Enhanced Blind Adaptive Narrowband Interference Suppression in DSSS

Habib Fathallah and Leslie A. Rusch

(fhabib@gel.ulaval.ca, rusch@gel.ulaval.ca)

Department of Electrical and Computer Engineering
Laval University, Québec, Canada G1K 7P4
(418) 656-2906, (418) 656-3159

Abstract - This paper addresses blind adaptive suppression of a digital narrowband interferer in direct sequence spread spectrum (DSSS) communications. We apply an adaptive suppression method for wideband interference proposed by Honig, Madhow and Verdú [1] to a narrowband interferer. We identify the eigenspaces of the system dynamics to: 1) significantly decrease convergence times via a new constraint on step size, 2) introduce a simple parameterization of the mean output energy and signal-to-interference ratio, and 3) identify modes of operation where the algorithm ceases to effectively cancel interference. We propose a new adaptive receiver that avoids the convergence anomalies, while capitalizing on the new step size for faster convergence.

1. Introduction

Due to the scarcity of frequency allocations, it has been proposed that spread spectrum (SS) be used for emerging personal communications services, overlaying this signal on existing frequency band occupants [2]. Recent research has investigated the application of multiuser detection theory [3,4,5] to a system with one true SS user to suppress the interference caused by a digital narrowband user. We adopt the model of [3] where each interfering bit is interpreted as a virtual SS user, forming a virtual code division multiple access (CDMA) system with $m+1$ users ($m$ narrowband bits per SS bit). Multiuser detection theory [4] gives us an optimal and many suboptimal receivers requiring information on some, or perhaps all, of the following: signature sequences, received energies, timing delays, etc. We focus on the blind multiuser detection algorithm proposed by Honig, Madhow and Verdú [1], which only requires knowledge of the spreading code of the desired user.

We first consider three fixed receivers. The matched filter or conventional receiver while suboptimal, only requires knowledge of the SS user’s spreading code. The second receiver is the decorrelating receiver, a linear receiver with asymptotic edge of the SS user’s spreading code. (We feature square waves at baseband, therefore we will have only partial narrowband bits at the beginning and end of the SS spreading code. As in [3] we assume the SS spreading code. The narrowband signal and the SS signal are asynchronous, therefore we will have only partial narrowband bits at the beginning and end of the SS spreading code. As in [3] we assume a CDMA system with $m+1$ virtual signature sequences, and the SS spreading code. (We feature square waves at baseband, however our remarks hold for arbitrary waveforms and carrier frequencies that are offset.) The first virtual user's signature sequence is constant during the first narrowband user's bit interval and zero everywhere else. Similarly each narrowband user's bit can be thought of as a signal arising from a virtual user with a signature sequence with only one non-zero interval. These form a set of orthogonal users, uncorrelated with one another. However, in general, the $i^{th}$ virtual user's signature sequence $s_i(t)$, taken to have unit energy, will have some cross correlation, $\rho_{ij}$, for $i$ from 1 to $m+1$, with the spread spectrum user: $\rho_{ij} = \int_0^T s_i(t)s_j(t)dt$ forming the vector $\rho$.

We assume that the received signal strength for both signals remains constant for the SS bit interval. Let $w_j$ be the received energy of the narrowband signal, and $w_0$ the received energy of
the SS user (including the processing gain). We will use the notation that the narrowband user data bits during the interval \((0,T)\) are \(b_1, \ldots, b_{m+1}\) or \(b\), and the SS bit is \(b_0\). The received signal during one bit interval of the SS user, \(t \in [0,T]\), is thus
\[
y(t) = \sqrt{w_0}b_0s_0(t) + \sqrt{w_i} \sum_{i=1}^{m+1} b_is_i(t) + \sigma n(t) \tag{1}
\]
Whereas a one-shot examination of the SS user is suboptimal, the complexity of the receiver is greatly reduced.

