Some Algorithms for the Attenuation of Broadband Noise

Michael Arciero

Department of Mathematical Sciences
United States Military Academy
West Point, New York

ARL Advisor: Thyagaraju Damarla

SEDD/Acoustics
U.S. Army Research Laboratory
Adelphi, Maryland

ABSTRACT:

We investigate several algorithms in the attenuation of broadband noise in both vehicle-mounted and ground arrays in an effort to increase target detection range. Many adaptive filtering techniques have been developed for use in noise cancellation applications when the noise statistics are unknown and non-stationary; or changing over time. We adapt variants of the least-mean-square (LMS) adaptive prediction algorithm to simulations as well as ARL test data. The computationally efficient block LMS algorithm is less sensitive to signal transients due to the use of averaging, and is therefore more stable. The normalized LMS algorithm gives more rapid convergence in highly non-stationary situations.

KEY WORDS: Noise cancellation, Least Mean Square algorithm.

CONTACT: Michael Arciero, USMA, West Point, NY 10996
Tel: (845) 938-2453 email: am9272@usma.edu

Thyagaraju Damarla, ARL, Adelphi, MD, 20783
Tel: (301)-394 1266 email: damarla@arl.mil

INTRODUCTION

One of a class of adaptive filtering algorithms, the least-mean-square (LMS) algorithm, introduced by Widrow and Hoff [1], has been widely used in noise cancellation and other signal processing applications due to its computational simplicity and robustness to changing signal statistics. A typical adaptive noise cancellation scheme is depicted in Figure 1. Assume that a discrete-time signal of interest, $y(n)$, is corrupted by a noise process, $z(n)$. The "primary" signal at the filter input is $y(n)+z(n)$. A noise "reference" signal, $x(n)$, consisting of some version of the noise process (represented here as a transfer function H) is then filtered and subtracted from the primary. The filter W attempts to match the primary signal by minimizing the filter output, $\varepsilon(n)$, in the mean-square sense. The justification for this will be seen shortly.

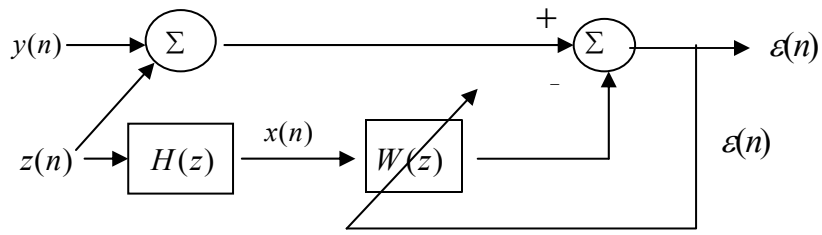


Fig. 1. Adaptive noise cancellation scheme.

If the filter, W, is implemented as a linear finite-impulse response transversal filter with L+1 tap weights, the problem can be viewed as that of predicting the primary signal on the basis of L+1 past samples of the reference. The prediction error is given by

$$\begin{aligned}\varepsilon(n) &= y(n) + z(n) - \sum_{k=0}^{m} w(k)\,x(n-k) \\ &= y(n) + z(n) - \mathbf{w}(n)\,\mathbf{x}(n)', \end{aligned} \qquad (0.1)$$

where $\mathbf{w}(n) = [w(0), w(1), w(2),..., w(L)]$ is the vector of filter weights, and $\mathbf{x}(n) = [x(n), x(n-1), x(n-2),..., x(n-L)]$. The n-dependence of the vector $\mathbf{w}(n)$ reflects the fact that the filter weights are adjusted with each input sample. Thus the filter "adapts" to changing or unknown noise statistics. The problem then is to judiciously choose the filter weights.

