# Development of Gradient Descent Adaptive Algorithms to Remove Common Mode Artifact for Improvement of Cardiovascular Signal Quality

Edward J. Ciaccio,[1,2] and Evangelia Micheli-Tzanakou[3]

[1]Department of Pharmacology PH7W318, Columbia University, 630 W 168th Street, New York, NY 10032, USA; [2]Department of Biomedical Engineering, Columbia University, New York, NY, USA; and [3]Department of Biomedical Engineering, Rutgers University, Piscataway, NJ, USA

**Abstract**—*Background*: Common-mode noise degrades cardiovascular signal quality and diminishes measurement accuracy. Filtering to remove noise components in the frequency domain often distorts the signal. *Method*: Two adaptive noise canceling (ANC) algorithms were tested to adjust weighted reference signals for optimal subtraction from a primary signal. Update of weight $w$ was based upon the gradient term of the steepest descent equation: $\nabla = \partial\xi/\partial w = \partial E[\varepsilon_k^2]/\partial w_k$, where the error $\varepsilon$ is the difference between primary and weighted reference signals. $\nabla$ was estimated from $\Delta\varepsilon^2$ and $\Delta w$ without using a variable $\Delta w$ in the denominator which can cause instability. The Parallel Comparison (PC) algorithm computed $\Delta\varepsilon^2$ using fixed finite differences $\pm \Delta w$ in parallel during each discrete time $k$. The ALOPEX algorithm computed $\Delta\varepsilon^2 \cdot \Delta w$ from time $k$ to $k + 1$ to estimate $\nabla$, with a random number added to account for $\Delta\varepsilon^2 \cdot \Delta w \to 0$ near the optimal weighting. *Results*: Using simulated data, both algorithms stably converged to the optimal weighting within 50–2000 discrete sample points $k$ even with a SNR = 1:8 and weights which were initialized far from the optimal. Using a sharply pulsatile cardiac electrogram signal with added noise so that the SNR = 1:5, both algorithms exhibited stable convergence within 100 ms (100 sample points). Fourier spectral analysis revealed minimal distortion when comparing the signal without added noise to the ANC restored signal. *Conclusions*: ANC algorithms based upon difference calculations can rapidly and stably converge to the optimal weighting in simulated and real cardiovascular data. Signal quality is restored with minimal distortion, increasing the accuracy of biophysical measurement.

**Keywords**—Adaptive noise cancellation, Cardiovascular, Mean squared error, Noise cancellation, Steepest descent.

Address correspondence to Edward J. Ciaccio, Department of Pharmacology PH7W318, Columbia University, 630 W 168th Street, New York, NY 10032, USA. Electronic mail: ciaccio@columbia.edu

## BACKGROUND

Cardiac and other biomedical signals are often degraded by noise, motion artifact, and proximate bioelectric sources.[6,16,30,31] These extraneous components, which can reduce the signal to noise ratio (SNR) to 1 or less, decrease the quality of the biological signal to the extent that quantitative and even qualitative measurements become difficult if not impossible to make. Subtle features in the biological signal of interest, whether they are of short time duration or of low amplitude or both, are often masked by both broadband noise and low frequency motion artifact. Efforts to reduce noise and motion artifact can increase the complexity and cost of the transduction device, while decreasing its utility, ease of use, and the rapidity with which it can be employed for recording of medical cardiovascular signals.[11] Following data recording, frequency domain filters are often used to remove noise and motion artifact; however, the frequency content of signal and noise often overlap, so that either portions of the noise content remain and/or portions of the frequency content of the signal are removed.[2,3] The latter results in distortion of the signal and loss of information content.[2,3] Noise cancellation[25,27–29] is the process whereby common-mode noise and artifact, which is the same on all input sensors except for differences in amplitude and perhaps in phase, are removed by subtractive methods. The primary input is a recording of the actual signal of interest, plus noise, much of which is often common-mode in biological recordings. The reference input or inputs consist of noise which is uncorrelated with the primary signal and contains common-mode components, and may contain additional components that are not common mode.[24,26]

In previous work, cancellation of common-mode artifact with a constant weighting was used to improve signal quality.[10] Two inputs were incorporated into the device: one from a primary and one from a reference piezoelectric sensor, with the piezoelectric transducers operating in compression mode. The primary signal consisted of a blood pressure tonometry signal which was acquired noninvasively by recording above the palpable radial artery pulse. The reference signal was acquired simultaneously by recording from a portion of the wrist lacking a palpable pulse. The common-mode noise components included mostly motion artifacts delivered to the piezoelectric transducers by patient movement and ambient vibrations, and from line frequency interference. Using a 1:1 subtraction, or some other fixed subtraction of reference from primary signals, most of the common-mode components were removed and the pulse restored. Yet, further measurement showed the best weighting for subtraction of common-mode noise was time-varying for this device. An adaptive algorithm was developed using fixed weight update increments for adjustment of the best weighting of the reference signal.[9] The performance index for weight adjustment was the estimated mean squared error (MSE). The weights were updated by only a small fixed amount during each adaptive iteration to simplify the process and stabilize the weight adjustment. However, the response to any rapid changes in common-mode characteristics that occurred was relatively slow. Herein, we study two gradient descent adaptive algorithms for cancellation of the additive noise present on cardiovascular signals.