2. Subspace Approach for fixed Receivers
2.1 Various Bases & Subspaces
We assume we sample the received signal \(N\) times in a SS bit interval and write the signature sequences as finite dimensional vectors. We define three sets of basis vectors for \(\Gamma\), the space (dimension \(N\)) spanned by all possible signature sequence vectors. The space spanned by all active signature sequences we call \(\Gamma_{\text{act}}\). For our virtual CDMA system there are \(m+2\) active users, therefore \(\Gamma_{\text{act}}\) has dimension \(m+2\). The orthogonal complement of \(\Gamma_{\text{act}}\) we call \(\Gamma_{\text{NO}}\) as it contains energy from noise only, with no energy from an active user.

The set of standard basis vectors, \(e_i\), we call \(B_c\). A second set of basis vectors, \(B_s\), for \(\Gamma_{\text{act}}\) is formed by the signature sequences themselves. These vectors span the space \(\Gamma_{\text{act}}\), have unit length, but are not orthogonal. The third set of basis vectors \(B_{ss}\) are formed from the eigenvectors of the matrix \(R_{ss}\), defined in section 3.1.

2.1.1 Covariance Matrix & Eigenspaces
In order to eventually generate the basis \(B_{ss}\) for \(\Gamma\), we examine the eigenvectors of the covariance matrix \(R_{ss}\) \((N\times N)\) of the received signal. Using the standard basis for the discrete-time (sampled) version of equation (1), we have the following expression for \(R_{ss}\)
\[
R_{ss} = E\{y y^T\} = w_0s_0s_0^T + w_i \sum_{i=1}^{m+1} s_is_i^T + \sigma^2I_N
\]
Using the basis of signature sequences we can write the components of \(R_{ss}\) in \(\Gamma_{\text{act}}\) as
\[
\begin{bmatrix}
R_{ss}
\end{bmatrix}_{\text{act}} = \begin{bmatrix}
w_0 + \sigma^2 & w_i \rho^T \\
w_i \rho & (w_i + \sigma^2)I_{m+1}
\end{bmatrix}_{B_i}
\]
We define
\[
\Delta^+ = \frac{1}{2}(1 - \alpha^2 \pm \sqrt{(1 - \alpha^2)^2 + 4 \alpha^2 \rho^T \rho}) \quad \alpha = \frac{w_i}{w_0}
\]
\[
\gamma_s = \left(\Delta^+ - \rho \rho^T\right)^{1/2} \quad \gamma_i = \left(\Delta^+ + \rho \rho^T\right)^{1/2}
\]
yielding an orthonormal set of eigenvectors with eigenvectors \(V_{0y}\) and \(V_{1y}\) spanning a subspace of dimension two that we call \(\Gamma_E\). It can be shown that \(\Gamma_E\) contains all the energy from the desired user and the effective energy from the interference and noise. We say effective energy, because we will show that the three fixed detectors examined each project the received signal onto this subspace, therefore it is only this subset of interference and noise energy which affects receiver performance.

\[
V_{0y} = \gamma_s \begin{bmatrix}
\Delta^- - \rho^T \\
\rho_1 & \cdots & \rho_{m+1} & \delta_i & \cdots & 0
\end{bmatrix}_{B_i} \quad \forall \ i \leq m+1
\]
\[
V_{1y} = \gamma_i \begin{bmatrix}
0 & \cdots & \rho_1 & \cdots & \rho_{m+1} & \delta_i & \cdots & 0
\end{bmatrix}_{B_i}
\]
\[
V_{by} = \gamma_i \begin{bmatrix}
0 & \cdots & \rho_1 & \cdots & \rho_{m+1} & \delta_i & \cdots & 0
\end{bmatrix}_{B_i}
\]
\[
V_{by} = \gamma_i \begin{bmatrix}
0 & \cdots & \rho_1 & \cdots & \rho_{m+1} & \delta_i & \cdots & 0
\end{bmatrix}_{B_i}
\]

The set of eigenvectors for \(2 \leq i \leq m+1\) are orthogonal to \(s_0\), and span a space we call \(\Gamma_{\text{IN}}\); the orthogonal complement of \(\Gamma_E\) relative to the space \(\Gamma_{\text{act}}\). The remainder of eigenvectors for \(m+2 \leq i \leq N\) span the space \(\Gamma_{\text{NO}}\). The eigenvalues associated with \(\Gamma_{\text{IN}}\) are functions of the interference power and the noise power. To recap, \(\Gamma = \Gamma_{\text{NO}} \oplus \Gamma_{\text{act}} = \Gamma_{\text{NO}} \oplus \Gamma_{\text{IN}} \oplus \Gamma_E\) where \(\oplus\) represents a direct sum.