If the noise process, $z(n)$, and the signal of interest $y(n)$ are uncorrelated, then the prediction error power is the mean-squared error given by the expected value

$$E(\varepsilon(n)^2) = E(y(n)^2) + E(z(n) - \mathbf{w}(n)\mathbf{x}(n)')^2. \qquad (0.2)$$

Thus the best predictor of the primary $y(n) + z(n)$ is attained precisely when the noise power is minimized. As is well-known (see, e.g., [2]) if the noise process is stationary, the filter weights which minimize (1) are given by the Wiener-Hopf equations:

$$\mathbf{w}_{opt} = \mathbf{R}_{xx}^{-1}\mathbf{P}_{xz}\ , \qquad (0.3)$$

where $[\mathbf{R}_{xx}]_{i,j} = [E(x(n+i)x(n+j)]$ is the matrix of noise covariances, and $\mathbf{P}_{xz} = E([y(n)x(n), y(n)x(n\text{-}1), y(n)x(n\text{-}2),..., y(n)x(n\text{-}L)])$ is the vector of cross-correlations.

In practice, both the signal of interest and the noise process are usually non-stationary with unknown statistics which must be estimated from the data samples. The basis of the LMS algorithm is the replacement of the unknown ensemble average in (1) with the sample

$$\varepsilon(n)^2 = y(n)^2 + (z(n) - \mathbf{w}(n)\mathbf{x}(n)')^2. \qquad (0.4)$$

For any n, $\varepsilon(n)^2$ is a quadratic function of the filter weights, with a unique minimum. Furthermore, the gradient, $\nabla(n) := grad(\varepsilon(n)^2)$, is a vector in w-space which points in the

direction of maximum increase of $\varepsilon(n)^2$. Starting with an initial weight vector, $\mathbf{w}(0)$, the LMS algorithm recursively adjusts the weights in the direction of maximum decrease of $\varepsilon(n)^2$:

$$\begin{aligned}\mathbf{w}(n+1) &= \mathbf{w}(n) - \mu\Delta(n) \\ &= \mathbf{w}(n) + 2\mu\varepsilon(n)\mathbf{x}(n)\end{aligned} \tag{0.5}$$

where the step size, μ, is a design parameter which serves to control the magnitude of the correction to $\mathbf{w}(n)$. In the case of stationary statistics, $\mathbf{w}(n)$ will converge to the optimal Wiener filter provided $\mu < 1/\lambda_{\max}$, where $\lambda_{\max}$ is the maximum eigenvalue of $\mathbf{R}_{xx}$, whose eigenvalues are all positive [2], [3]. In the non-stationary case, the quadratic error surface $\varepsilon(n)^2$ evolves over time, and the algorithm continuously seeks the changing minimum.

The use of the LMS algorithm in attenuating broad band vehicle self-noise in roof-mounted listening arrays was discussed in [4]. In that setting, the noise consists of narrow band "tones" created by the cylinders of the diesel engine, as well as broad band components. A reference of the noise can be obtained from acoustic sensors in the engine compartment. The LMS algorithm performs well in attenuating the high-power tones generated by the engine diesel engine, but performs poorly on the broad band components. As noted in [4], one reason for this is that the gradient, $\nabla(n)$, increases most rapidly in the direction of the eigenvector corresponding to large eigenvalues, which correspond to large noise components. The tones of the diesel engines tested use on the battlefield are typically 30-40 dB stronger than the broad band components. A practical limitation is that it is difficult to obtain a noise reference which is highly correlated the noise in the primary. We see from (0.2) that if $x(n)$ and $z(n)$ are uncorrelated, the filter "shuts off", as all the filter weights are zero.

In [4] a non-LMS based approach was proposed for attenuating broad band noise components. In this paper, we apply variations of the LMS algorithm to test data obtained from both vehicle-mounted and ground-based listening arrays.

ADAPTIVE PREDICTION FILTERS

One application of the LMS algorithm proposed by Widrow, et al. [5] in the context of noise cancellation is that of a prediction filter. These "adaptive prediction filters" can be effective when the noise de-correlates over time much more rapidly than the signal of interest [6]. As is well known, this is true when the noise process has a much wider band width than the signal of interest. For example, a single complex sinusoid $x(n) = e^{in\omega}$ has covariance sequence $c(k) = E(x(n)x(n-k)) = e^{ik\omega}$, which has unit magnitude for all k, while white noise samples separated by one or more lags are uncorrelated. Thus, all covariances are zero except the central one. It would seem likely that this implementation of the LMS algorithm would be effective in canceling the broad-band components of vehicle self-noise. The narrow band components could then be attenuated using the LMS with noise reference as described above.

