

Revision 1; date: 06.16.04

Elvir Causevic  
Robert E. Morley  
M. Victor Wickerhauser  
Arnaud E. Jacquin

---

ABSTRACT

---

FAST WAVELET ESTIMATION OF WEAK BIOSIGNALS

Wavelet based signal processing has become commonplace in the signal processing community over the past decade and wavelet based software tools and integrated circuits are now commercially available. One of the most important applications of wavelets is in removal of noise from signals, called *denoising*, accomplished by thresholding wavelet coefficients in order to separate signal from noise. Substantial work in this area was summarized by Donoho and colleagues at Stanford University, who developed a variety of algorithms for *conventional* denoising. However, conventional denoising fails for signals with low signal-to-noise ratio (SNR). Electrical signals acquired from the human body, called *biosignals*, commonly have below 0 dB SNR. Synchronous linear averaging of a large number of acquired data frames is universally used to increase the SNR of weak biosignals. A novel wavelet-based estimator is presented for fast estimation of such signals. The new estimation algorithm provides a faster rate of convergence to the underlying signal than linear averaging. The algorithm is implemented for processing of auditory brainstem response (ABR) and of auditory middle latency evoked potential response (AMLR) signals. Experimental results with both simulated data and human subjects demonstrate that the novel wavelet estimator achieves superior performance to that of linear averaging.

## List of Figures

**Figure 1:** Left: typical ABR waveform averaged over 8,192 frames and filtered by a BP Butterworth filter with linear phase (100-1,500Hz). Right: Typical single frame of ABR data, overlaid on the averaged waveform shown with the darker line. Note the peak-to-peak amplitude of the unaveraged single-frame signal.

**Figure 2:** Results of low pass filtering compared to conventional denoising applied to noisy sine waves.

**Figure 3:** Cyclic-shift tree denoising (CSTD) algorithm example for  $N = 8$  frames. The depth of the tree is  $K = 3$ . Each level contains exactly 8 frames, consisting of a different combination of frames from the level immediately above it.  $fd_{12}$  denotes the “denoising operation”  $fd_{12} = \text{den}(f_{12}, \delta_1)$ .

**Figure 4:** Threshold function selection plotted for 6 decreasing functions, a constant function, and one increasing function  $(\sqrt{2})^k$ .

**Figure 5:** Left: Performance of the novel wavelet algorithm compared to linear averaging 512 data frames (Top: linearly averaged 512 frames, middle: CSTD wavelet denoised 512 frames, bottom: original waveform (“true value”). Right: Variance comparison between linear averaging (top) and novel wavelet algorithm (bottom).

**Figure 6:** Performance of CSTD algorithm compared to linear averaging 256 data frames. (a): Template of AMLR evoked potential waveform from Spehlmann; (b): linear average of 8192 AMLR frames; (c): Single frame consisting of AMRL model plus WGN; (d): Linear average of 256 frames; (e): Result of CSTD algorithm.

**Figure 7:** Left: Comparison of Linear Averaging to CSTD for 512 frames of ABR data. (Top: linearly averaged 512 frames, middle: CSTD wavelet denoised 512 frames, bottom: linearly averaged 8,192 frames (“true value”). Right: Variance comparison between linear averaging and CSTD, as a function of number of frames.

**Figure 8:** Comparison between linear averaging and CSTD for AMLR. Left: consecutive averages of 256 frames of 110 ms each. Right: Results of CSTD algorithm applied to EEG frames covering the temporal range [7.5 ms, 110 ms] (Note: In the interval [0, 7.5 ms] the signal displayed is simply the linear average of 256. It is of little importance for AMLR interpretation). Color code for leads: Blue: “Fp1,” Green: “Fp2,” Red: “F7,” Cyan: “F8,” Magenta: “Fz.”

**Figure 9:** Color-coded compressed evoked potential array (CEPA) for averaging method. Each horizontal line in an image represents a reconstructed mid-latency evoked potential. Wave Na (See also Figure 6) is clearly identifiable by the vertical band of blue at latency around 15 ms. Wave Pa appears as a red band at latency around 25 ms. Waves Nb and Pb should respectively appear as vertical bands of cyan-blue and yellow-red but are not as easily identifiable.

**Figure 10:** Color-coded compressed evoked potential array (CEPA) for CSTD method. Each horizontal line in an image represents a reconstructed mid-latency evoked potential. Wave Na (See also Figure 6) is clearly identifiable by the vertical band of blue at latency around 15 ms. Wave Pa appears as a red band at latency around 25 ms. Waves Nb and Pb should respectively appear as vertical bands of cyan-blue and yellow-red but are not as easily identifiable.

## 1.0 INTRODUCTION

Accurate and rapid biosignal acquisition and estimation is critically important in a clinical environment. One example is universal neonatal hearing screening. As reported by the National Center for Hearing Assessment and Management, of the 4,000,000 infants born in the U.S. annually 12,000 have permanent hearing loss, corresponding to an average occurrence of three per 1,000 [19, 20, 21]. There are now 36 States in the U.S. that mandate that hearing screening be performed on all newborns. This presents a significant opportunity for increasing accuracy and decreasing test times for millions of tests conducted in the U.S. alone. The *auditory brainstem response* (ABR), generated by the brainstem in response to auditory stimuli, is widely used in hearing screening [14].

In this work, ABR signals will be analyzed using a novel wavelet-based noise suppression algorithm. The ABR is a good example of a weak biosignal since its amplitude is commonly under a microvolt, and it is commonly contaminated by noise with amplitude on the order of millivolts. It is a member of the evoked potential (EP) class of biosignals, which is present in an EEG recorded from electrodes on the human scalp in the presence of a periodic acoustic stimulus (click) presented in the ear canal via earprobe. This biosignal occurs during a time window of approximately 10 milliseconds after the onset of the stimulus. The first auditory related neural recordings were reported as early as 1929 by Berger, and 1930 by Weaver and Bray [13]. The key results derived by studying ABRs are directly applicable to many weak, repetitive biosignals such as the acoustic middle-latency evoked potential response (AMLR) or other evoked potentials. Wavelet-based techniques have already been successfully used for the denoising of visual and auditory late response (ALR) evoked potentials [24], which have significantly higher signal-to-noise ratio (SNR) than either ABR or AMRL signals.

In this paper, we define an *EEG frame* (or simply *frame*) to be a set of electrical potentials recorded at sampling frequency  $f_s$  during the time period spanned between two consecutive stimuli. The

recorded signal  $E[n]$  is usually modeled as a time series which is the sum of the deterministic evoked potential signal  $S[n]$  and noise  $R[n]$ :

$$E[n] = S[n] + R[n], \quad (1.1)$$

where  $R$  represents the biological noise contributed by the EEG and instrument noise. This noise is usually modeled as a zero-mean random process which averages-out to zero when sufficiently long averages are used. A typical ABR waveform collected from a subject in this work, averaged over 8,192 frames, and smoothed with a Butterworth BP filter (100 Hz – 1,500 Hz) is shown on the left side of Figure 1. The right side of Figure 1 shows the same averaged ABR waveform — the dark line close to the center of the graph, but overlaid by a single frame of the unaveraged ABR recording.

Conventional wavelet denoising (WD) has been demonstrated to be an effective algorithm across a wide range of signal processing problems [8, 9, 24] when the signal-to-noise ratio (SNR) of the signal being denoised is relatively high. However, an application of conventional denoising to signals with small SNR fails, as will be demonstrated in Section 3.4. The goal of this work is to be able to obtain a fast estimate of the underlying weak biosignal corrupted by noise, which significantly outperforms the usual method of linear averaging synchronized to the stimulus.

The organization of the paper is the following. Section 2 describes the signal acquisition procedure. Section 3 reviews classical results of estimation theory and wavelets and motivates the rest of the paper. Section 4 is devoted to an algorithmic presentation of the novel denoising method proposed in this paper. Section 5 contains experimental results using both simulated signals and real signals from human subjects.

## **2.0 SIGNAL ACQUISITION AND PROCESSING SYSTEM**

### **2.1.1 Experiment Design — ABR**

Condensation acoustic click stimuli (rectangular pulses) were presented to the ear with a probe. The clicks had power of 65 dB peak-equivalent SPL (dB SPL) and 100  $\mu$ s duration. The repetition rate

was set to 37 per second, or one click every 27.03 ms. This rate was selected to match the normative data given in [14, 22]. The click stimuli were calibrated using a standard procedure and IEC standard [11, 6, 10].

Three pre-gelled surface electrodes were attached to each subject's head, one each on the left and right mastoid ( $M_1$  and  $M_2$ ), the bone behind the ear, and one applied to the vertex ( $C_z$ ), at the highest point on the head. The electrodes were located per standard EEG and ABR configuration [19, 6]. The electric potentials were recorded differentially between the vertex and the ipsilateral mastoid process (behind the ear containing the probe). The contralateral mastoid electrode (behind the other ear) was used as a reference. Inverted common mode voltage between the vertex and the ipsilateral electrode was fed into the reference electrode to reduce 60 Hz power line interference [2, 12, 19].