The canonical form of a receiver is defined in [1] as the sum of the desired user’s spreading code, \(s_0\), plus a vector orthogonal to \(s_0\), i.e., \(c = s_0 + \beta v^\perp\). The vector \(s_0\) can be written as a linear combination of \(V_{0y}\) and \(V_{1y}\), and therefore falls in \(\Gamma_E\). We construct \(v^\perp\) in \(\Gamma_E\), orthogonal to \(s_0\), and together they span \(\Gamma_E\).

\[
V^\perp = \frac{1}{\sqrt{1 - \rho^T \rho}} \begin{bmatrix}
\rho^T & - \rho^T
\end{bmatrix}_{B_i}
\]

As the three fixed receivers we will examine all project the received signal onto \(\Gamma_E\), they can be parameterized in canonical form as \(c = s_0 + \beta v^\perp\).

2.2 Parametrized MOE and SIR
The matched filter receiver has \(c_{\text{MF}} = s_0\). The decorrelating detector for the virtual CDMA system is given by [3]:
\[
c_{\text{DEC}} = \frac{1}{\sqrt{1 - \rho^T \rho}} \begin{bmatrix}
1 & - \rho^T
\end{bmatrix}_{B_i} = s_0 + \frac{\rho^T \rho}{\sqrt{1 - \rho^T \rho}} \cdot v^\perp = s_0 + \beta_{\text{DEC}} \cdot v^\perp
\]
which demonstrates that these two detectors do indeed fall in \(\Gamma_E\). The detector minimizing the mean square error [3,4] is
\[
c_{\text{MMSE}} = \frac{1}{1 - \rho^T \rho + \sigma^2/w_i} \left(\begin{bmatrix}
1 + \sigma^2/w_i
\end{bmatrix} - \rho^T\right)_{B_i}
\]
\[
= s_0 + \frac{\rho^T \rho}{1 - \rho^T \rho + \sigma^2/w_i} \cdot v^\perp = s_0 + \beta_{\text{MMSE}} \cdot v^\perp
\]
In [1], Honig, et al., demonstrated that the receiver which minimizes the MSE also minimizes the mean output energy (MOE), defined as \(\text{MOE}(c) = E\{\langle c, y\rangle^2\}\). We use our subspace partition for an alternate derivation of \(c_{\text{MMSE}}\).

Any two vectors \(c\) and \(y\) can be decomposed into their compo-
ments falling into each of the spaces $\Gamma_E$, $\Gamma_{IN}$, and $\Gamma_{NO}$. For the received signal $y$, only the component in $\Gamma_E$ will contain energy from the desired user. We will minimize the noise and interference energy in the output by choosing the components of $c_{\text{MMSE}}$ in $\Gamma_{IN}$ and $\Gamma_{NO}$ to be zero. Therefore $c_{\text{MMSE}}$ falls in the two-dimensional space $\Gamma_E$ and can be written as $c_{\text{MMSE}} = s_0 + \beta_{\text{min}} v^\perp$.

$$MOE(\beta) = w_0 + \sum_{i} c_i c_i > + \sigma^2 < c_i c_i > - \beta w_j \sqrt{\rho^i (1 - \rho^i)}$$

We differentiate the MOE to find the minimizing value of $\beta$. 

$$\beta_{\text{min}} = \frac{\sqrt{\rho^i (1 - \rho^i)}}{1 - \rho^i + \sigma^2} = \beta_{\text{MMSE}}\left(\frac{1}{w_j}\right)$$

which matches the previous result in (3).

### Mean Output Energy

![Mean Output Energy](image)

**Figure 1 MOE vs. interference power and parameter**

In Figure 1 we plot the MOE as a function of $\beta$ and the interference power, tracing also $\beta_{\text{DEC}}$ and $\beta_{\text{MMSE}}$. The adaptive version of the MMSE receiver minimizes the MOE. As the MOE is convex in $\beta$, the algorithm will not be trapped in local minima. However, the function is much less convex as the interference power approaches the desired user’s power, as is evident in Figure 1 and is nearly flat for very weak interference.