## METHOD

### Design of Equations for Adaptive Noise Cancellation (ANC)

Adaptive filtering methods are useful for cancellation of correlated signal components.[25,27–29] The process is called adaptive noise cancellation (ANC). When common-mode noise is present on the recorded biomedical signal $d$ (termed the primary or desired signal) and it is also on a reference signal $x$, then the reference signal can be subtracted from the primary to remove the common-mode noise. If the common-mode noise is in-phase and the same amplitude on both signals, and if it is time-invariant in both phase and amplitude, it would be completely cancelled by 1:1 subtraction:

$$\varepsilon = d - x \tag{1}$$

where $\varepsilon$ is the error signal. The error signal is a measure of the similarity between the desired and reference signals, and can be used to develop a performance criterion, or index, for determining how well the common-mode noise is cancelled by the subtraction. The 'desired' signal $d$ is so termed because the reference signal is weighted to maximize its similarity to $d$.[25,27–29] If the common-mode noise is in phase but not exactly 1:1 in amplitude on both signals, and is time invariant in both phase and amplitude, then a scaling factor or weight $w$ can be used to best adjust the reference signal for subtraction from the primary signal:

$$\varepsilon = d - w \cdot x \tag{2a}$$

$$\varepsilon = d - y \tag{2b}$$

where $y$ is termed the output signal.[26] For example, for complete cancellation of common-mode noise that is twice the amplitude on $d$ as compared with $x$, the optimal weight ($w^*$) would be 2. When the optimal weighting is constant over time, it can be tuned manually either in hardware or software for best noise cancellation. A problem arises however when the optimal weighting is time varying, i.e., the amplitude ratio of common-mode noise on the primary and reference varies over time. If we use the mean squared error (MSE), symbolized by $\xi$, as the performance index for best weighting:

$$\xi = E[\varepsilon^2] \tag{3}$$

where $E$ is the expectation operator, then a plot of $\xi$ versus $w$ would be parabolic with the optimal weighting $w^*$ occurring at the minimum (Fig. 1). For parabolic functions, the equation of steepest descent can be used for weight update from any arbitrary starting point[25]:

$$w_{k+1} = w_k - \mu \nabla_k \tag{4}$$
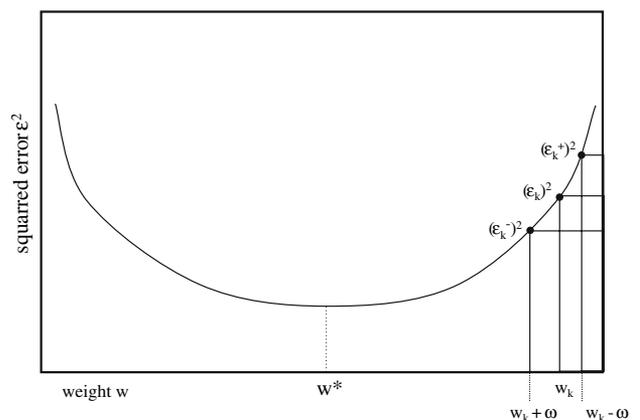


**FIGURE 1. Schematic of the method of parallel comparison (PC). The squared error is used as the performance index. Adaptation is based upon finite differences in weighting.**

where $w$ is the weight used to adjust the reference signal amplitude during discrete time $k$, $\mu$ is a convergence coefficient, and the gradient $\nabla_k$ is given by:

$$\nabla_k = \partial E[\varepsilon_k^2]/\partial w_k = \partial \xi_k/\partial w_k \qquad (5)$$

with $\partial$ denoting the partial derivative. Since the expectation operator is taken over all time, for practical implementation the gradient must be estimated. One way to do this is to use the difference equation[27]:

$$\nabla_k \approx \Delta\varepsilon_k^2/\Delta w_k \qquad (6)$$

which is the squared error difference divided by the weight difference at time $k$. Therefore substituting Eq. (6) into Eq. (4) and expanding the differences:

$$w_{k+1} = w_k - \mu\Delta\varepsilon_k^2/\Delta w_k \qquad (7a)$$

$$\Delta\varepsilon_k^2 = \varepsilon_k^2 - \varepsilon_{k-1}^2 \qquad (7b)$$

$$\Delta w_k = w_k - w_{k-1} \qquad (7c)$$

Equation (7) provides the direction (sign) for weight update automatically. A major problem with estimating the gradient using Eq. (7a) is that the term in the denominator $\Delta w_k$ can approach zero, which will make the weight update subject to instability. One way to increase stability is to drop the $\Delta w_k$ from the denominator using the following paradigm:

$$\nabla_k \approx \Delta\varepsilon_k^2 \qquad (8a)$$

$$s = \text{sign } (\Delta w_k) \qquad (8b)$$

$$w_{k+1} = w_k - \mu \cdot s \cdot (\Delta\varepsilon_k^2) \qquad (8c)$$

Use of Eq. (8) will result in more rapid convergence far from the optimal weighting and slower convergence near the optimal weighting as compared to Eq. (7).

An alternative method for computation of $\varepsilon_k$, rather than determining error difference between successive times, is to make the computation at each discrete time $k$ based upon finite differences[9,10,27]:

$$\varepsilon_k^+ = d_k - (w_k + \omega) \cdot x_k \qquad (9a)$$

$$\varepsilon_k^- = d_k - (w_k - \omega) \cdot x_k \qquad (9b)$$

$$\Delta\varepsilon_k = \varepsilon_k^+ - \varepsilon_k^- \qquad (9c)$$

$$\Delta w_k = 2\omega \qquad (9d)$$

$$w_{k+1} = w_k - \mu \cdot [(\varepsilon_k^-)^2 - (\varepsilon_k^+)^2]/2\omega \qquad (9e)$$

where $\omega$ is a constant, finite difference in weighting used to compute the error at discrete time $k$. Thus

the sign of $[(\varepsilon_k^-)^2 - (\varepsilon_k^+)^2]$ determines the direction of weight update. The concept is illustrated in Fig. 1: discrete changes $\pm\omega$ in weight $w$ shift the error so that the gradient can be estimated based upon $[(\varepsilon_k^+)^2 - (\varepsilon_k^-)^2]/2\omega$. We term this adaptive method the parallel comparison or PC algorithm, because errors based on finite differences in weight are computed in parallel at discrete time $k$.