The data acquisition system collected 1,024 samples per frame, corresponding to approximately 21.3 ms of a recording following a stimulus. A total of 8,192 frames of data were collected for each testing condition, in response to 8,192 click stimuli. The validity of each ABR recording was verified using the standard  $F_{sp}$  method [12, 7], which is proportional to SNR.

Valid ABR recordings from a total of ten ears from six volunteer subjects were used. The subjects were adults, five males (27, 33, 48, 54, and 57 years old), one female (30 years old). All of the subjects participated on a voluntary basis with signed consent forms under approval from an Institutional Review Board at Everest Biomedical Instruments Company under the direction of a certified audiologist. Hearing sensitivity in each of the ears was normal based on standard audiometry (thresholds better than 20 dB HL), otoacoustic emissions screening, or independent clinical ABR evaluations. A total of 8,192 frames of 10.67 ms of post-stimulus data were analyzed for each recording. Data were purposefully collected in a very noisy office environment, since the goal of our work was to perform rapid signal estimation in such a non-ideal environment. This is in sharp contrast with most other published data, for which experiments were run in a sound-attenuating booth.

### 2.1.2 Experiment Design — AMLR

Since AMRL responses are not the main focus of this work, we only used one test subject (namely one of the authors). SynAmps<sup>2</sup>™ amplifiers from Neuroscan were used for data acquisition, with a sampling frequency of 5 kHz. EEG data from the five electrodes Fp1, Fp2, F7, F8 and Fz, all placed on the forehead was collected with a reference electrode placed on the left mastoid (ear lobe). The acoustic stimuli consisted of 1 kHz tone bursts with a duration of 10 ms and modulated by a cosine envelope. The stimulus repetition rate was 9.2 Hz and stimulus level was 85 dB SPL. While the EEG recording lasted for five minutes, the stimulus was played for the first four minutes only, so that we could verify that the mid-latency Evoked Potential returns to baseline in the absence of stimuli. The subject was relaxed and awake during the entire recording. In addition, preprocessing of the EEG was applied in the form of downsampling to 2.5 kHz followed by bandpass filtering with a 170-tap FIR filter with passband 18-250 Hz.

## 3.0 THEORETICAL MODEL DEVELOPMENT

### 3.1 ABR Signal Modeling and Linear Estimation Theory

We make three key assumptions in the model described by Equation (1.1): 1) that the true value of the ABR signal  $S$  at a fixed time delay after the onset of the stimulus remains constant from frame to frame, 2) that this ABR signal is “smooth,” and 3) that the additive noise  $R$  is white, and gaussian (AWGN). All three assumptions have been empirically verified for the last 30 years of research since the first ABR recordings [17, 6, 16, 12, 15].

A possible estimator of the signal  $S$  corrupted by AWGN noise is the linear average of  $N$  observed frames  $E_i[n] = S[n] + R_i[n]$ :

$$\hat{A} = \frac{1}{N} \sum_{i=0}^{N-1} E_i[n] = S + \frac{1}{N} \sum_{i=0}^{N-1} N_i[n],$$

(3.1)

where  $i$  ( $0 \leq i \leq N-1$ ) indicates frame number and  $n$  is the time sample index. The noise variance for this sample mean estimator is calculated as follows:

$$\text{var}\{\hat{A} - S\} = \frac{1}{N^2} \text{var}\left\{\sum_{i=0}^{N-1} R_i\right\} = \frac{1}{N^2} N\sigma^2 = \frac{\sigma^2}{N},$$

(3.2)

where  $\sigma^2$  is the variance of the random noise process and  $N$  is the number of frames averaged. Moreover, it can be shown that this estimator is the minimum variance, unbiased estimator (MVUE) since it achieves the Cramer-Rao lowerbound (CRLB)(See [3] for details). Thus, linearly averaging  $N$  frames of ABR data reduces the variance of the noise by a factor of  $1/N$ , and improves the SNR by:

$$10 \log_{10}(N) \text{ dB}. \quad (3.3)$$

Experimental results presented in Sections 5.1.2 and 5.1.4 verify this theoretically expected result. As will be demonstrated in Sections 4 and 5, the novel wavelet denoising algorithm proposed here improves both the Cramer-Rao lower bound for the variance and the MSE as compared to the sample linear estimator.

### 3.2 Wavelet transform for the ABR Signal

The reader is referred to [18, 24, 29] for an introduction on wavelets. The classical, critically sampled, discrete wavelet transform (DWT) implemented in this work can be viewed as processing the original signal using a bank of “constant-Q” filters applied successively. The first application of the filters decomposes the signal into high-pass (HP) and low-pass (LP) components. These components are then decimated by a factor of 2, and the low-pass component is then further decomposed into another set

of highpass and lowpass components, and so on, until the signal is fully decomposed. At each of the decomposition levels, the coefficients can be either set to zero or reduced in magnitude, such that a particular feature of the signal is affected upon reconstruction. If the LP and HP filters (H and G respectively) used for decomposition are quadrature mirror filters (QMFs), and if their biorthogonal complements are used for reconstruction with proper treatment of endpoints, then a perfect reconstruction in phase and amplitude can be achieved.

The biorthogonal 9-7 wavelet was chosen because it can be implemented using simple and short FIR filters, and still allow perfect reconstruction. In addition, this wavelet is symmetric and has 4 vanishing moments, which makes it suitable for distinguishing between the expected regularity of the smooth ABR signal vs. the absence of regularity for the rough AWGN noise. The biorthogonal 9-7 wavelet is also the wavelet used for EEG wavelet signal processing in research literature [15], the JPEG 2000 image compression standard, and is also used by the FBI for the national fingerprint storage database [29]. Implementation details for the DWT transform using this wavelet can be found in [18].

### **3.3 Effect of SNR on the Performance of Conventional Denoising**

We now demonstrate, using simulated signal data consisting of eight cycles of a sine wave with peak amplitude of one, that a critical condition necessary for conventional denoising to work is that the SNR of the measured signal is fairly high (approximately 20 dB), thereby establishing that conventional wavelet denoising is not applicable to the estimation of weak biosignals. This confirms results reported by Coifman and Wickerhauser in [4].

By performing a visual comparison using Figure 2, it appears that for noisy sine waves with SNR of  $-20$  dB and  $-10$  dB, the lowpass filtering approximately estimated the shape of the sine wave waveform. We also see that for these signals, conventional denoising produced constant level signals of very small amplitude, with bursts of high amplitude noise spikes. This is because too many wavelet coefficients were set to zero, removing the noise, but also removing the salient features of the signal. For



the 0 dB and +10 dB noisy sine waves, the LP filtering produced a smooth estimate of the final waveform, with no high frequency noise riding on it, while the conventional denoising still demonstrated relatively large spikes of high frequency noise. It is only when the SNR of the noisy sine wave is +20 dB, that we see the performance of conventional denoising approaching that of a low pass filter. The following section introduces an algorithm which overcomes these limitations.

#### 4.0 FAST WAVELET ESTIMATION — CYCLIC SHIFT TREE DENOISING (CSTD)

This section describes a novel wavelet denoising algorithm which utilizes information contained in all of the individual ABR data frames, as opposed to using only the single linear average of the acquired data frames. As demonstrated in the above Section, the conventional wavelet-based denoising algorithm applied to an ABR signal with “low SNR” (i.e. less than 0 dB) “denoises too much,” and significantly distorts the morphology of the underlying signal of interest. Relevant characteristics of the existing conventional denoising algorithm are as follows:

1. Wavelet coefficients thresholded to zero at each wavelet decomposition level
2. SNR of the original signal must be large prior to applying denoising ( $> +10$  dB)
3. All denoising operations are performed on a single data vector

The first characteristic is in common with the novel algorithm proposed. The second characteristic is a limitation that is overcome by the novel algorithm, whose performance is demonstrated for signals with SNR less than zero. The third characteristic of conventional wavelet denoising is expanded here from a single data vector to all the available data, or in the case of ABR, from denoising a single, final linearly averaged frame, to step-by-step denoising of all of the available frames and their recombinations in a tree-like fashion.

We recall that the goal of this work is to reduce the variance of the ABR signal estimator as a function of the number of frames, beyond that of the optimum linear estimator. We propose to achieve this by creating a large number of recombinations, denoted by  $M$ , of the original  $N$  frames. If these new

$M > N$  frames are statistically independent and the estimator is unbiased, this will decrease the variance of the estimator.

We choose to recombine the frames by using two *adjacent* frames and calculating their linear average. This method is chosen for its simplicity, computational stability, and well-understood behavior. Each *dyadic linear average* is then *denoised*, which creates a new frame. This recombination process is a tree-like process, in which new levels of recombined frames are created. The average and denoise operation creates frames at level  $k$ , which are no longer a linear combination of frames from level  $k-1$ .