Consider a narrowband signal, $y(n)$, corrupted by wideband additive noise, $z(n)$. The adaptive prediction filter scheme depicted in Figure 2, a modification of that in Figure. 1, uses a delayed version of the primary signal, depicted as the Δ -block in Fig. 2, as a reference. A "bulk delay" parameter, Δ, is chosen large enough that the $z(n)$ is de-correlated with the delayed signal, $x(n-\Delta) = y(n-\Delta) + z(n-\Delta)$. As in the standard LMS, samples $x(n-\Delta)$ are then used to "predict" the primary, with the error given in (0.1).

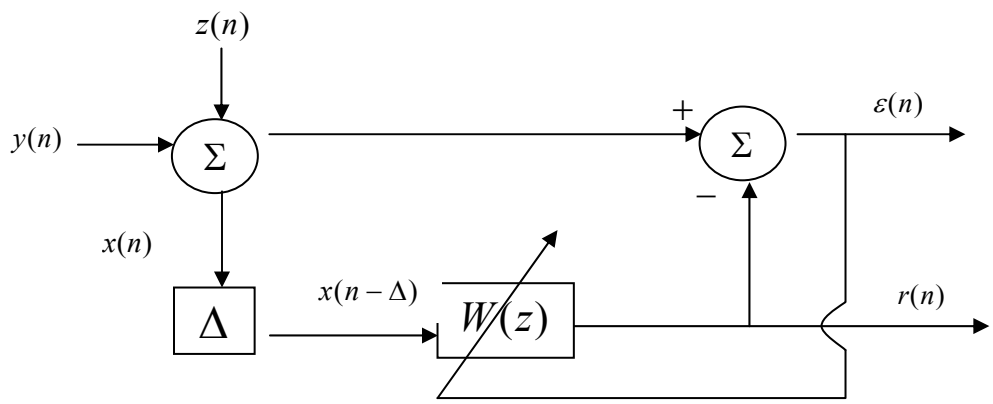


Figure 2. Adaptive prediction scheme

The mean squared error (0.2), and the squared error (0.4) become, respectively,

$$E(\varepsilon(n)^2) = E(z(n)^2) + E(y(n) + \mathbf{w}(n)\mathbf{x}(n-\Delta)')^2 \tag{1.1}$$

$$\varepsilon(n)^2 = z(n)^2 + (y(n) + \mathbf{w}(n)\mathbf{x}(n-\Delta)')^2 \ . \tag{1.2}$$

Thus the filter is optimized when the output, $r(k)$, most closely matches the narrowband signal of interest, $y(n)$. The filter weight update recursion (0.5) becomes

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu\varepsilon(n)\mathbf{x}(n-\Delta) \ , \tag{1.3}$$

where

$$\varepsilon(n) = x(n) - \mathbf{w}(n)\mathbf{x}(n-\Delta)' \ . \tag{1.4}$$

We will consider two variations of the adaptive prediction filter applied to test data obtained from sensors mounted on the HMMV and M113 vehicles, as well as ground-based array sensors. First, we will apply the filter to simulated data and address the performance.