A second method to compute the weight update using difference equations is to include $\Delta w_k$ in the numerator rather than in the denominator, so that from Eq. (4):

$$w_{k+1} = w_k - \mu\Delta\varepsilon_k^2 \cdot w_k \qquad (10)$$

This method has the advantage, as compared to Eq. (8), of automatic computation of the sign of the weight update from the product of $\Delta\varepsilon_k^2$ with $\Delta w_k$, while still maintaining stability when $\Delta w \to 0$. However, the effect of having $\Delta w_k$ in the numerator will be as follows. The speed of convergence when $w$ is far from the optimal weighting will be further increased as compared with Eq. (8), because both $\Delta\varepsilon_k^2$ and $\Delta w_k$ will be large per unit step. The speed of convergence when $w$ is near the optimal weighting will be diminished as compared with Eq. (8), because $\Delta\varepsilon_k^2$ and $\Delta w_k$ will be small per unit step. Yet, convergence speed can be improved as $w \to w^*$ by adding a random number to the update equation:

$$w_{k+1} = w_k - \mu\Delta\varepsilon_k^2 \cdot \Delta w_k + R \qquad (11)$$

where $R$ is a random number with truncated limits.[17,21] The limits of $R$ are truncated such that it will have a large effect near the bottom of the parabola, where $\mu\Delta\varepsilon_k^2 \cdot \Delta w_k$ in Eq. (11) would be expected to be small, but a negligible effect at the sides of the parabola, where $\mu\Delta\varepsilon_k^2 \cdot \Delta w_k$ would be expected to be large. Equation (11) describes the ALOPEX algorithm. ALOPEX is an acronym for ALgorithm Of Pattern EXtraction, which refers to its use in pattern recognition.[17,21] Herein, Eq. (11) is used for weight update of a single reference input, but it can also be used to update multiple weights $j = 1, \ldots, n$ when there are $n$ reference inputs.[17,21] For added stability, we computed the ALOPEX weight update using:

$$w_{k+1} = w_k - \mu\Sigma_k\Delta\varepsilon_k^2 \cdot \Sigma_k\Delta w_k + R \qquad (12)$$

where the summation sign $\Sigma_k$ denotes that averages of $\Delta\varepsilon_k^2$ and $\Delta w_k$ over $k = 10$ previous iterations was used.

### Experiments

The ANC algorithms described in Eq. (9) (PC) and Eq. (12) (ALOPEX) were tested. The PC algorithm was implemented with two reference inputs:

$$\varepsilon = d - (w_1 x_1 + w_2 x_2) \qquad (13)$$

We used a triangle wave (tri) of intermediate period ($k = 25$ sample points) to simulate a blood pressure pulse, and sinusoids with longer and shorter periods (selected as $k = 100.5$ and $13.2$ sample points so as not to be multiplicatively related to tri) to simulate common-mode motion artifact and line frequency interference, respectively. The signal and common-mode artifact were added to the inputs as follows:

$$d = .45 \text{ tri}(2\pi k/25.0) + 2.8 \sin(2\pi k/13.2)$$
$$+ 2.65 \sin(2\pi k/100.5) \qquad (14a)$$

$$x_1 = 1.4 \sin(2\pi k/13.2) + 2.65 \sin(2\pi k/100.5) \quad (14b)$$

$$x_2 = 5.6 \sin(2\pi k/13.2) + 2.65 \sin(2\pi k/100.5) \quad (14c)$$

where $d$ is the primary input, $x_1$, $x_2$ are reference inputs, and noises $n_1 = \sin(2\pi k/13.2)$ and $n_2 = \sin(2\pi k/100.5)$. The noise amplitudes were made several times larger than that of the signal (Eq. 14a; SNR$\sim$1:6). Weights $w_1$ and $w_2$ were initialized to zero. For faster convergence, we used two values for the convergence coefficient $\mu$: 0.003 and 0.00003 with the threshold for change from large to small $\mu$ being the point at which the squared error halved from its value at initialization.

The ALOPEX algorithm was implemented as described in Eq. (12). A simulated blood pressure pulse shape (bpp) was used as the primary signal, with a common-mode sinusoid added as noise. The signal and common-mode noise contributions to the inputs were as follows:

$$d = 1.7 \text{ bpp}(2\pi k/80) + 0.7 \sin(2\pi k/28) \qquad (15a)$$

$$x_1 = 0.7 \sin(2\pi k/28) \qquad (15b)$$

We ran this simulation after initializing weight $w_1$ to zero and using a value $\mu = 0.001$.

The PC and ALOPEX algorithms were then tested on real data using a single noise cancellation weight. An electrogram recording acquired from the epicardial surface of a canine heart during normal sinus rhythm was used as the cardiovascular signal. The recording was bipolar and the electrodes were constructed of 1 mm silver disks attached to silver wires, with a bipolar electrode spacing of 3.5 mm. The digitization rate was 1 kHz. Additional details concerning the data acquisition system used for recording cardiac electrograms have been described elsewhere.[11]

## RESULTS

Figure 2 shows the signals used for the PC algorithm simulation. The primary input $d$ is the larger signal and it is composed of noise components with periods $T = 13.5$ and $T = 100.5$ sample points in addition to the simulated cardiovascular signal with period $T = 25$ sample points (see Eq. 14). The error signal $\varepsilon$ was computed from Eq. (9) formulated for a two weight system. Since the weights are adaptively updated, ideally they will converge so that the triangle component of the primary (the simulated cardiovascular signal), which is not common-mode, will appear at the error signal without additive noise. Convergence using the PC algorithm required approximately 100 iterations (100 sample points) for the signal amplitudes shown. Some common-mode noise still appears on the ANC (i.e., error) signal, however its SNR $\gg$ 1 and it is stable over time.