#### 4.1 A Frame Recombination Method: Cyclic-Shift Tree Denoising

The proposed algorithm produces an array of new frames by dyadically averaging and denoising. We refer to this method as *cyclic-shift tree denoising* (CSTD). It was motivated by [1, 5]. The CSTD algorithm creates a tree (or array) of frames of width  $N$  and depth:

$$K = \log_2 N. \quad (4.1)$$

At each level  $k$ , ( $1 \leq k \leq K$ ), along the depth of the tree, two adjacent frames at level  $k-1$  (dyads) are averaged and denoised, and a new level is created. We extend this new level to include not only dyadic averages of adjacent frames at level  $k-1$ , but also dyadic averages of *acyclical shift* by one frame at level  $k-1$ . This procedure is illustrated in Figure 3 for  $N = 8$  frames.

Two items are important to note about the CSTD algorithm. First, without denoising, frame recombinations using cyclic-shift tree averaging yield simple linear combinations of frames at each new level. A linear average of frames at any particular level is identical to the linear average of frames at all other levels. The second item to note is that at the bottom of the CSTD tree, each frame is identical to every other frame. This is because the cyclic shift algorithm, without denoising, assures that each frame at the bottom is the linear average of **all** the frames at the original level, and that they are included in that average exactly once. An examination of the bottom level of Figure 3 confirms this fact for  $N = 8$ .

Thus, the bottom  $N$  frames are each an identical linear average of all of the top  $N$  frames, but taken through a different path.

With denoising however, data in each frame is recombined, in a nonlinear fashion, through a different path. In other words, a recombined and denoised frame is not a linear combination of frames at the previous level, because of the wavelet coefficient thresholding. This is a powerful feature of the cyclic-shift tree algorithm as we will demonstrate in Section 5.

## 4.2 Intermediate Denoising With Variable Thresholds

Denoising consists of thresholding wavelet coefficients [8, 9, 10, 26], where two different types of thresholds are used. First, thresholding is applied within a single wavelet-transformed frame (DWT). At each level (or scale) of the wavelet decomposition, a different threshold is applied which affects coefficients at different scales differently. Wavelet decomposition scales corresponding to higher frequencies are thresholded with a larger threshold, and as the scale increases with additional levels of decomposition, and the features of the signal are more prominent, fewer coefficients are set to zero. This thresholding is implemented in the novel algorithm as suggested in standard wavelet denoising literature, such that the threshold level drops from scale to scale by a factor of  $2^{i/2}$ , where  $i$  is the wavelet decomposition level. The reader is referred to the work of Donoho [8-10] for details on conventional wavelet denoising.

The second type of variable thresholding is applied across CSTD levels as depicted in Figure 3, and is unique to our method. Wavelet coefficients are thresholded for each frame at that level with a different initial threshold, denoted  $\delta_l$ . The threshold function is really a function of two variables:  $\delta_{k,w}$ , where  $k$  is the index corresponding to the CSTD algorithm level, and  $w$  is the index corresponding to the particular wavelet scale within a single frame, at level  $k$  of CSTD. We will only be examining the variation of  $\delta$  as a function of the first index (CSTD level  $k$ ) and use the notation  $\delta_k$  for simplicity. When we discuss thresholding, we exclusively refer to CSTD level-dependent thresholding applied to our novel algorithm. Both types of thresholding play an important role in the effectiveness of the algorithm.

We have many choices when selecting the underlying function that relates  $\delta$  to the CSTD level. We decided to select the function  $\delta_k$  from a range of commonly used signal processing functions in search of one that minimizes the variance of the estimator. Intuitively, we wanted to select a monotonic function, either a strictly-increasing, strictly-decreasing, or a constant function, because the behavior of the CSTD is consistent from level to level. In addition to the constant function  $\delta_k = \Delta$ , we evaluated the six functions and their inverses listed in Equation 4.2 below, where  $k$  denotes the level in the tree. Each of these functions describes the relationship between  $\delta$ 's at intermediate levels. Threshold  $\delta_1$ , applied at the first level (initial delta),  $\delta_2$  is calculated using one of the forms of Equation 4.2 with respect to  $\delta_1$ ,  $\delta_3$  is also calculated with respect to  $\delta_1$ , and so on through  $\delta_6$ .

$$\delta_k = \begin{cases} k \text{ and } \frac{1}{k} \\ k^2 \text{ and } \frac{1}{k^2}, \\ \log(k) \text{ and } \frac{1}{\log(k)}, \\ e^k \text{ and } \frac{1}{e^k}, \\ (\sqrt{2})^k \text{ and } \frac{1}{(\sqrt{2})^k}, \\ \cos\left(\frac{\pi k}{4K}\right) \text{ and } \frac{1}{\cos\left(\frac{\pi k}{4K}\right)}, \quad 0 \leq \frac{\pi k}{4K} \leq \frac{\pi}{4}. \end{cases} \quad (4.2)$$

Starting with a large delta and decreasing it at each level yields substantially lower variance and RMS error. Furthermore, this strategy makes sense considering that frames that are deep in the tree consist of averages of more frames than those frames higher in the tree and are therefore expected to have a higher SNR, thereby requiring a lower denoising threshold.

Figure 4, obtained for a typical human subject's ABR recording, demonstrates that the decreasing function  $2^{-k/2}$  yields the best overall results in terms of rms error. It also demonstrates that there exists a minimum for each particular threshold function, which depends on the choice of the initial threshold.

This means that the choice of initial threshold below which to set the coefficients to zero is very important, because a very large threshold keeps too many noise-related coefficients, and a very small threshold eliminates coefficients related to the information-bearing signal.

### 4.3 CSTD Performance Evaluation

The overall algorithm, which combines the ABR signal acquisition description from Section 2, and the novel CSTD algorithm developed in this Section is as follows:

1. Acquire ABR data from human subjects over 8,192 frames, using the data acquisition system described in Section 2. (Details on human subjects are provided in Section 2).
2. Create an array of wavelet coefficients by performing a DWT on each frame of the 8,192 frames, arranged as successive original frames  $[1, 2, \dots, K]$ .
3. Create the “final average” by linearly averaging all 8,192 frames, and filtering to obtain a smooth signal to be used as the “true” ABR signal.
4. Denoise the array of wavelet coefficients using CSTD, to obtain a total of  $M = N \cdot \log_2(N)$  new frames, each denoised differently. (i.e., at most  $8,192 \cdot 13 = 106,496$  frames).
5. Linearly average all different denoised reorderings of frames to which CSTD has been applied to obtain a sequence of  $N$  frames.
6. Linearly average the  $N$  frames to obtain one frame of wavelet coefficients
7. Perform the IDWT on this averaged frame to obtain time domain samples.
8. Calculate the variance and the RMS error between the linear average and final average and compare to the variance and RMS error between denoised average and final average, for an increasing number of frames.

This algorithm was iterated for various functions  $\delta_n$ , and various starting values to obtain the minimum variance and error. The algorithm was applied to simulated data with known SNRs and to real ABR data collected from human subjects. The results will be reviewed in Section 5

## Rate of Convergence of the Estimator

Next we consider the rate of convergence of the novel wavelet denoising estimator, defined as the reduction of noise power as a function of the number of acquired ABR frames. We recall from Section 3.2 that linear averaging reduces the variance of the noise by a factor of  $c/N$ . The goal of the novel CSTD algorithm is to reduce the variance by  $c/M$ , where  $M > N$ .

We start with the assumption that the new  $M$  frames are statistically independent to establish a theoretical best case, even though this assumption is clearly not valid. If this assumption held, the SNR improvement (in dB) obtained by linearly averaging these  $M$  frames would be, according to Eq. 3-4:

$$10 \log_{10} (M) = 10 \log_{10} (N) + 10 \log_{10} (K), \quad (4.3)$$

which represents an additional improvement of  $10 \log_{10} (K)$  dB over the improvement brought about by simply averaging the  $N$  original frames. As will be presented in Section 5.2.2, the novel wavelet denoising algorithm does achieve excellent performance when  $N * \log_2(N)$  new (denoised) frames are created out of the original  $N$  frames. While the assumption that the  $M$  new frames are statistically independent will be shown to be incorrect, we will also see that the novel wavelet denoising comes fairly close to the theoretical limit. We next examine the performance of this estimator with simulated and human subject data.

## 5.0 EXPERIMENTAL RESULTS

### 5.1 Simulated Data Experiments

#### 5.1.1 Experiment Design— ABR

Simulations were carried out using a Matlab-generated sine wave to which white gaussian noise was added sample by sample. We followed a method by Coifman and Wickerhauser [4], which they presented for denoising medical signals and images. To approximately match the amplitude and frequency characteristics to that of an ABR signal, a sine wave was created with amplitude one and with eight full cycles per 10.66 ms frame (750 Hz). A random noise vector was created with a Gaussian

distribution, with a density of  $N(0,1)$ , referred to as a WGN vector. This vector was generated over 8,192 frames, so that a different noise sample was added to each sine wave frame. Next, the noisy sine wave signals were generated for simulation by adding the scaled WGN vector to the sine wave vector in order to match any given SNR. The starting phase of the noise-free sine wave was the same for each simulated EEG frame (i.e. no phase jitter was introduced).