FILTER PERFORMANCE

Let $x(n)$ consist of a sinusoidal signal in white Gaussian noise,

$$x(n) = \Sigma_{i=1}^{m} A_i \sin(\omega_i + \varphi_i) + z(n), \tag{2.1}$$

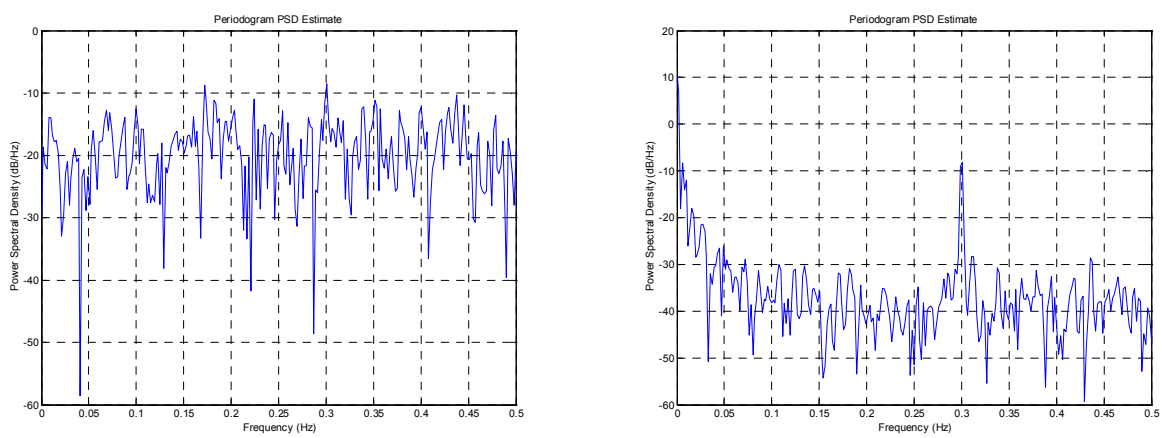
where A_i are amplitudes, φ_i are independent uniform random phases, ω_i are normalized frequencies in cycles/sample, and $z(n)$ is the noise process. The assumption on the random phases is necessary to ensure that $x(n)$ is a stationary process. Let $P_I(\omega)$ and $P_O(\omega)$ denote the power spectral densities of the input and output to the filter, respectively. First consider the situation with one sinusoid corresponding to $m = 1$ in (2.1). To evaluate performance, we will use the following expression [6] for the filter gain:

$$G(\omega) = \frac{P_O(\omega_1)P_I(\omega)}{P_O(\omega)P_I(\omega_1)}; \qquad |\omega - \omega_1| >> \frac{2\pi}{L} \tag{2.2}$$

Expressed in dB, (2.2) gives a ratio of signal-to-noise ratios of output to input spectra, and thus give the increase, in dB, of the sinusoid relative to the background white noise. If υ^2 is the variance of the white noise process, the gain, (2.2) can be expressed [6]

$$G(\omega) = 1 + \frac{a^2}{\mu \upsilon^2 L}, \qquad |\omega - \omega_1| >> \frac{2\pi}{L}, \tag{2.3}$$

where $a = \dfrac{\frac{L}{2}\sigma}{1 + \frac{L}{2}\sigma}$ and $\sigma = \dfrac{A^2}{2\upsilon^2}$ is proportional to the input signal-to-noise ratio. Figure 3 shows periodogram estimates of the input and output spectra, respectively, for a simulated sinusoid of the form (2.1) with normalized frequency ω=0.3, $L = 1024$, $\mu = .0015$, Δ=1, and $\sigma = \upsilon^2 = 1$. Note that Δ=1 is sufficient to de-correlate the process $z(n)$.

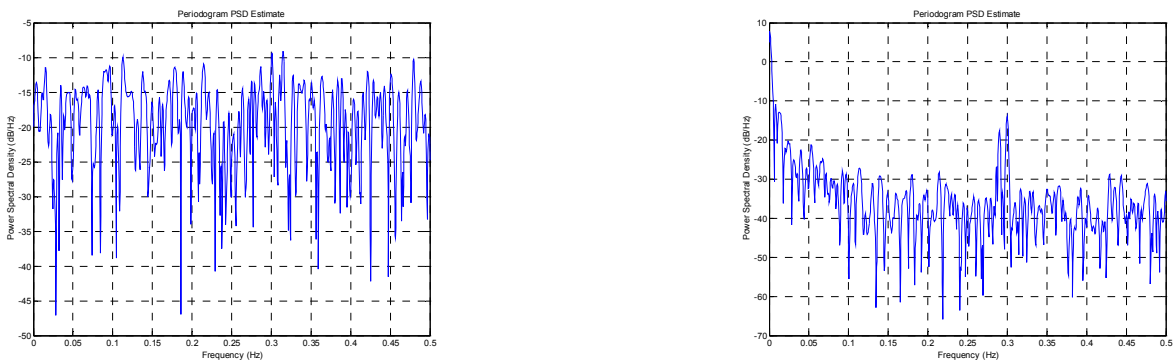


(a) (b)

Figure 3. Filter input, (a), and output, (b), power spectra for sinusoidal signal with ω=.3, L=1024, μ=.0015, and σ=υ^2=1 .