This parameterization, $c = s_0 + \beta v^\perp$, can also be used to facilitate calculation of the signal-to-interference ratio, SIR.

$$SIR(\beta) = \frac{w_0 + \sum_{i} c_i c_i > + \sigma^2 < c_i c_i > - \beta w_j \sqrt{\rho^i (1 - \rho^i)}}{1 - \rho^i + \sigma^2}$$

$\beta_{\text{min}}$ gives a peak in the SIR for very high values of interference power, but exhibits no optimality for weak interferers.

### 3. Subspace Approach for Blind Adaptation

#### 3.1 System Dynamics - Eigenspaces of $R_{xy}$

As mentioned earlier, the receiver that minimizes the MOE also minimizes the MSE. The following gradient descent algorithm was proposed to adaptively minimize the convex function $MOE$.

$$c(n) = c(n-1) - \mu \cdot \langle c(n), y(n) \rangle \cdot (y(n) - \langle s_0, y(n) \rangle \cdot s_0)$$

where $n$ refers to the iteration or bit interval. We define $v(n) = \langle y(n) - s_0 s_0^\perp \rangle$, the product of the Householder matrix and the received signal, yielding

$$c(n) = [I_N - \mu v(n) y^\perp(n)]c(n-1)$$

Taking expectations of both sides we have

$$E[c(n)] = [I_N - \mu E[v(n)y^\perp(n)]E[c(n-1)] = [I_N - \mu R_{xy}]E[c(n-1)]$$

The components of $R_{xy}$ in $\Gamma_{\text{act}}$ can be written in $B$, as

$$[R_{xy}]_{\text{act}} = \begin{bmatrix} -w_j \rho^i \rho & -(w_j + \sigma^2) \rho^i \\ w_j \rho & (w_j + \sigma^2)I_{m+1} \end{bmatrix}$$

Because of the special form of the Householder matrix we would expect the eigenvectors of $R_{xy}$ to follow those of $R_{yy}$, except those eigenvectors falling in $\Gamma_E$, i.e., except in directions containing energy from $s_0$. Indeed our calculations show that the two matrices have the same eigenvectors for $2s \leq N-1$. The first two eigenvectors on the other hand are different.

$$V_{iv} = \frac{1}{\gamma_i} \begin{bmatrix} 1 + \frac{\sigma^2}{w_j} \rho^i \\ 1 - \rho^i \end{bmatrix}, \quad \gamma_i = \sqrt{\left(1 + \frac{\sigma^2}{w_j}\right)^2 + \left(\frac{\sigma^2}{w_j}\right)^2}$$

$$V_{iv} = \frac{1}{\gamma_i} \begin{bmatrix} \rho^i \\ -\rho^i \end{bmatrix}$$

with eigenvalues $\lambda_i = 0$ and $\lambda_i = w_j/\rho^i + \sigma^2$. We make two important observations: the first eigenvector is a multiple of the MMSE receiver in (3), and the second eigenvector is the vector $V^\perp$ defined in (2); together they span the space $\Gamma_E$. Note that these eigenvectors are not orthogonal (because $R_{xy}$ is not symmetric), but that $V_{iv} \perp s_0$.

#### 3.2 Tap Weight Trajectory & New Step Size

Having analyzed the matrix $R_{xy}$, we are now prepared to follow the trajectory of the mean tap weight,

$$t(n) = E[c(n)] = [I_N - \mu R_{xy}]t(n-1)$$

The initial value of the tap weight is taken to be the desired user’s spreading code $s_0$. We will express the matrix $R_{xy}$ in a basis of its eigenvectors $B_{\text{act}}$, to yield a diagonal form. Writing $s_0$ in the basis $B_{\text{act}}$, we arrive at
\[ \tau(n) = [I_n - \mu R_n] s_o \]

\[ = [I_n - \mu R_n] \begin{bmatrix} 1 & \cdots & 0 \\ 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 0 \end{bmatrix} \begin{bmatrix} y'_0 - y'_1 \\ \cdots \\ y'_0 - y'_{N-1} \end{bmatrix} \]