### 5.1.2 Results of CSTD Estimator with Simulated ABR Data

The novel denoising algorithm was applied to the three noisy sine waves described in the previous section. The results for the  $-20$  dB noisy signal represented the worst case, and it is the only case discussed here in detail. The left side of Figure 5 demonstrates the performance of the algorithm when applied to the noisy sine wave with an SNR of  $-20$  dB, after only 512 frames have been averaged. The waveforms are shown in the following order: the top waveform is the linear average of 512 frames, the middle waveform is the novel wavelet estimate of the signal, and the bottom waveform is the actual signal being estimated. The right side of Figure 5 shows the variance of the linear estimator (top) vs. the variance of the novel wavelet estimator (bottom), as a function of number of frames averaged.

The numerical results corresponding to Figures 5 are in Table 5-1. The linear averaging and wavelet denoised variances were calculated and compared, as were the SNRs for the two approaches.

Table 5-5.1 Comparison of linear averaging to novel wavelet denoising for 512 data frames

Frames	Linear variance	Denoisd variance	Variance ratio	Linear SNR (dB)	Denoise SNR (dB)	SNR Improvement
2	23.6363	2.0081	11.77	-16.74	-6.03	10.71
4	12.6286	0.9706	13.01	-14.02	-2.87	11.14
8	5.9670	0.6821	8.75	-10.76	-1.34	9.42
16	2.9670	0.3341	8.88	-7.72	1.76	9.48
32	1.6339	0.1610	10.15	-5.13	4.93	10.06
64	0.8000	0.0827	9.67	-2.03	7.82	9.86
128	0.4052	0.0479	8.46	0.92	10.20	9.28
256	0.1902	0.0271	7.02	4.21	12.68	8.47
512	0.1023	0.0184	5.56	6.90	14.35	7.45

The variance for the denoising algorithm was reduced by a factor of 9.25 on average, as compared to that of linear averaging. The theoretically expected SNR improvement for linearly averaging 512 frames is

27.1 dB. For a  $-20$  dB signal, the resulting SNR after averaging together 512 frames is therefore expected to be 7.1 dB. We see from the table that linear averaging produced a signal with SNR of 6.9 dB, similar to the expected value, while the novel wavelet algorithm more than doubled that, and produced an estimate with a SNR of 14.35 dB. Note that this is somewhat lower than the theoretically expected value of 16.64 dB (computed from Equation 4.3), which we would expect if the assumption of statistical independence of the  $M$  frames was valid. When the algorithm was run over a larger number of frames, i.e. 4,000 or more, the performance of the linear averaging approached that of the wavelet denoising, in terms of error variance.

Thus, when we compare the linear averaging process applied by itself, to the CSTD process, which also uses linear averaging as an integral part, we see that the amount of improvement is generally reduced with an increasing number of frames. However, with ABR signal processing, we are concerned with SNR improvement for a small number of frames (i.e., 512 vs. 8,192), hence the CSTD algorithm does offer a significant improvement when compared to conventional linear averaging.

### 5.1.3 Experiment Design—AMLR

We modeled a typical acoustic mid-latency evoked response (AMLR) from the standard primer of Spehlman *et al.* [27], as shown in Figure 6(a). The AMLR typically occurs in the first 80 ms following the onset of the stimulus. Mid-latency evoked potentials are of particular interest with regard to measuring depth of anesthesia for which being able to unambiguously measure the latencies and amplitudes of components Na, Pa and Nb is desirable [25]. We stored this ideal response as a vector of values sampled at 4 kHz, corresponding to a frame size of 320 samples. As in the ABR experiment, we generated a large vector of 8,192 frames and several random noise vectors to obtain different signal-to-noise ratios.

### 5.1.4 Results of CSTD Estimator with Simulated AMLR Data

Figure 6(b-e) illustrates the performance of the CSTD algorithm on noisy AMRL signals. We first note that the average of all 8,192 frames of AMRL is very close to the model, as we would expect



for such a long average. We also note from this figure that the CSTD algorithm results in reconstructed wave shapes with significantly better fidelity to the noise-free AMRL template than those obtained by simply averaging 256 frames.

SNR improvements of the CSTD result versus linear averaging are listed in Table 5-2 for averages of 256 and 512 frames. Again, we can easily verify that linear averaging improves the SNR by the predictable amount given by Equation 3.6 and that the SNR improvements which result from CSTD come close to the theoretical maximum value of Equation 4.3.

Table 5-2 Comparison of linear averaging to CSTD denoising for 256 and 512 AMLR 512 data frames

Frames	SNR noisy signal	Linear SNR (dB)	Denoise SNR (dB)	SNR dB Improvement
256	-15	9.2	14.5	9.3
256	-20	4.4	12.9	8.5
256	-25	-1.4	9.9	11.3
512	-15	11.8	15	3.2
512	-20	7.7	15.0	7.3
512	-25	2.2	11	8.8

## 5.2 Human Subjects Experiments

### 5.2.1 CSTD Estimator Experiment Results with Human Subject Data— ABR

Figure 7 presents the results of applying, for one subject, the novel wavelet denoising to the ABR waveform presented in the introduction. A total of 512 frames were processed, taken from the first quarter of the available 8,192 frames. The bottom waveform in the plot on the left side of Figure 7 has been averaged over 8,192 frames and smoothed by a bandpass filter (100 Hz -1,500 Hz). Still in this same plot, we can see in the middle waveform showing the denoised signal estimate, the peak V in the same location as in the bottom waveform taken to be the true ABR signal (i.e., approximately 7.5 ms). We observe that the denoised waveform is smooth, and conclude that locating peaks on it by a human expert, or perhaps even having an automated peak detection algorithm perform this task, would be substantially more accurate than attempting the same on the linear average of the same 512 frames.

The right side of Figure 7 shows the change in variance (for one subject) for linear averaging (top) vs. novel wavelet denoising (bottom), as a function of the number of frames averaged. Variance was computed using mean-squared-error (MSE), where linearly averaged 8,192 frames constituted the “true signal.” We can see that the variance of this wavelet denoising algorithm is reduced substantially relative to linear averaging. Table 5-3 summarizes the results for these waveforms.

Table 5-3 Denoised signal root-mean-square error (rms; i.e. square root of the mse) compared to linear averaging variance for up to 512 frames.

Frames	$\sigma_{\text{denoised}}$	$\sigma_{\text{averaged}}$	Ratio
2	1.85492	3.60118	1.94
4	0.91770	1.39518	1.52
8	0.92050	1.06191	1.15
16	0.53540	0.66465	1.24
32	0.19162	0.33016	1.72
64	0.02722	0.08958	3.29
128	0.01584	0.04042	2.55
256	0.01341	0.02653	1.98
512	0.00490	0.01264	2.58

### 5.2.2 CSTD Estimator Experiment Results with Human Subject Data—AMLR

Figure 8 shows eight successive reconstructions of the AMRL using simple averaging of 256 frames and the CSTD algorithm. Note that in the case of CSTD processing, each data frame consisted of 256 samples spanning the interval 7.5-110 ms (This interval was chosen since the first 7.5 ms are of little use for interpretation of the AMRL response). Grand averages of 1818 frames are also displayed in the lower part of the figure. Figures 9 and 10 show the waveforms in the form of *compressed evoked potential arrays* (CEPAs). Each image in Figure 9 (and in Figure 10) corresponds to a particular electrode. Each horizontal line of any particular image in the CEPA represents a color-coded waveform. The temporal overlap between two consecutive lines is of 200 frames (78% overlap). This graphical display allows us to visualize the stability of the reconstructed Evoked Potential signal in time and to visualize the amount of noise left in the denoised responses. Troughs are represented by colors going

from cyan to dark blue as their amplitude increases. Peaks are represented by colors going from yellow to dark red.

Figures 8-10 show that the CSTD algorithm significantly improves the reconstructed AMRL, especially if we focus our attention to the Na-Pa-Nb morphology of the signal. Looking at bands after time 00:04:00 (which indicates the end of the first four minutes) in Figures 9 and 10, we can also visualize the amount of residual EEG and instrument noise which is left in the responses.

## 6.0 CONCLUSION AND FURTHER WORK

This work presents a novel fast wavelet estimator, for application to weak biosignals such as ABRs and AMRLs. Linear averaging of many signal frames is universally used to increase the SNR of Evoked Potential Responses. Our goal was to develop a wavelet based estimator that would increase the SNR faster than a linear averaging estimator. The concept of wavelet denoising was utilized to create the new algorithm. Conventional wavelet denoising is the process of setting wavelet coefficients that represent the noise to zero, while keeping the coefficients that represent the signal, using a threshold function. We established that conventional wavelet denoising fails for signals with SNR below zero, because too many wavelet coefficients are reduced to zero.