The filter gain appears to be in agreement with the predicted gain $G(\omega) \approx 20$ dB. Figure 4 shows the ability of this same filter to separate closely spaced sinusoids. Again, the filter gain $G(\omega) \approx 20$ dB for each sinusoid. This corresponds to the fact that the spectra (and hence the gain,

$G(\omega)$) for multiple sinusoids may be obtained as a superposition of those corresponding to single sinusoids [7].



(a) (b)

Figure 4. Filter input, (a), and output, (b), power spectra for sinusoidal signal with $\omega_1 = .3$, $\omega_2 = .29$, $L = 1024$, $\mu = .0015$, and $\sigma = \upsilon^2 = 1$.

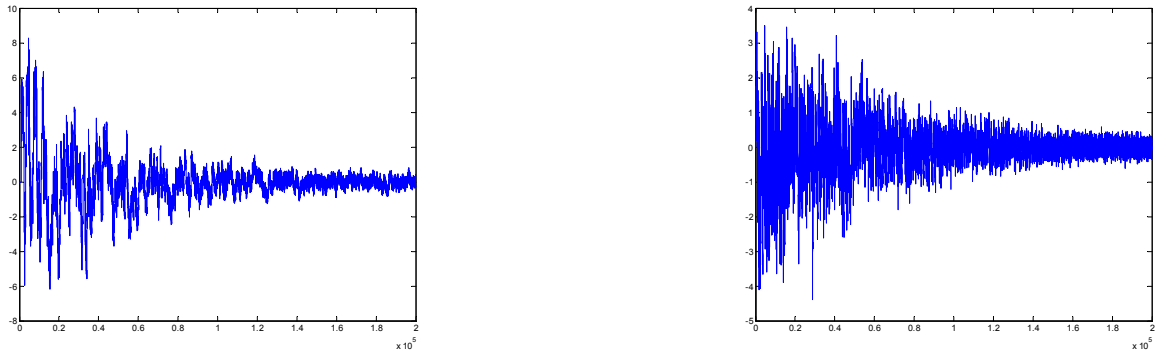
Filter Parameters

Note that (1.2) holds only if the noise correlation time (in samples) is less than the bulk delay parameter, In order that the filter be effective, Δ should be chosen less than the signal correlation time, in samples, otherwise the filter weights are zero, as observed in the previous section. For the simulations above, the choice of Δ is not critical; any value greater than zero will suffice. The optimal filter order is given by [6]

$$L_{opt} = f_s / \alpha \,, \qquad (2.4)$$

where f_s is the sampling frequency and α is the signal bandwidth. In the case of sinusoids, the bandwidth is zero and the choice of $L = 1024$ above is somewhat arbitrary. Indeed, similar reductions in broadband noise were obtained with L ranging from 256 to 2048, as long as μ was chosen sufficiently small, in accordance with (2.3). A larger value for L may be necessary to resolve closely spaced sinusoids [6]. Note that as L increases (2.3) holds for values of ω closer to ω_1. The choice of μ is critical, as it controls the convergence and ability of the filter to track changing statistics. To guarantee stability, it suffices that $\mu < 1/trace(\mathbf{R_{zz}})$, for example [2].

Figure 5 shows the error, $\varepsilon(n)$, for the signal depicted in Figure 3 for two sets of values of filter order and step size, with all initial filter weights equal one. We see that the convergence is quite slow in both cases. In practice, the weights can be "trained" initially to the data by running the algorithm repeatedly on the first few seconds of data, using the converged values of the filter weights as new initial weights. As noted below, this effect can also be achieved with the *normalized* LMS algorithm.



