Analysis of a Decision Directed Beamformer

A. Lee Swindlehurst, Member, IEEE, Soumendra Daas, and Jiankan Yang

Abstract—In this paper, we study a technique for using decision direction to extract digital signals from antenna array data. The algorithm alternates between 1) estimating and demodulating the received signals and 2) using the resulting bit decisions to regenerate the signal waveforms and recompute the beamformer weights. An analysis of the (asymptotic) symbol error rate performance of the algorithm for the case of M-ary PSK signals is included, along with several representative simulation examples.

I. INTRODUCTION

T

HE problem of extracting communication signals using an array of sensors is one of increasing importance to modern communications. For example, as the demand for bandwidth and time slots in mobile cellular radio systems increases, the use of multiple antennas at the cellular base station has been considered as a way of providing an extra degree of spatial discrimination beyond the coarse cell structure. Such super-directivity (sometimes referred to as spatial division multiple access, or SDMA) could potentially allow frequency re-use within the same cell, decrease cell-to-cell interference, and provide a significant increase in capacity (e.g., see [1]-[4]). As such systems switch to a digital format, methods for accurate multi-sensor reception of digitally modulated signals will be required.

Most conventional techniques for co-channel signal estimation using antenna arrays (i.e., beamforming) require that the directions of arrival (DOA's) of the signals be determined before the beamformer weights can be computed. On the other hand, a number of so-called blind beamforming algorithms have been developed that exploit the temporal rather than the spatial structure of the signals. These include the various SCORE algorithms [5], [6], the constant modulus approach [7], [8], and other property restoration algorithms [9]. The term property restoration refers to the fact that these algorithms force the signal estimates to have certain structural properties that the actual signals are known to possess (e.g., constant modulus).

Perhaps the most structured of all signals are those that are digitally modulated, where all of the uncertainty in the signal's value at some time is due only to synchronization and which of a finite alphabet of symbols has been transmitted. In this paper, we present a property restoration algorithm for digital signals that exploits a priori knowledge of the signals' modulation format, pulse shape, and baud and carrier frequencies for improved estimation performance. Our approach is decision directed, in that symbol decisions made on a preliminary signal estimate are used to generate a new set of beamformer weights, and an updated signal estimate (see also [10], [11]). This method is related to the LS-CMA algorithm of [12], although not only is the constant modulus property enforced at each step but the pulse shape over each baud based on the symbol decision for that baud as well. Two similar iterative techniques recently presented by Talwar, et al. [13], [14], alternate between using a least squares beamformer and either projecting the resulting signal estimate onto the nearest symbol, or using enumeration to approximate the maximum likelihood solution. A method alternating between least squares estimators of the signal waveform and beamformer weights has also appeared in [10], [17].

In addition to presenting the algorithm and its implementation, we also conduct a symbol error rate analysis of its performance for the case of M-ary PSK signals. After some background material in the following section, the algorithm and its performance are described in Sections III and IV.

II. DATA MODEL AND RELEVANT ALGORITHMS

Consider an array of m sensors having arbitrary positions and characteristics that receives the waveforms of d narrowband (co-channel) signals. The vector of complex sensor outputs is denoted \( \mathbf{x}(t) \), and is modeled by the following familiar equation:

\[
\mathbf{x}(t) = [\mathbf{a}(\theta_1) \mid \cdots \mid \mathbf{a}(\theta_d)] \begin{bmatrix} s_1(t) \\ \vdots \\ s_d(t) \end{bmatrix} + \mathbf{n}(t) \]

The columns of the \( m \times d \) matrix \( \mathbf{a} \) are the so-called steering or propagation vectors of the array, and are denoted as \( \mathbf{a}(\theta_i), i = 1, \cdots, d \). These vectors describe the array response to a unit waveform with parameter(s) \( \theta_i \), which include the DOA of the signal. The \( d \)-vector \( \mathbf{s}(t) \) is composed of the complex waveforms (in-phase and quadrature components) of the signals received at time \( t \), and the \( m \)-vector \( \mathbf{n}(t) \) accounts for additive measurement noise. The noise term is modeled as a zero-mean, stationary, complex random process that is uncorrelated with any of the signals. It is further assumed to be temporally and spatially white