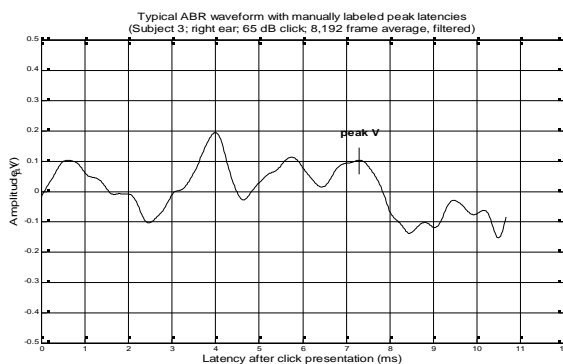
We developed methods to recombine a set of  $N$  original data frames to create a large number of new frames, each of which was denoised using a variable threshold function. We presented an algorithm, called cyclic-shift tree denoising (CSTD), for the creation of new frame combinations. This algorithm used the original  $N$  frames of data and produced  $N \cdot \log_2(N)$  new frames. The new frames were derived by averaging pairs of adjacent original frames, followed by denoising each average. New levels of frames were thereby created, and at each level, a different denoising threshold was employed. Since denoising is a non-linear operation, the CSTD algorithm produced new frames that were not simply a linear combination of the original frames.

The novel algorithm has several limitations. Like linear averaging, it cannot be applied to a single frame of data, hence multiple measurements of the same signal must be made. It also requires a number of original frames equal to a power of two. Also, the signal being estimated must be fairly constant across frames as well as smooth, when compared to the noise that corrupts it, because this is the assumption underlying wavelet denoising. Finally, the algorithm in its current implementation requires that all the frames of data be collected and stored prior to the application of the algorithm (i.e., for processing 512 frames, all of the 512 frames must be available in memory).

Further work should address the influence of the choice of wavelet transform on the CSTD algorithm. Also, a formal comparison with more traditional methods of Evoked Potential extraction could be performed.

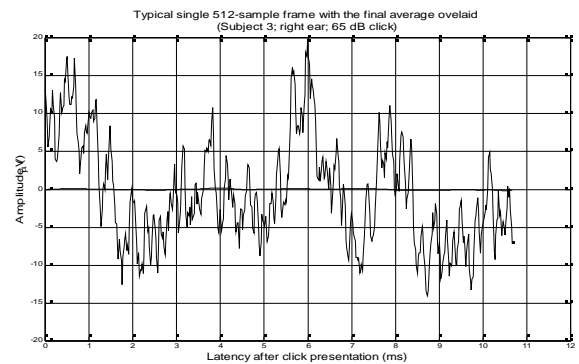
## Figures

$\mu\text{V}$



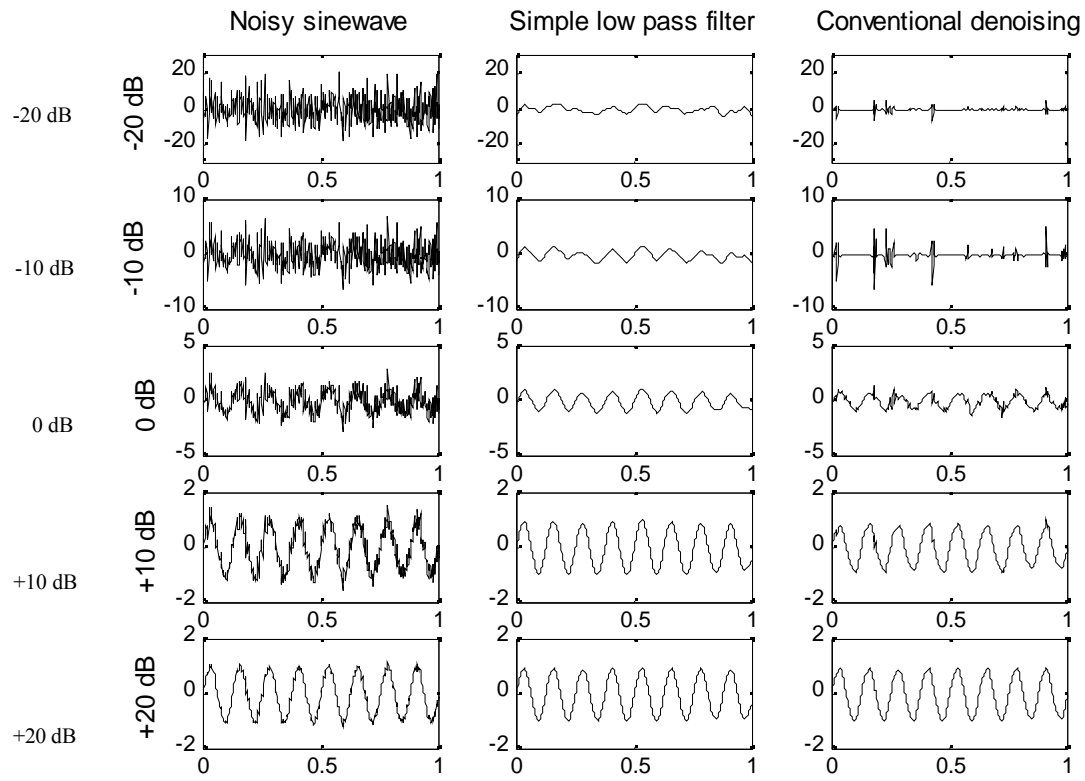
Latency after click presentation (ms)

$\mu\text{V}$

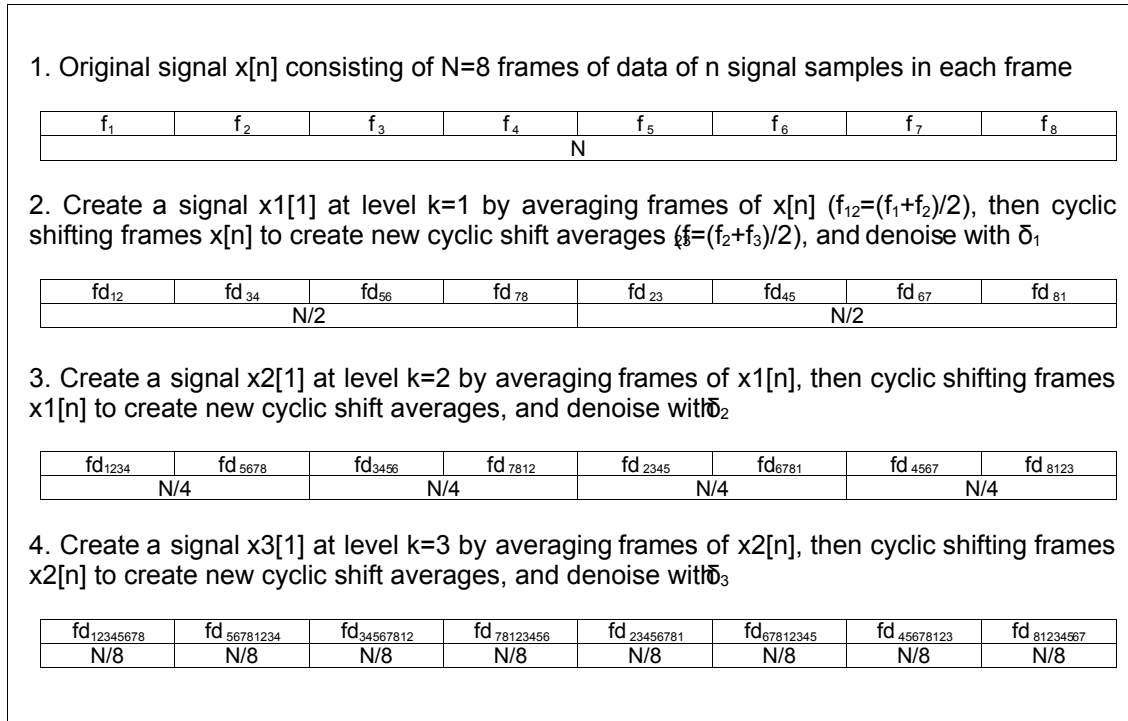


Latency after click presentation (ms)