(a) (b)

Figure 5. Filter error for one sinusoid in WGN, 200,000 samples, initial filter weights =1, (a) L=1024, μ = .0015 , and (b) L=256, μ = .015, Δ = 1.

NORMALIZED AND BLOCK ALGORITHMS

Two variations of the prediction filter of (0.2) and Figure 2 will be considered. The *normalized* LMS algorithm (NLMS) [8] adjusts the step size parameter with each sample by replacing μ in (0.5) with $\mu(n) = \dfrac{\rho}{\mathbf{x}(n)\mathbf{x}(n)^T}$. Thus (1.3) becomes

$$\mu(n) = \frac{\rho}{\mathbf{x}(n-\Delta)\mathbf{x}(n-\Delta)^T} \quad . \tag{3.1}$$

The recursion (0.5) for the filter weights then becomes

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\rho}{\mathbf{x}(n-\Delta)\mathbf{x}(n-\Delta)^T}\varepsilon(n)\mathbf{x}(n-\Delta)\,, \tag{3.2}$$

where ρ is a design parameter which controls the convergence rate. Several interpretations for the choice (3.1) are given in [8]. For ρ =1, (3.1) minimizes the *a-posteriori* error, $\varepsilon^+(n) = x(n) - \mathbf{w}(n+1)\mathbf{x}(n-\Delta)'$. Additionally, since the filter *misadjustment*, defined as a ratio of error to minimum error, $\varepsilon(n)/\varepsilon_{\min}$, is proportional to squared energy of the signal, which suggests choosing as in (3.1). Lastly, it can be shown that (3.2) is proportional to performing the recursion (1.3) repeatedly on each sample until the filter weights converge. These properties would seem to give the NLMS algorithm more rapid convergence in a stationary environment and the ability to adapt quickly to non-stationary environments.

The *block* LMS (BLMS) algorithm [8] adjusts the filter weights after every "block" of data samples, using as an estimate of the gradient in (0.5) and an average of errors (0.1) for the past k samples. The recursion (1.3) becomes, for the kth block of length M ,

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{2\mu\Sigma_{i=0}^{M-1}\varepsilon(kM+i)\mathbf{x}(kM-\Delta+i)}{M} \quad . \qquad (3.3)$$

The chief advantage of the using (3.3) over standard LMS is that the averaging of the errors results in gradient vectors, and thus filter weights, with less variance. Thus, larger step size can be used without the filter becoming unstable. The trade-off is that BLMS will adapt more slowly in highly non-stationary environments.

HMMV and M113 TEST DATA

In this section we show the results of applying the adaptive prediction filter in (1.3) to test data obtained at Aberdeen Proving Ground in September of 2001. Noise cancellation was performed on a signal obtained from a microphone mounted on the roofs of the HMMV and M113 vehicles. The sampling frequency for this test data is 2048 samples/sec. Filter parameters:
$L = 2048,\ \mu = .0015,\ \Delta = 25.$

HMMV at Idle.

Figure 6 (a) shows power spectra from filter input and output from the HMMV at idle. This situation is approximately stationary. The spectra are taken from .5 seconds of data several seconds into the filtering process. We see from the filter error in Figure 6 (b) that the filter error stabilized after approximately 1.5 sec. The error is initially zero due to the fact that filter uses samples delayed by $L+1+\Delta$=21074 lags to predict the current sample. We see that the tone near 50 Hz=0.02 cycles/sample is 25-30 dB greater than the surrounding broad band noise "pedestal" at the filter input. We see an apparent broad band attenuation $G(\omega) \approx 40$ dB in the filtered signal.

The 50 Hz tone is created by the HMMV diesel engine, and can subsequently be cancelled via conventional LMS [4]. The resolution exhibited in Figure 4 (b) enables the filter to resolve vehicle self-generated tones and those created by a similar target vehicle.

The stationarity of the signal is apparent from the time-frequency plots of filter input and output in Figures 6 (c) and (d).