In Fig. 3 the path of weights $w_1$ and $w_2$ for convergence are shown. Each point in the scatterplot denotes an iterative weight update. Initially both weights were set to zero. The weight update involves large steps at first (larger spacing between points), and then progressively smaller steps (smaller spacing between points) toward convergence to the optimal weighting. Upon convergence, $w_1 = 0.66$ and $w_2 = 0.34$; with this combination of weights:

$$0.66 \cdot x_1 + 0.34 \cdot x_2 \sim 2.8 \sin(2\pi t/13.2)$$
$$+ 2.65 \sin(2\pi t/100.5)$$

Thus the weighted sum of the two reference inputs $x_1$ and $x_2$, upon convergence, approximately equals
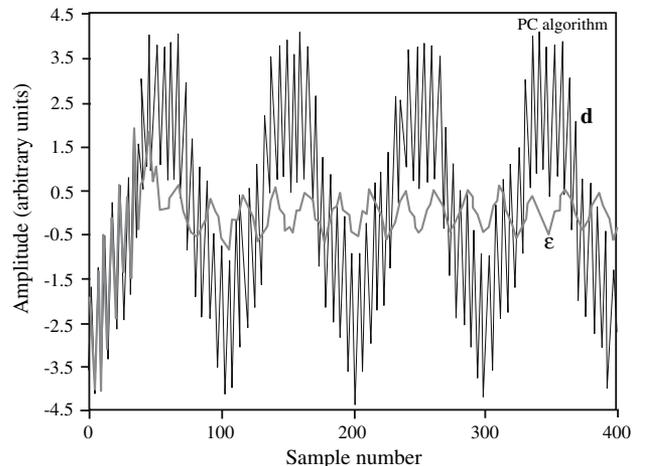
**FIGURE 2. Signals used for parallel comparison (PC) testing. The desired signal d (large deflections, black) and error signal ε (smaller deflections, gray) are shown. Complete convergence occurs at approximately the 100th iteration (100 sample points).**
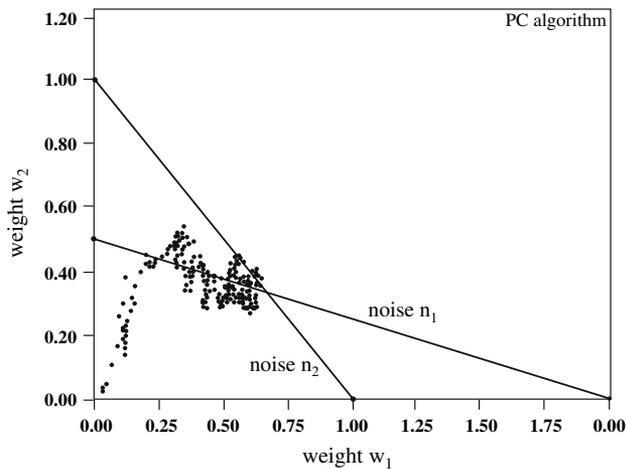
FIGURE 3. Scatterplot of the weight convergence for the parallel comparison (PC) algorithm with two reference inputs. Each point shows the values of the two weights at a particular iteration $k$. The update proceeds from $(w_1, w_2) = (0,0)$ toward the line labeled 'noise $n_1$', and then proceeds along the line labeled noise $n_1$ toward the intersection between the two lines.



FIGURE 4. Signal + noise (top), signal (middle), and signal + noise with adaptive noise cancellation (bottom) using the ALOPEX algorithm. Traces are offset for clarity.

and therefore will cancel by subtraction the additive noise on the primary input (Eq. 14a). The lines in Fig. 3 show how the gains are related for common-mode noise cancellation. The line extending from $w_1 = 2$ to $w_2 = 0.5$ shows the value of the weights that will completely cancel the common-mode noise $\sin(2\pi t/13.2)$ (Eq. 14). The line extending from $w_1 = 1$ to $w_2 = 1$ shows the value of the weights that will completely cancel the common-mode noise $\sin(2\pi t/100.5)$ (Eq. 14). From an initialized weighting of zero, during the update process the weights converge to the line in which the common-mode noise $\sin(2\pi t/13.2)$ is completely cancelled. Upon arrival at this line, further adaptive update involves converging to the second line (optimal cancellation of the common-mode noise $\sin(2\pi t/100.5)$) while approximately maintaining optimal cancellation of the other common-mode noise.

In Fig. 4, the signals used for the ALOPEX algorithm simulation are shown. The simulated blood pressure trace is at the center. The desired signal is offset at the top of the figure. Although the main features of the simulated pulse can still be observed (systolic and diastolic components and dicrotic notch) they are skewed and probably unusable for quantitative or even qualitative analysis. The error signal, being the signal + noise with ANC, is shown offset at the bottom of the figure. Much of the noise artifact is removed. Small ripples are caused by flux of the weighting about a locus centered upon the optimal weight value.

Figure 5 shows the convergence process for the signals of Fig. 4, based on Eq. (12). Initially, because

the weighting is not optimal, the squared error term $\Sigma_k \Delta \varepsilon_k^2$ is large. It is also pulsatile because the window $\Sigma_k$ slides across a periodic signal. Its amplitude is also relatively large initially since the weighting for cancellation of common-mode artifacts is initially incorrect. As the weights converge to optimal, the $\Sigma_k \Delta \varepsilon_k^2$ decreases in value and its amplitude diminishes. Upon convergence after $\sim$2000 sample points, the mean $\Sigma_k \Delta \varepsilon_k^2$ remains constant and contains only a small pulsative component.

