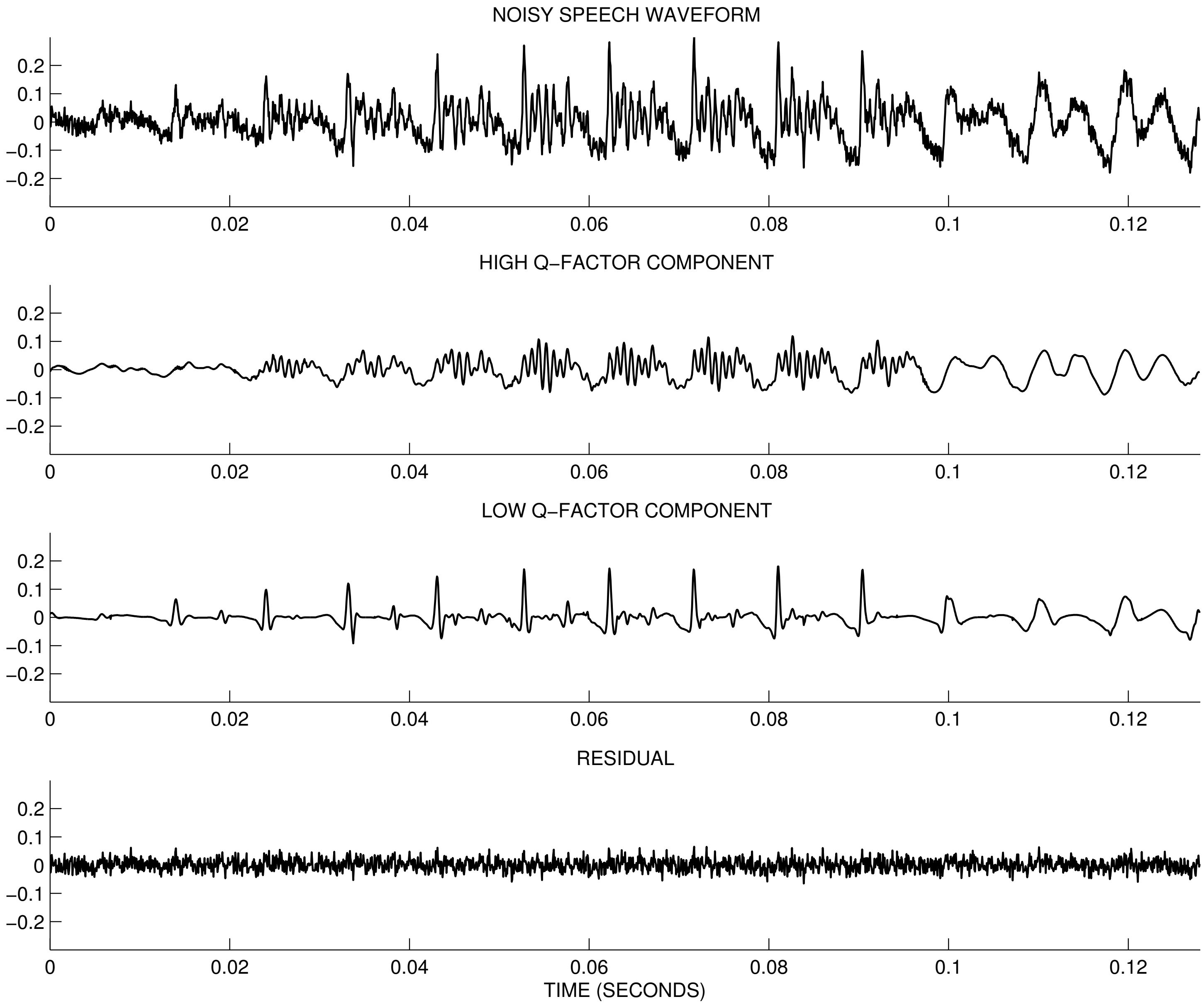
$$\mathbf{w}_i \leftarrow \mathbf{d}_i + \mathbf{u}_i, \qquad i = 1, 2 \tag{20}$$

$$\text{repeat} \tag{21}$$

where $\mathrm{soft}(x, T)$ is the soft-threshold rule with threshold T,

$$\mathrm{soft}(x, T) = x \max(0, 1 - T/|x|).$$

Note: no matrix inverses; only forward and inverse transforms.


NOISY SPEECH WAVEFORM
HIGH Q−FACTOR COMPONENT
LOW Q−FACTOR COMPONENT
RESIDUAL
0.2
0.1
0
−0.1
−0.2
0
0.02
0.04
0.06
0.08
0.1
0.12
TIME (SECONDS)