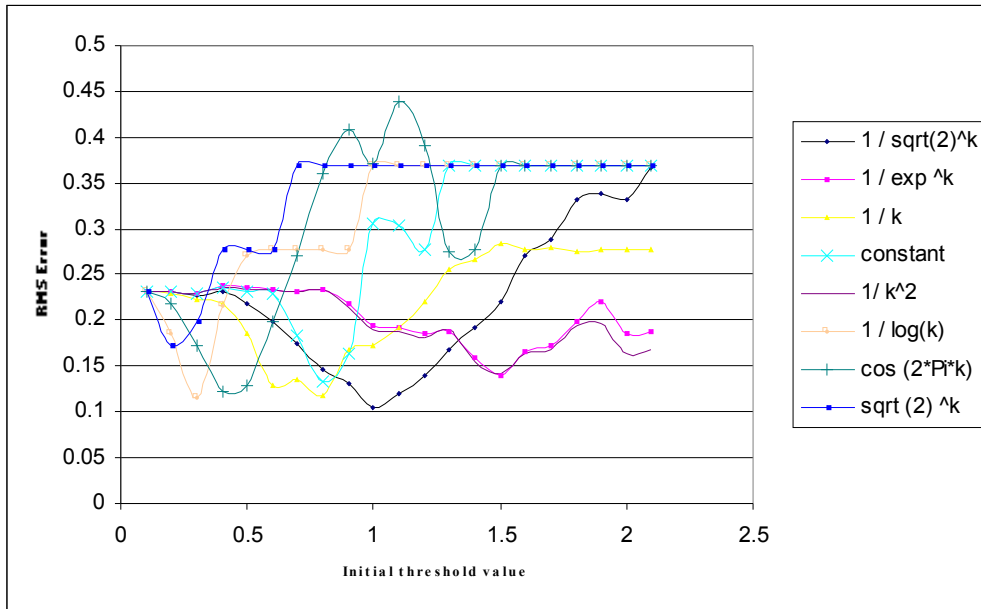
**Figure 1:** Left: typical ABR waveform averaged over 8,192 frames and filtered by a BP Butterworth filter with linear phase (100-1,500Hz). Right: Typical single frame of ABR data, overlaid on the averaged waveform shown with the darker line. Note the peak-to-peak amplitude of the unaveraged single-frame signal.



**Figure 2:** Results of low pass filtering compared to conventional denoising applied to noisy sine waves.

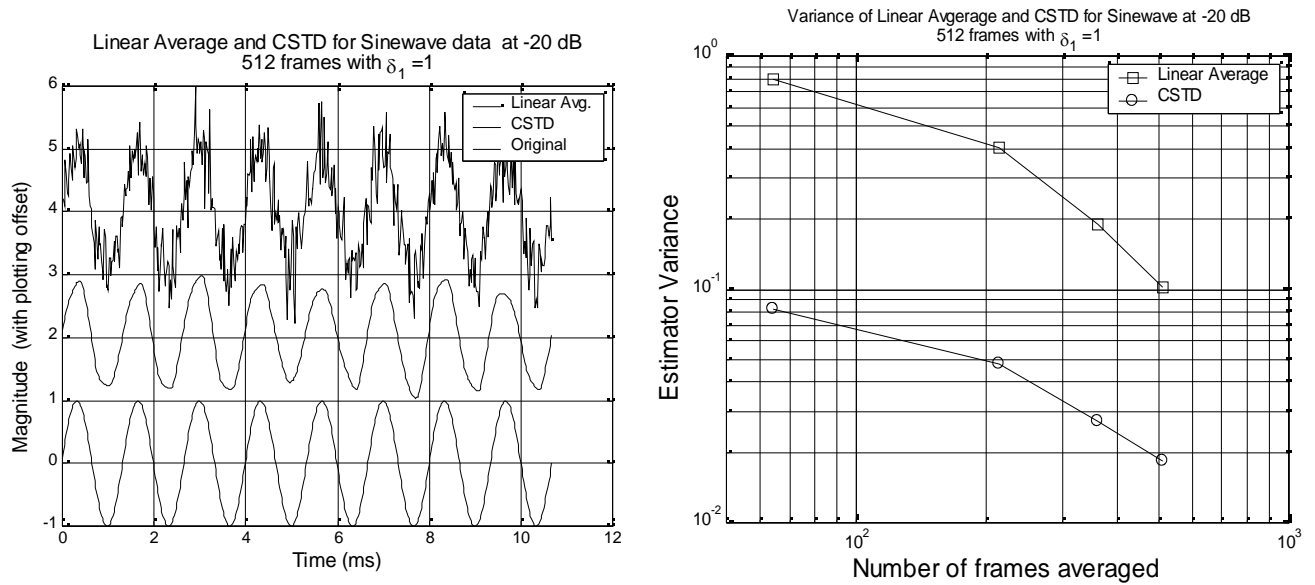


**Figure 3:** Cyclic-shift tree denoising (CSTD) algorithm example for  $N = 8$  frames. The depth of the tree is  $K = 3$ . Each level contains exactly 8 frames, consisting of a different combination of frames from the level immediately above it.  $fd_{12}$  denotes the “denoising operation”  $fd_2 = \text{den}(f_{12}, \delta_1)$ .

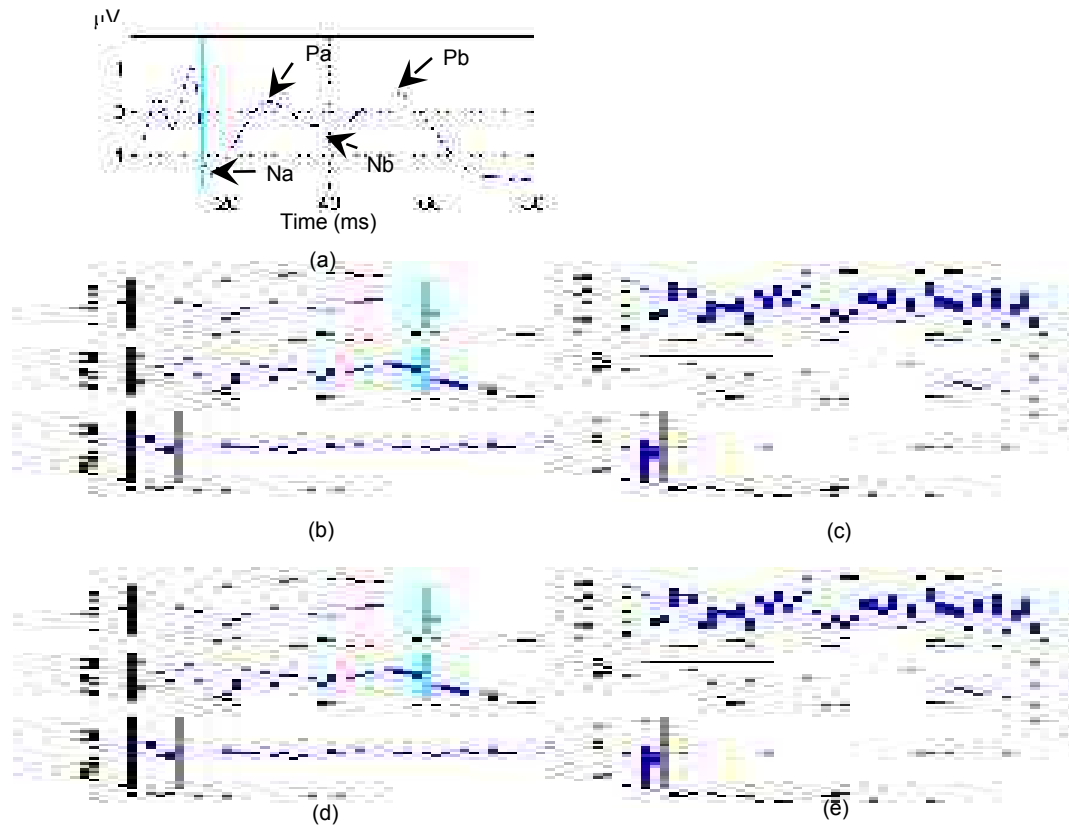


**Figure 4:** Threshold function selection plotted for 6 decreasing functions, a constant function, and one increasing function ( $\sqrt{2}^k$ ).