The ALOPEX and PC algorithms were then tested on canine electrogram data. In Fig. 6a is shown the recorded electrogram signal over a 1600 ms interval. The heart rate, which was in sinus rhythm but rapid, is $\sim$3/s (period $T \sim$ 325 ms). An 18 Hz triangle wave ($T \sim$ 55 ms) with an amplitude of 2 mV was added as noise. The signal + noise is shown in panel (b), and the signal itself is completely embedded in the noise except at the times of the two largest peak deflections from panel (a), where two small deflections can be observed. Noise cancellation using ALOPEX is shown in panel (c). The weight was initialized at $w_o = 1.5$, with constant convergence coefficient $\mu = 0.3$, and a Gaussian random number $R$ truncated at $\pm$1E$-$4. Convergence to the optimal weighting is complete before 100 ms. Only a short segment of the triangle wave noise appears in panel (c) prior to convergence. The result of the PC algorithm when used for adaptive noise cancellation on the same signal + noise is shown in panel (d). The parameters included an initial weighting $w_o = 1.5$, with a constant finite difference $\pm \omega = \pm 0.01$ and constant convergence coefficient $\mu = 0.01$. The PC algorithm took slightly longer to converge than ALOPEX, but was still complete in
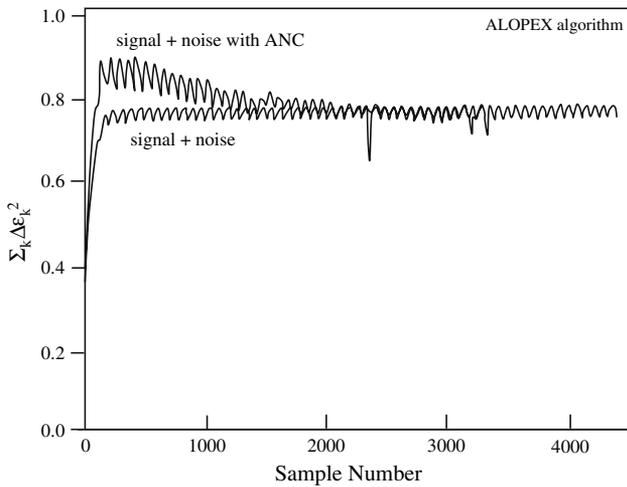
**FIGURE 5. Convergence of the optimal weighting using ALOPEX for adaptive noise cancellation (signals shown in Fig. 4). The sum of squares error converges to that of the pulse without added artifact after approximately 2000 sample points (25 simulated heart cycles).**

~100 ms. Both algorithms show stability for noise cancellation even though the signal is highly pulsatile (panel (a)) and the SNR is initially ~1:5 (panel (b)).

In Fig. 7a is shown a longer sequence of the electrogram signal used for processing, and below it is shown the frequency spectrum (Fig. 7b). The electrogram signal with additive noise is shown in Fig. 7c, and in Fig. 7d its frequency spectrum is shown,

centered on 18 Hz. In Fig. 8 is shown sequences of the ANC (i.e., error) signal for ALOPEX (Fig. 8) and for the PC algorithm (Fig. 8), with their respective frequency spectra beneath (panels (b) and (d)). The frequency spectra are similar, although not the same. They mainly differ in that the PC algorithm harmonics are slightly less than those for the ALOPEX algorithm. The frequency spectrum when ALOPEX is used for ANC (Fig. 8b) most resembles the original frequency spectrum of the signal (Fig. 7b).

## DISCUSSION

In this study two algorithms were tested for their ability to cancel common-mode noise in cardiovascular signals. The algorithms are rooted in the equation for steepest descent but modified to allow difference equations to approximate the MSE gradient. Stability is maintained by preventing $\Delta w \rightarrow 0$ in the denominator.

### Convergence to the Optimal Weighting

Pulsatile and periodic signals, along with several additive common mode noise components, were used to test the PC and ALOPEX algorithms, and in each case convergence was rapid and stability was maintained. In the case of real data, the electrogram signal used for ANC testing was pulsatile with varying
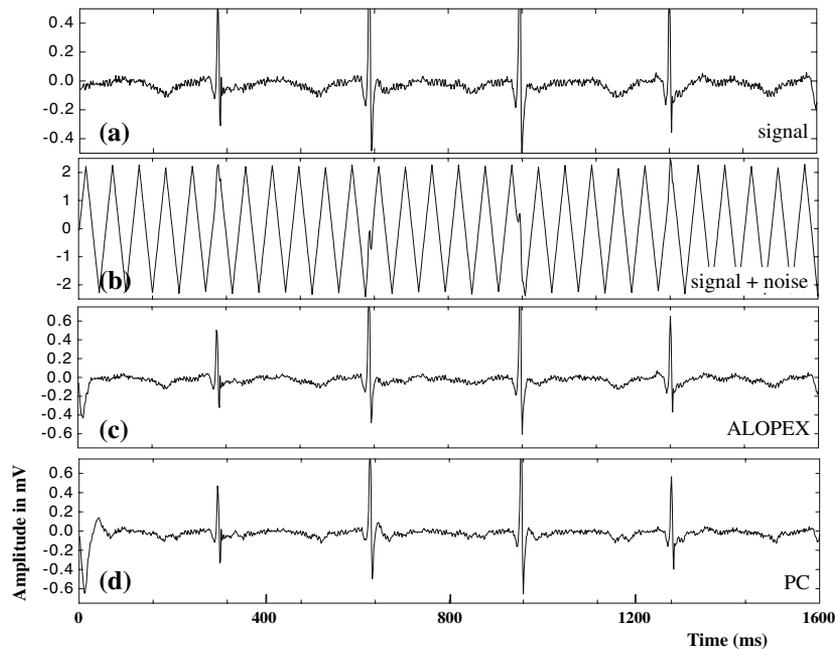