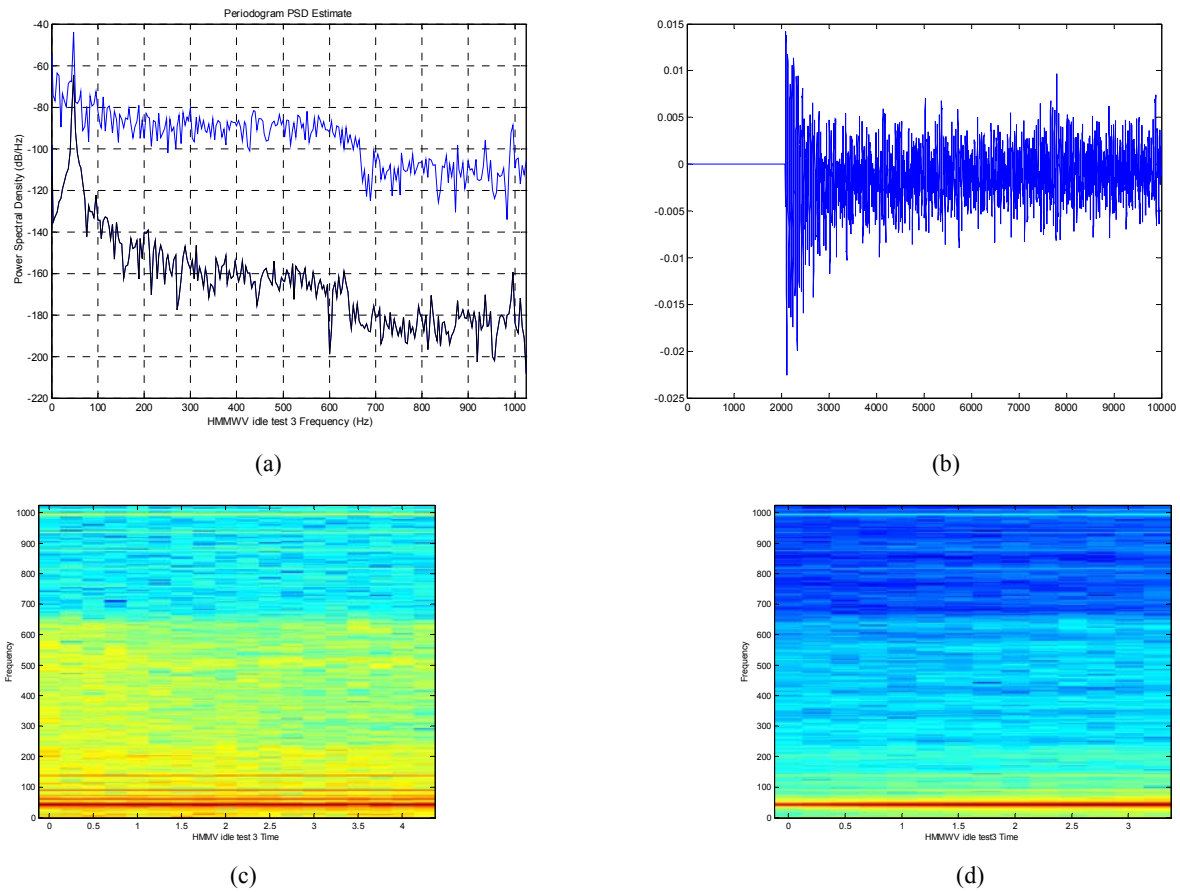
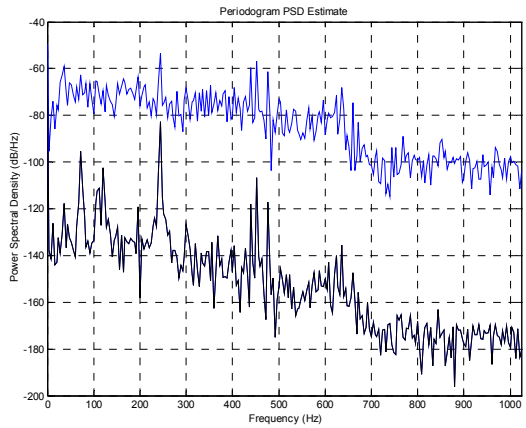