\[ = \frac{1}{1 - \rho^2 r + \sigma^2/w_j} \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y'_0 - y'_1 \\ \cdots \\ y'_0 - y'_{N-1} \end{bmatrix} \]

\[ = \frac{1}{1 - \rho^2 r + \sigma^2/w_j} \begin{bmatrix} y'_0 - y'_1 \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y'_0 - y'_1 \\ \cdots \\ y'_0 - y'_{N-1} \end{bmatrix} \]

\[ = \frac{\gamma'_0}{1 - \rho^2 r + \sigma^2/w_j} V_n - \frac{\gamma'_1}{1 - \rho^2 r + \sigma^2/w_j} V_n \]

The mean tap weight vector falls entirely in the subspace \( \Gamma_E \), therefore we can write it in parameterized form as

\[ \tau(n) = s_o + \frac{\sqrt{\rho^2 r (1 - \rho^2 r)}}{1 - \rho^2 r + \sigma^2/w_j} \left( 1 - \left[ 1 - \mu w_j (1 - \rho^2 r) + \sigma^2 \right] \right) V_n \]

Given \( \lambda_{o_0} = 0 \), the only eigenvalue that effects the stability of the algorithm is \( \lambda_{i_1} \), so that

\[ \mu < \frac{2}{w_j (1 - \rho^2 r) + \sigma^2} \]

ensures convergence of the mean tap vector. As \( n \) tends to infinity (4) approaches the MMSE detector.

### 3.3 Convergence Anomalies & a New Detector

Using our subspace approach to analyze the stochastic gradient algorithm, we can examine new performance tests for the convergence of the algorithm by projecting the mean tap weights onto the three subspaces \( \Gamma_E, \Gamma_{\text{AN}} \) and \( \Gamma_{\text{NO}} \). Ideally the algorithm should only have a non-zero projection onto the space \( \Gamma_E \), and this should asymptotically approach the MMSE detector.

\[ \langle \tau(n), V_n \rangle = \frac{\gamma'_0}{1 - \rho^2 r + \sigma^2/w_j} \left( 1 - \left[ 1 - \mu (w_j (1 - \rho^2 r) + \sigma^2) \right] \right) \]

\[ \rightarrow \frac{\gamma'_0}{1 - \rho^2 r + \sigma^2/w_j} \]

\[ \langle \tau(n), V_n \rangle = \frac{\gamma'_1}{1 - \rho^2 r + \sigma^2/w_j} \left( 1 - \left[ 1 - \mu (w_j (1 - \rho^2 r) + \sigma^2) \right] \right) \]

\[ \rightarrow \frac{\gamma'_1}{1 - \rho^2 r + \sigma^2/w_j} \]

\[ \langle \tau(n), V_n \rangle = 0 \quad \text{for} \quad 2 \leq i \leq N-1 \]

In section 4 we examine these projections for various interference powers and step sizes.

We noted previously that the adaptive algorithm will not be trapped by local minima, but it is easily lead astray by spurious noise samples when the interference power is smaller than the SS power (see Figure 1). Also, the gradient is inherently noisy, so that the longer the algorithm is allowed to run, the more likely an outlier noise sample will occur and force the gradient to follow the noise into the space \( \Gamma_{\text{NO}} \). Once displaced into \( \Gamma_{\text{NO}} \) the algorithm has great difficulty in returning to \( \Gamma_E \). This problem is alluded to in [7], where a step size that varies as the inverse of the iteration number is suggested to effectively stop adaptation after a certain point.

In view of these convergence anomalies, we propose a new detector which takes advantage of the new step size for quicker convergence, while avoiding poor adaptation when the interference power is weak. There are two factors that determine the algorithm’s continuing adaptation: iteration number and interference power. When the interference power is very weak, i.e., undetected, the matched filter is nearly optimal, and a control forces the proposed receiver to the matched filter.