**FIGURE 6. Test of PC and ALOPEX algorithms on canine electrogram data. (a) Signal; (b) Signal + noise; (c) Restored signal after ANC using ALOPEX; and (d) Restored signal after ANC using the PC algorithm. For both algorithms, convergence occurs in <100 ms.**
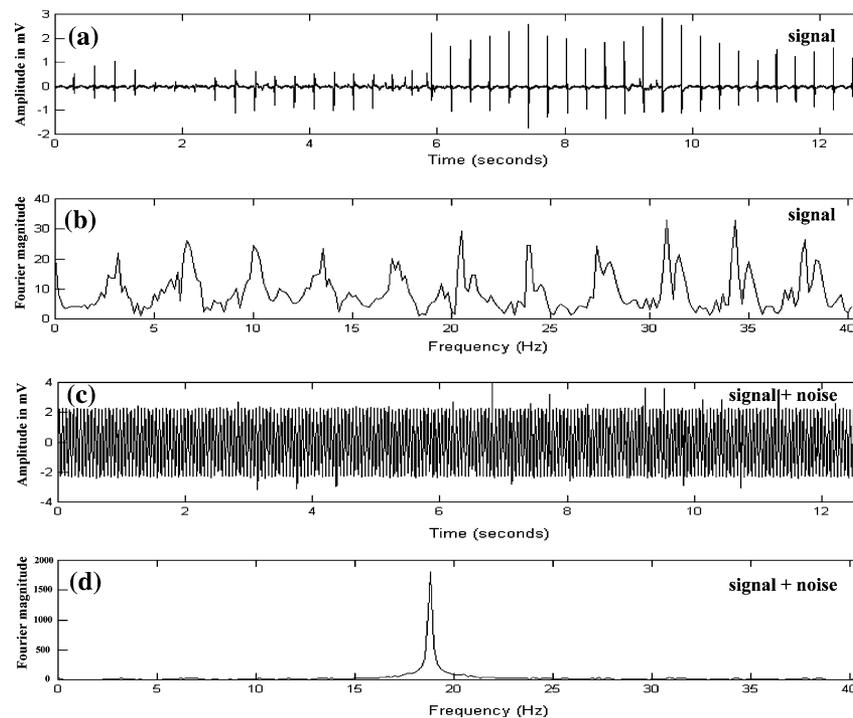
**FIGURE 7.** Signal and noise time series and frequency spectra. (a) Signal is shown over an interval of 12.4 s (first 1.6 s are shown expanded in Fig. 6a); (b) Frequency spectrum of this signal (first 8192 sample points); (c) Signal + noise (first 1.6 s are shown expanded in Fig. 6b); and (d) Frequency spectrum of the signal + noise.

amplitude and slightly varying rate (Fig. 7a). Yet, the stability of both algorithms for ANC was maintained at instances when large deflections occurred as well as during the intervening intervals (Fig. 8). The actual speed of convergence and stability of these ANC algorithms depend on the magnitude and interrelationships of the weight update parameters as well as the characteristics of the signal and noise. The PC algorithm uses finite differences in squared error to estimate the gradient in parallel during each time epoch, while the ALOPEX algorithm computes the squared error difference between successive time epochs to estimate the gradient. Since they rely upon differing methods for weight update, use of identical convergence coefficients for each would not be expected to result in identical rate of convergence nor the same degree of stability.

Figure 3 shows the convergence of the PC algorithm when using a two reference input (and two-weight) system. It reveals an interesting property of the convergence process. One of the additive noises is completely cancelled by the two-weight system when their values reside anywhere along either of the lines. Both of the additive noises are completely cancelled at the intersection of these lines, which is thus the point of optimal weighting. From an initial weighting of $w_1 = 0$, $w_2 = 0$, the update process converges first so as to completely cancel noise $n_1$ (the approach of the weight update, as represented by the points in the scatterplot, is approximately perpendicular to the line). There is then some overshoot toward the line at which complete cancellation of noise $n_2$ occurs. However, eventually the update centers again about the noise $n_1$ line and is directed toward the intersection of the lines. Near the intersection there is greater scatter because convergence then alternates between the two lines. These observations are in accord with the fact that the amplitude of noise $n_1$ is somewhat greater than noise $n_2$ (Eq. 14) and thus it makes a greater contribution to the squared error (greater squared contribution to the weight update).

During the ALOPEX simulation (Figs. 4 and 5), portions of the shape of the simulated cardiovascular signal (the systolic peak and the secondary peak following the dicrotic notch) were similar in amplitude and width to the additive noise that can be observed to be riding on the signal in the top trace of Fig. 4. The similarity of signal and noise at these moments would suggest that the algorithm might attempt to converge the weights so as to cancel the two major peaks of the primary rather than the common-mode artifact. Yet as the bottom trace in Fig. 4 illustrates, upon convergence the simulated blood pressure signal is mostly restored. If however the period of the common-mode
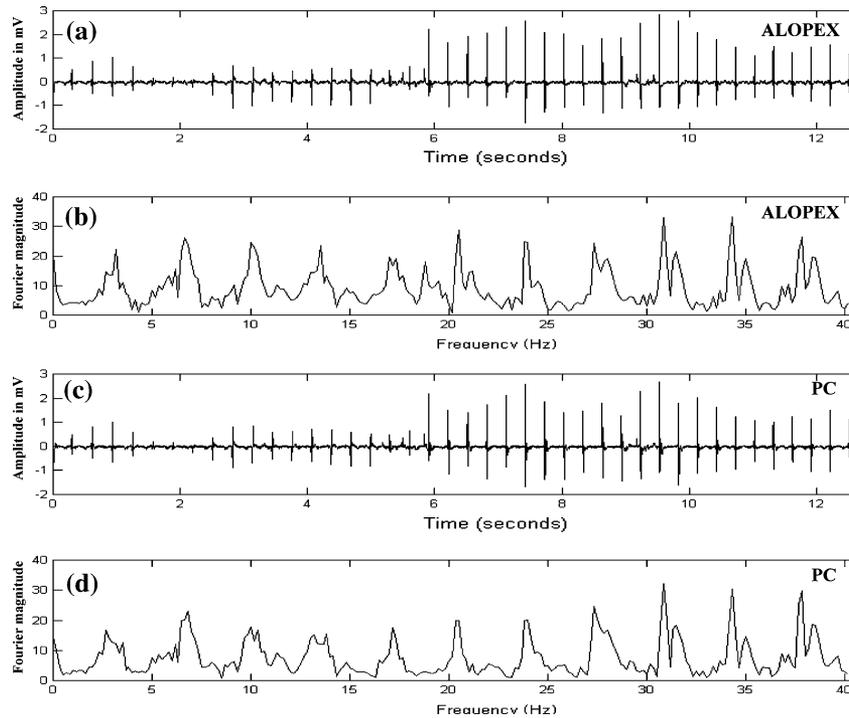
**FIGURE 8.** ANC signals using ALOPEX and PC algorithms, time series and frequency spectra. (a) Restored signal with ANC using ALOPEX; (b) Frequency spectrum; (c) Restored signal with ANC using the PC algorithm; and (d) Frequency spectrum. First 1.6 s of panels (a) and (c) are also shown in Fig. 6 panels (c) and (d), respectively.

artifact were such that its peaks always aligned with the peaks of the simulated blood pressure signal (i.e., if these signals were correlated), some cancellation of the blood pressure pulse signal would occur. This point emphasizes a requirement for adaptive algorithms such as these, that the common-mode artifact be uncorrelated with respect to the biomedical signal being filtered. Additionally there are three downward spikes in the squared error after 2000 sample points (Fig. 5). The downward spikes must be the result of some cancellation of the pulse itself and are caused by a transient similarity of the pulse signal to the weighted common-mode reference signal. However after complete convergence at ~4000 sample points, no such spikes occurred.