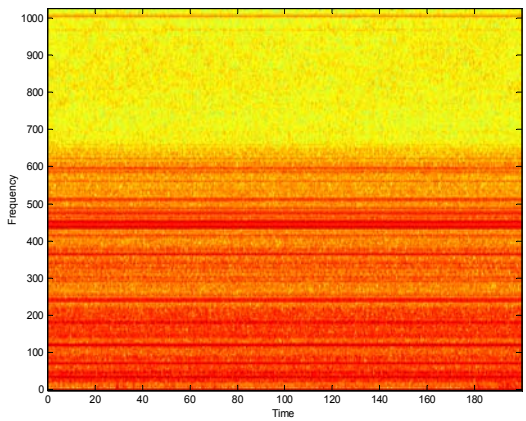
Figure 6. (a) HHMV roof mic and filter output. (b) Filter error for 0.5 seconds of data. (c) Spectrogram of 5 seconds of input. (d) Spectrogram of 5 seconds of filter output. L=1024 μ = .0015, Δ = 25.

M113 at Idle

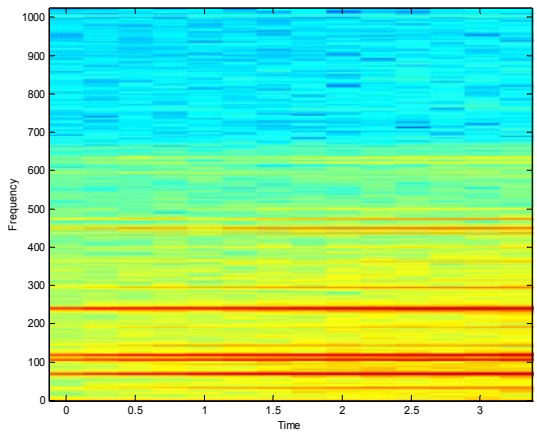
Figure 7 (a) shows power spectra from filter input and output from the M113 at idle. Again, as is apparent from the time-frequency plot of the filter input, Figure 7 (b), this situation is approximately stationary. The spectra are taken from 0.5 seconds of data several seconds into the filtering process. The M113 diesel engine has several tonal components. Broadband attenuation is again apparent from Figures 7 (b) and 7 (c).



(a)



(b)



(c)

Figure 6. (a) M113 roof mic and filter output. (b) Spectrogram of 180 seconds of input. (c)(c) Spectrogram of 4 seconds of filter output. L=1024 μ = .0015, Δ = 25.

REFERENCES

[1] Widrow, B.,Hoff, M.E. *Adaptive Switching Circuits*, IRE WESCON Conv. Rec.,part 4 pp. 96-104, 1960.

[2] Widrow, B., Stearns, S. *Adaptive Signal Processing*, Prentice Hall, NJ., 1985.

[3] Kay, S. *Modern Spectral Estimation*, Prentice Hall, NJ. 1988.

[4] Arciero, M., "An approach to attenuation of broadband platform noise" Proceedings, Tenth Annual ARL and USMA Technical Symposium, Aberdeen Proving Ground, Maryland, November 2002.

[5] Widrow, B, Glover, J., McCool, J., Kaunitz, J., Williams, C., Hearn, J., Zeidler, R., Dong, E. Jr., Goodin, R., "Adaptive noise cancellation: principles and applications", Proc. IEEE, Vol 63, pp. 1692-1716, Dec. 1975.

[6] Zeidler, R., "Performance analysis of LMS adaptive prediction filters", Proceedings of the IEEE, Vol. 78, no. 12, Dec. 1990, pp. 1781-1806.

[7] Zeidler, R., Satorius, E., Chabries, D., and Wexler, H. "Adaptive enhancement of multiple sinusoids in uncorrelated noise", IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-26, No. 3, June, 1978.

[8] Fahrang-Borougeny, B., *Adaptive Filters: Theory and Applications,* Wiley, New York, 1998.