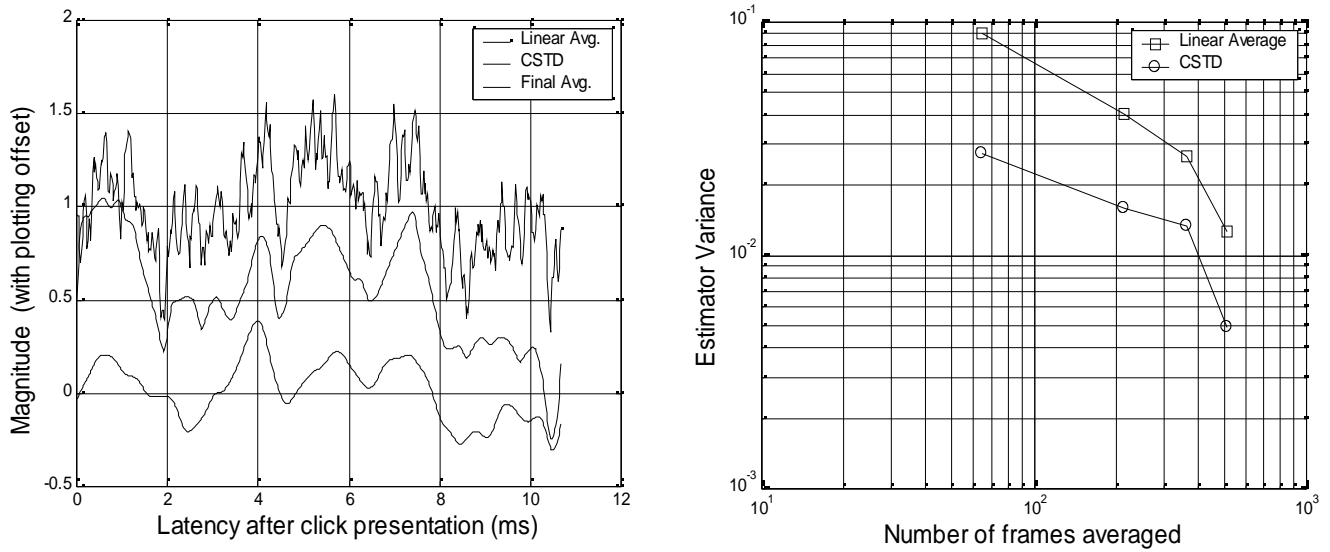




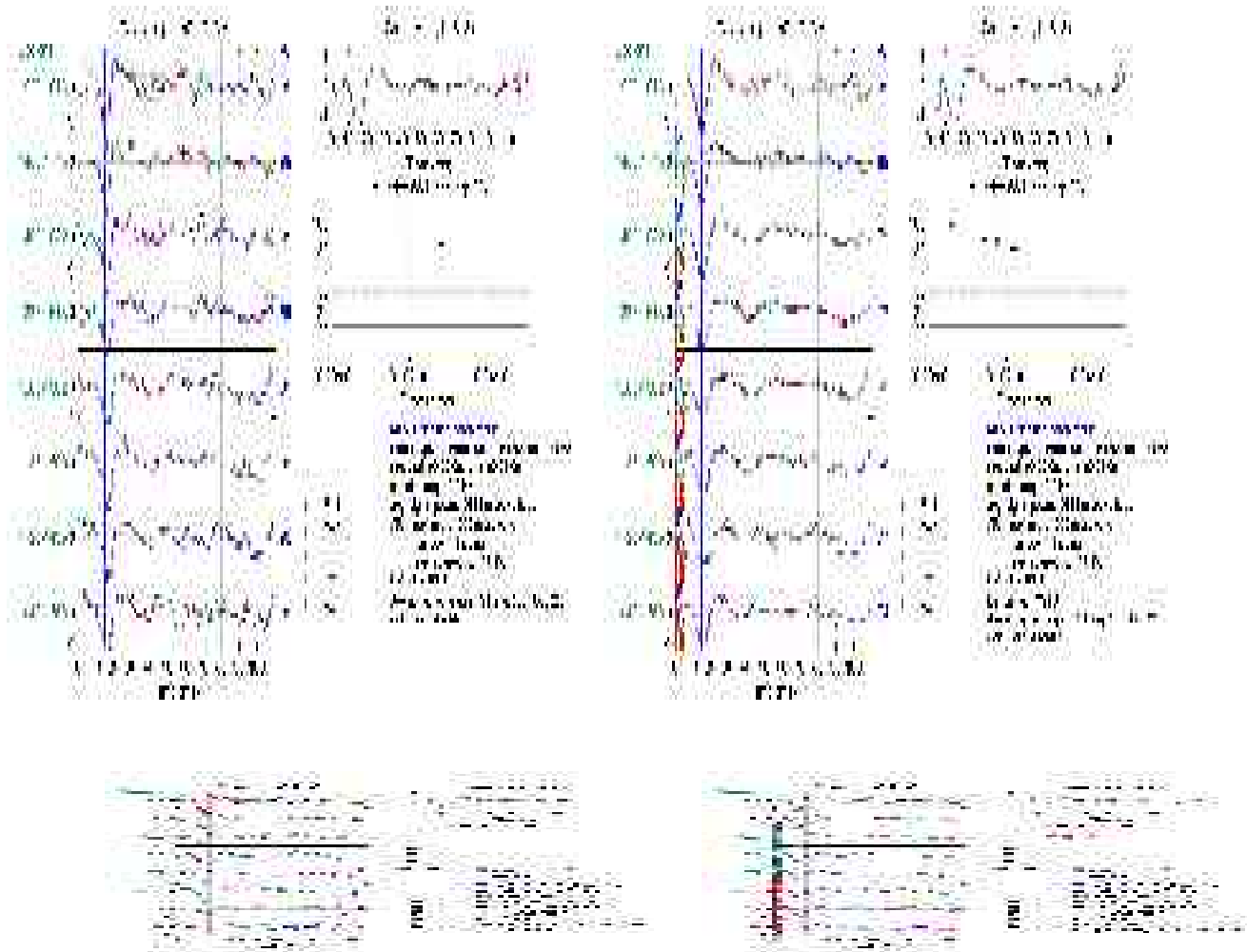
**Figure 5:** Left: Performance of the novel wavelet algorithm compared to linear averaging 512 data frames (Top: linearly averaged 512 frames, middle: CSTD wavelet denoised 512 frames, bottom: original waveform (“true value”). Right: Variance comparison between linear averaging (top) and novel wavelet algorithm (bottom).



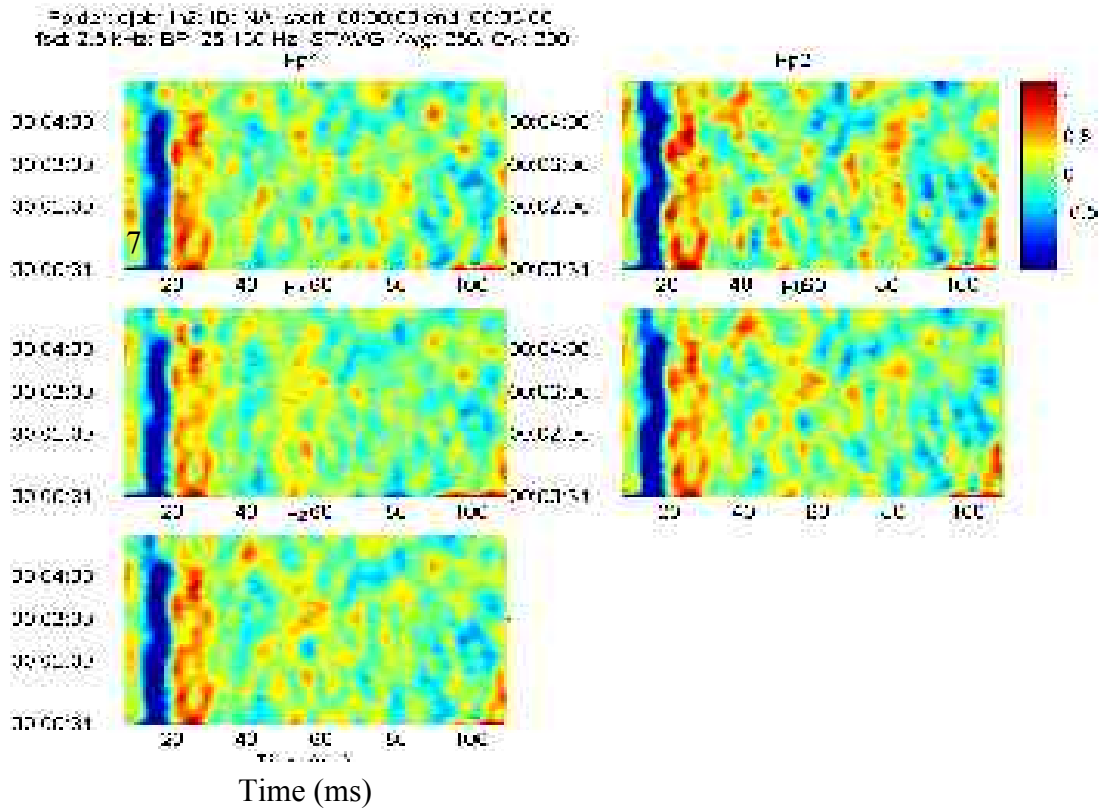
**Figure 6:** Performance of CSTD algorithm compared to linear averaging 256 data frames. (a): Template of AMLR evoked potential waveform from Spehlmann; (b): linear average of 8192 AMLR frames; (c): Single frame consisting of AMRL model plus WGN; (d): Linear average of 256 frames; (e): Result of CSTD algorithm.