Although the common-mode noise was adaptively cancelled in these simulations, difference algorithms for gradient descent can also be used for adaptive pattern matching (APM)[22,23]. In pattern matching mode a template, or model signal, is weighted to adjust its form for best fit with an input signal based on the mean squared error or some other criterion of similarity.[5,7,8,12–14] At convergence to the optimal weighting, the final weighting is a quantitative measure of the degree of similarity between the two signals. APM has been shown to be useful in cardiovascular research for prediction of the location of reentrant ventricular tachycardia[7,8,12–14] and for region-finding in biomedical magnetic resonance images.[5] In fact, ANC can be considered to be a special case of APM in which the pattern (reference signal) is matched to the input (desired signal) discrete point by discrete point for an indefinite interval.[5,7,8,12–14] The error signal when convergence is complete is a measure of the dissimilarity (uncorrelated components) between pattern and input.

### Difference Algorithms versus LMS

Perhaps the most commonly used algorithm for adaptive noise cancellation and for some forms of pattern matching is the Least Mean Squares (LMS) algorithm[25,27–29]:

$$w_{k+1} = w_k - 2\mu\varepsilon_k x_k \qquad (16)$$

Advantages of LMS include ease of implementation, and optimal performance under practical conditions.[25,27–29] Thus, the LMS algorithm has enjoyed widespread application. The average performance of LMS is equal to that of Newton's method when both algorithms are tested with statistically nonstationary input signals.[26] Yet, the convergence coefficient $\mu$ for LMS must be carefully selected to prevent instability in the weight adjustment and to prevent cancellation of

signal components as well as common-mode noise. Although the weight update can be normalized by dividing the right-hand-term in Eq. (16) by the reference signal average, that average is computed over a predefined interval. Thus, transient spikes and sharp shifts in the signal baseline may not immediately be accounted for in the normalization process. This increases the probability that sudden changes in the performance surface will be erroneously interpreted as requiring a drastic change of weighting, causing cancellation of portions of the signal itself as well as common-mode noise, and/or incorrectly weighting the common-mode noise component after weight adjustment to account for the transient spike or sharp baseline shift.

Typically, LMS relies on a single reference input which is shifted by increments of one sample point to provide delayed inputs (tapped delay line) for cancellation of common-mode noise with phase adjustment,[25,27–29] but this presupposes that a single reference input accounts for all common-mode noise components and the relationship between them. In contrast, the ANC algorithms described herein can use multiple reference inputs, potentially increasing stability, and more accurately representing common-mode noise present on the desired signal. Use of multiple references would be expected to reduce the effect of a transient uncorrelated spike or baseline shift occurring on any one of them, because it is their summation (and thus weighted average) that forms the subtractive component for ANC. These algorithms were also shown to converge rapidly even for a SNR $\ll 1$.

ALOPEX has been used previously for rapid and stable convergence in a variety of settings including combinatorial optimization,[19] pattern matching,[1] fuzzy clustering of auditory neuronal responses,[4] detection of multiple sclerosis with visual evoked potentials,[15] and blood cell identification using neural networks.[20] When multiple reference inputs are used for image processing, parallel processing elements can be incorporated, one for each input, to speed convergence.[18] Herein, ALOPEX was shown to be useful for adaptive noise cancellation under conditions of SNR $\ll 1$. Incorporation of multiple reference inputs would be expected to increase the SNR of the filter output (error signal) while minimally increasing convergence time.

### Limitations and Future Directions

The simulations that were provided may not precisely mimic neither actual cardiac signals nor common-mode noise. Thus, the PC and ALOPEX algorithms may not perform as well (in terms of speed of convergence, or in terms of stability to large oscil-lations in noise cancellation weights) when they are used for ANC with certain real signals. Yet, the common-mode artifacts on each input were relatively large and of varied frequency, suggesting that these algorithms are robust to variations in signal and noise characteristics and to statistical nonstationarity. In the tests, a DC offset or bias was not added to the signals. Had a bias been added to the primary and/or the reference signals for either the PC or ALOPEX methods, another reference input composed of DC signal only would be required for complete cancellation of the bias. The bias reference input $x_b$ would be set to an arbitrary and constant value such as unity ($x_b = 1$).

The PC and ALOPEX algorithms as written do not adjust for phasic differences between the common-mode noises on each input. Phase shifts can occur for a variety of reasons and cause delay in the arrival of common mode components at one input with respect to another.[2,3] By using a tapped delay line[25,27–29] for each of the reference inputs, it would be possible to cancel common-mode artifact with phase shift, the subject of future research. A comparison of ANC algorithms to frequency domain filters is also a subject for future study.

### REFERENCES

[1]Ademoglu, A., and E. Micheli-Tzanakou. Analysis of pattern reversal visual evoked potentials (prvep) in Alzheimer's disease by spline wavelets. *IEEE Trans. Biomed. Eng.* 44:881–890, 1997.
[2]Akay, M. Biomedical Signal Processing. San Diego: Academic Press, 1994.
[3]Akay, M. Detection and Estimation Methods for Biomedical Signals. San Diego: Academic Press, 1996.
[4]Anderson, M. J., and E. Micheli-Tzanakou. Auditory stimulus optimization with feedback from fuzzy clustering of neuronal responses. *IEEE Trans. Inform. Technol. Biomed.* 6:159–170, 2002.
[5]Angelini, E. D., and E. J. Ciaccio. Optimized region finding and edge detection of knee cartilage surfaces from magnetic resonance images. *Ann. Biomed. Eng.* 31:336–345, 2003.
[6]Camps-Valls, G., M. Martinez-Sober, E. Soria-Olivas, R. Magdalena-Benedito, J. Calpe-Maravilla, and J. Guerrero-Martinez. Foetal ECG recovery using dynamic neural networks. *Artif. Intell. Med.* 31:197–209, 2004.