If the interference power is strong, the algorithm has a good chance of converging to the MMSE receiver. Once enough iterations have passed for convergence, the proposed receiver averages over several dozen bits to find the mean value for the tap weights and use this as a fixed receiver (alternately we could use this as the initial state for a decision-directed version of the LMS algorithm). If the interference power degrades significantly the proposed receiver also stops the adaptation, as the likelihood of drifting out of \( \Gamma_E \) and into \( \Gamma_{\text{NO}} \) is quite great.

### 4. Simulation Results

In Figure 2 we present Monte Carlo simulation results vs. the theoretical calculations of the probability of error for three fixed receivers: the matched filter (MF), the decorrelating detector (DEC), and the MMSE detector. All simulations use an \( m \)-sequence of length 63 for the true spread spectrum user, a noise power 6 dB down from the despread CDMA signal (as proposed in [2] during field trials), interference powers that vary from parity with the CDMA signal to 40 dB stronger, and values for \( m \) of 1, 2, 4 and 8.

![Figure 2 Probability of error vs. Near-far ratio: theoretical values as points, simulation values as curves.](image-url)
We see that simulation and theory match well. The decorrelating and MMSE detectors yield very similar performance. Figure 2 also shows results for the stochastic gradient and the new receiver using the old step size. For most plots the proposed receiver achieves the performance of the decorrelating or MMSE detector. The gradient’s performance improves as the interference power increases.

To arrive at reliable estimates of the probability of error it was necessary to run the algorithm for millions of bits. This allowed adequate opportunity for the gradient algorithm to be lead astray by noise and significantly differ from the MMSE detector, even for very strong interference. Contrast this with the proposed detector. This detector ran adaptively at an interference power of 30 dB for 1000 bits. The tap weight values were then averaged over the last 100 iterations. This receiver (now fixed) was used to calculate the probability of error. At the time the gradient algorithm was stopped, it had formed a good estimate of the MMSE detector.

For small interference power, Figure 3c, leads to ideal performance with all energy projected onto $\Gamma_E$. However for small interference power energy is displaced from $\Gamma_E$ into $\Gamma_{I\&N}$, i.e., the gradient algorithm is tracking noise samples and not the desired user. While [1] and [5] alluded to the inefficiency of the gradient and the desirability of switching quickly to a decision-directed LMS detector, we now see that when the interference power is weak it is unlikely that the gradient algorithm will ever produce meaningful output, motivating the proposed receiver.

In Figure 3 the energy of the adaptive receiver was projected onto the partition of $\Gamma$ and plotted versus the iteration number. A strong interference power, Figure 3c, leads to ideal performance with all energy projected onto $\Gamma_E$. However for small interference power energy is displaced from $\Gamma_E$ into $\Gamma_{I\&N}$, i.e., the gradient algorithm is tracking noise samples and not the desired user. While [1] and [5] alluded to the inefficiency of the gradient and the desirability of switching quickly to a decision-directed LMS detector, we now see that when the interference power is weak it is unlikely that the gradient algorithm will ever produce meaningful output, motivating the proposed receiver.

Figure 4 presents curves for theoretical values of the mean tap weight versus Monte Carlo simulations for an interference power of 30 dB and $m = 4$. We project the tap weights as proposed in equation (6) onto the subspaces $\Gamma_E$ and $\Gamma_{I\&N}$. It is clear that the new step size allows convergence in dozens of iterations, versus hundreds of iterations for the old step size. Given that we wish to switch out of the stochastic gradient adaptation as soon as possible, the new step size criteria is an important improvement to the algorithm. Note though that we also see in these plots the additional excess noise caused by the faster convergence.

5. Conclusion

By examining the eigenspaces of the received signal we exploited the special form of the narrowband interference to uncover a two dimensional subspace $\Gamma_E$. This subspace is of key importance as all the signal energy is contained in this subspace, while only portions of the interference and noise energies are present. We demonstrate that the receivers investigated project the received signal onto this subspace. We showed that the adaptive version of the MMSE receiver is, on average, constrained to this subspace. This allowed us to identify less restrictive stability constraints on the adaptation step size. We also explain why for weak interference the adaptive algorithm is susceptible to outlier noise samples which lead the adaptive receiver to leave the plane $\Gamma_E$. Finally, we proposed a new adaptive receiver that avoids the convergence anomalies while capitalizing on the new step size for faster convergence.

6. References