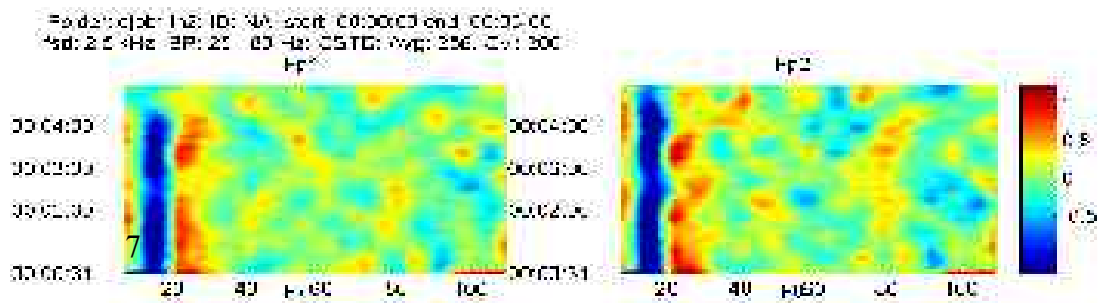
**Figure 7:** Left: Comparison of Linear Averaging to CSTD for 512 frames of ABR data. (Top: linearly averaged 512 frames, middle: CSTD wavelet denoised 512 frames, bottom: linearly averaged 8,192 frames (“true value”). Right: Variance comparison between linear averaging and CSTD, as a function of number of frames.

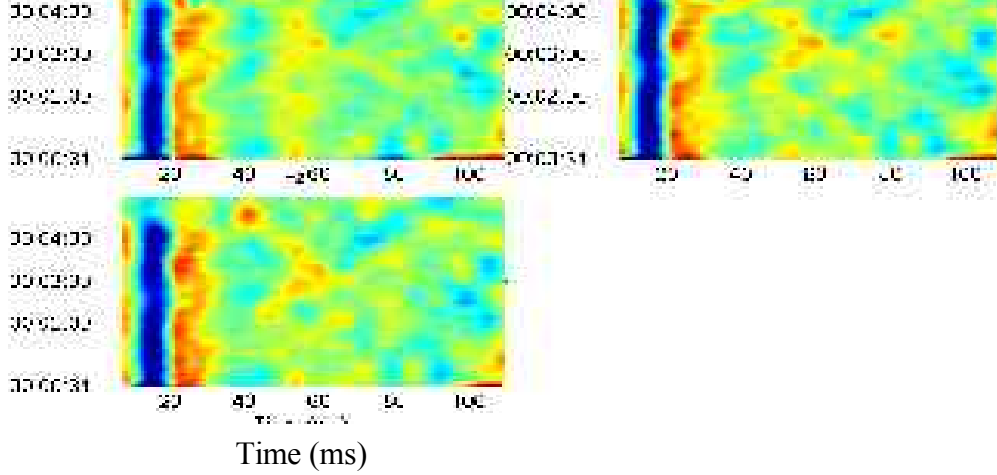


**Figure 8:** Comparison between linear averaging and CSTD for AMLR. Left: consecutive averages of 256 frames of 110 ms each. Right: Results of CSTD algorithm applied to EEG frames covering the temporal range [7.5 ms, 110 ms] (Note: In the interval [0, 7.5 ms] the signal displayed is simply the linear average of 256. It is of little importance for AMLR interpretation). Color code for leads: Blue: “Fp1,” Green: “Fp2,” Red: “F7,” Cyan: “F8,” Magenta: “Fz.”



**Figure 9:** Color-coded compressed evoked potential array (CEPA) for averaging method. Each horizontal line in an image represents a reconstructed mid-latency evoked potential. Wave Na (See also Figure 6) is clearly identifiable by the vertical band of blue at latency around 15 ms. Wave Pa appears as a red band at latency around 25 ms. Waves Nb and Pb should respectively appears as vertical bands of cyan-blue and yellow-red but are not as easily identifiable.





**Figure 10:** Color-coded compressed evoked potential array (CEPA) for CSTD method. Each horizontal line in an image represents a reconstructed mid-latency evoked potential. Wave Na (See also Figure 6) is clearly identifiable by the vertical band of blue at latency around 15 ms. Wave Pa appears as a red band at latency around 25 ms. Waves Nb and Pb should respectively appears as vertical bands of cyan-blue and yellow-red but are not as easily identifiable.

## References

- [1] Beylkin, G. (1992). On the representation of operators in bases of compactly supported wavelets. *SIAM Journal of Numerical Analysis*. 6-6:1716-1740.
- [2] Burke, M.J., and Gleeson, D.T. (2000). A micropower dry-electrode ECG Preamplifier. *IEEE Transactions on Biomedical Engineering*. Volume 47:2.
- [3] Causevic, E. (2001). *Fast Wavelet Estimation of Weak Biosignals*. Ph.D. Thesis, Washington University.
- [4] Coifman, R. R. and M. V. Wickerhauser (1995). Adapted waveform ‘de-noising’ for medical signals and images. *IEEE Engineering in Medicine and Biology* 14((5) September/October): 578-586.

- [5] Coifman, R.R., Donoho, D.L. (1995). Translation invariant denoising. *Technical Report 475*. Department of Statistics, Stanford University.
- [6] Don, M., & Elberling, C. (1996). Use of quantitative measures of auditory brain-stem response peak amplitude and residual background noise in the decision to stop averaging. *Journal of the Acoustical Society of America* 99: 491-499.
- [7] Don, M., Elberling, C. and Waring, M. (1984). Objective detection of averaged auditory brainstem responses. *Scandinavian Audiology* 13: 219-228.
- [8] Donoho, D.L. (1995). De-noising by soft thresholding. *IEEE Transactions on Information Theory*. Volume 41: 613-627.
- [9] Donoho, D.L. et al. (1995). Wavelet shrinkage: Asymptopia?. *Journal of Royal Statistics Society B*. 57(2):301-369.
- [10] Donoho, D.L. and Johnstone, I. (1994). Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, December, 81:425-455.
- [11] IEC [International Electrotechnical Commission]. (1994). Auditory test signals of short duration of audiometric and neuro-otological purposes. *International Standard IEC 645-3*. 1<sup>st</sup> Ed. Geneva, Switzerland.
- [12] Elberling, C., & Don, M. (1984). Quality estimation of averaged auditory brainstem responses. *Scandinavian Audiology* 13: 187-197.
- [13] Everest Biomedical Instruments Company. (2001). AudioScreener OAE+ABR Engineering Design Documentation. St. Louis, Missouri.
- [14] Hall, J.W.III (1992). *Handbook of Auditory Evoked Responses*. Boston, Allyn and Bacon.
- [15] Hoppe, U., et al. (2001) An automatic sequential recognition model for cortical auditory evoked potentials. *IEEE Transactions on Biomedical Engineering* 48:154-164.
- [16] Hyde, M.L., Davidson, M.J., and Alberti, P.W. (1991). Auditory test strategy. In J.T. Jacobson & J.L. Northern (Eds.), *Diagnostic Audiology* (pp. 295-322). Austin, TX:Pro-ed.
- [17] Jewett, D.L. and Williston, J.S. (1971). Auditory evoked far fields averaged from the scalp of humans. *Brain*, 4, 681-696.
- [18] Mallat, S. (1998). *A Wavelet Tour of Signal Processing*. Academic Press. San Diego California.
- [19] Misulis, E.K. (1997). *Essentials of Clinical Neurophysiology*. 2<sup>nd</sup> ed. Butterworth-Heinemann. Newton MA.
- [20] National Center for Hearing Assessment and Management (2001) [www.infanthearing.org](http://www.infanthearing.org).
- [21] Norton, S.J. et al. (2000). Identification of neonatal hearing impairment: summary and recommendations. *Ear & Hearing* (21):529-535.

- [22] Parthasarathy, T., P. Borgsmuller, et al. (1998). "Effects of repetition rate, phase, and frequency on the auditory brainstem response in neonates and adults." *Journal of the American Academy of Audiology* 9: 134-140.
- [23] Rao, M.R., Bopardikar, A.S. (1998). *Wavelet Transforms: Introduction to theory and applications*. Addison Wesley. Reading, Massachusetts.
- [24] Quiroga, R.Q., 2000: *Obtaining Single Stimulus Evoked Potentials with Wavelet Denoising*, *Physica*, 145: 278-292.
- [25] Rundshagen, I., Schnabel, K., Schulte am Esch, J., (2002) *Impaired explicit memory after recovery from propofol/sufentanil anaesthesia is related to changes in the midlatency auditory evoked response*, *British Journal of Anaesthesia*, 89 (3): 376-381.
- [26] Sendur, L., Selesnick, I. W. (2002), *Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency*, *IEEE Trans. On Signal Processing*, 50(11): 2744-2756.
- [27] Spehlmann, Misulis, K. E., Fakhoury, T., *Evoked Potential Primer*, 2001, Elsevier Science & Technology Books.
- [28] Vetterli, M., Kovacevic, J., (1995), *Wavelets and Subband Coding*, Prentice Hall.
- [29] Wickerhauser, M. V. (1994). *Adapted Wavelet Analysis from Theory to Software*. A.K. Peters. Natick Massachusetts.