[7]Ciaccio, E. J. Localization of the slow conduction zone during reentrant ventricular tachycardia. *Circulation* 102:464–469, 2000.

[8]Ciaccio, E. J., A. W. Chow, D. W. Davies, A. L. Wit, and N. S. Peters. Localization of the isthmus in reentrant circuits by analysis of electrograms derived from clinical noncontact mapping during sinus rhythm and ventricular tachycardia. *J. Cardiovasc. Electrophysiol.* 15:27–36, 2004.

[9]Ciaccio, E. J. and G. M. Drzewiecki. Array sensor for arterial pulse recording-reduction of motion artifact. In Proceedings of the 1988 Fourteenth Annual IEEE Northeast Bioengineering Conference, 10–11 March 1988, pp. 66–69.

[10]Ciaccio, E. J., G. M. Drzewiecki, and E. H. Karam. Algorithm for reduction of mechanical noise in arterial pulse recording with tonometry. In Proceedings of the 1989 Fifteenth Annual IEEE Northeast Bioengineering Conference, 27–28 March, 1989, pp. 161–162.

[11]Ciaccio, E. J., A. E. Saltman, O. M. Hernandez, R. J. Bornholdt, and J. Coromilas. Multichannel data acquisition system for mapping the electrical activity of the heart. *Pacing Clin. Electrophysiol.* 28:826–838, 2005.

[12]Ciaccio, E. J., M. M. Scheinman, V. Fridman, H. Schmitt, J. Coromilas, and A. L. Wit. Dynamic changes in electrogram morphology at functional lines of block in reentrant circuits during ventricular tachycardia in the infarcted canine heart: a new method to localize reentrant circuits from electrogram features using adaptive template matching. *J. Cardiovasc. Electrophysiol.* 10:194–213, 1999.

[13]Ciaccio, E. J., M. M. Scheinman, and A. L. Wit. Relationship of specific electrogram characteristics during sinus rhythm and ventricular pacing determined by adaptive template matching to the location of functional reentrant circuits that cause ventricular tachycardia in the infarcted canine heart. *J. Cardiovasc. Electrophysiol.* 11:446–457, 2000.

[14]Ciaccio, E. J., A. L. Wit, M. M. Scheinman, S. M. Dunn, M. Akay, and J. Coromilas. Prediction of the location and time of spontaneous termination of reentrant ventricular tachycardia for radiofrequency catheter ablation therapy. *J. Electrocardiol.* 28(Suppl):165–173, 1995.

[15]Dasey, T. J., and E. Micheli-Tzanakou. Detection of multiple sclerosis with visual evoked potentials – an unsupervised computational intelligence system. *IEEE Trans. Inform. Technol. Biomed.* 4:216–224, 2000.

[16]Ferrara, E. R., and B. Widrow. Fetal electrocardiogram enhancement by time-sequenced adaptive filtering. *IEEE Trans. Biomed. Eng.* 29:458–460, 1982.

[17]Harth, E., and E. Tzanakou. Alopex – stochastic method for determining visual receptive-fields. *Vision Res.* 14:1475–1482, 1974.

[18]Melissaratos, L., and E. Micheli-Tzanakou. A parallel implementation of the ALOPEX process. *J. Med. Syst.* 13:243–252, 1989.

[19]Micheli-Tzanakou, E. Neural networks and optimization techniques in neural waveform analysis. *J. Med. Life Sci. Eng.* 13:58–67, 1994.

[20]Micheli-Tzanakou, E., H. Sheikh, and B. Zhu. Neural networks and blood cell identification. *J. Med. Syst.* 21:201–210, 1997.

[21]Tzanakou, E., R. Michalak, and E. Harth. ALOPEX process – visual receptive-fields by response feedback. *Biol. Cybernet.* 35:161–174, 1979.

[22]Widrow, B. Rubber-mask technique 1. Pattern measurement and analysis. *Pattern Recog.* 5:175–197, 1973.

[23]Widrow, B. Rubber-mask technique 2. Pattern storage and recognition. *Pattern Recog.* 5:199–211, 1973.

[24]Widrow, B. Thinking about thinking: the discovery of the LMS algorithm. *IEEE Signal Process. Mag.* 22:100–106, 2005.

[25]Widrow, B., J. R. Glover Jr, J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, E. Dong Jr, and R. C. Goodlin. Adaptive noise cancelling. Principles and applications. *Proc. IEEE* 63:1692–1716, 1975.

[26]Widrow, B., and M. Kamenetsky. Statistical efficiency of adaptive algorithms. *Neural Networks* 16:735–744, 2003.

[27]Widrow, B., and J. M. McCool. Comparison of adaptive algorithms based on methods of steepest descent and random search. *IEEE Trans. Anten. Prop.* 24:615–637, 1976.

[28]Widrow, B., J. M. McCool, M. G. Larimore, and C. R. Johnson Jr. Stationary and nonstationary learning characteristics of the LMS adaptive filter. *Proc. IEEE* 64:1151–1162, 1976.

[29]Widrow, B., and S. D. Stearns. Adaptive Signal Processing. Upper Saddle River, NJ: Prentice Hall, 1985.

[30]Yelderman, M., B. Widrow, J. M. Cioffi, E. Hesler, and J. A. Leddy. ECG enhancement by adaptive cancellation of electrosurgical interference. *IEEE Trans. Biomed. Eng.* 30:392–398, 1983.

[31]Zarzoso, V., and A. K. Nandi. Noninvasive fetal electrocardiogram extraction: blind separation versus adaptive noise cancellation. *IEEE Trans. Biomed. Eng.* 48:12–18, 2001.