\[
\mathbb{E}\{\mathbf{n}(t)\mathbf{n}^T(s)\} = \sigma_n^2 \mathbf{I}_{m,s} \]

\[
\mathbb{E}\{\mathbf{n}(t)\mathbf{u}^T(s)\} = 0 \]

where \( \mathbb{E}\{\cdot\} \) denotes expectation, and \( \delta_{t,s} \) is the Kronecker delta.
The general problem addressed in this paper is the estimation of one or more of the signal waveforms at \( N \) distinct sample points \( S = [s(1), \ldots, s(N)] \), using the received data \( X \)

\[
X = A(\theta)S + N
\]

where \( X \) and \( N \) are defined similarly to \( S \). This is typically done by forming a linear combination of the array outputs, as in

\[
\hat{S} = W^*X
\]

where \( W = [w_1, \ldots, w_d] \) and \( w_i \) is referred to as the beamformer weight vector for the \( i \)th signal. There are a number of methods available for choosing the weight matrix \( W \), each with a different optimality criterion and a different set of assumptions about what a priori information is available. We mention two here, the least-squares (LS) and minimum mean-squared error (MMSE) approaches.

The LS algorithm finds the signal estimate that, in the LS sense, best matches the received data given an estimate of the steering matrix \( A = A(\theta) \)

\[
\hat{S}_{LS} = \arg \min_S \|X - \hat{A}S\|^2_F
\]

where in this case \( \hat{W}_{LS} = \hat{A}^* = \hat{A}^*(\hat{A}^*\hat{A})^{-1} \). If the noise is temporally white and Gaussian, then it is easy to show [15] that \( \hat{S}_{LS} \) corresponds to the maximum likelihood (ML) estimate of \( S \). On the other hand, the MMSE weight vector is calculated to be

\[
\hat{W}_{MMSE} = \arg \min_W \mathbb{E}\{\|\hat{W}^*X - \hat{S}\|^2_F\} = R_{xx}^{-1}R_{xs}
\]

where

\[
R_{xx} = \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} x(t)x^*(t) = AR_{ss}A^* + \sigma_n^2 I
\]

\[
R_{ss} = \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} s(t)s^*(t)
\]

\[
R_{xs} = \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} x(t)s^*(t) = AR_{xs}
\]

In this case, the weight vector depends on the signals themselves through \( R_{ss} \) or \( R_{xs} \), and thus the MMSE method cannot be implemented directly without knowledge of \( S \). Typically, \( R_{ss} \) is replaced by a sample average, and \( R_{xs} \) is replaced by some other suitable estimate (e.g., see [16]) or is calculated using a known training sequence.

III. A DECISION DIRECTED APPROACH

In the approach considered herein, the column of \( R_{xs} \) corresponding to the signal of interest (SOI) is estimated by assuming that the SOI is digitally modulated, and that an initial (perhaps crude) estimate of the SOI is available for demodulation. The resulting symbol stream is then remodulated to generate a "clean" approximation of the SOI, which can be used as a reference signal in estimating the appropriate column of \( R_{xs} \). An outline of the algorithm is given in the following:

1) Obtain an initial estimate of the SOI by either a conventional DOA-based or blind beamformer. Demodulate the signal to estimate the transmitted symbols, and use the symbol decisions to generate an improved estimate of the transmitted signal. Denote this estimate as \( \hat{s}(t) \).

2) Compute an estimate of the MMSE beamformer weights for the SOI

\[
\hat{w} = R_{xx}^{-1}R_{xs}\hat{s}
\]

3) Compute the signal estimate \( \hat{s}(t) = \hat{w}^*x(t) \)

4) Demodulate \( \hat{s}(t) \) to obtain an estimate of the transmitted symbols. For the \( k \)th iteration, let the estimate of the transmitted symbols be denoted by the vector \( \hat{q}_k \).

5) Using \( q_k \) as the modulating symbol stream, reconstruct a unit amplitude estimate of the transmitted signal. Denote this estimate as \( \hat{s}_k(t) \).

6) Set \( \hat{s}_k(t) = \hat{s}_k(t) \).

7) If needed, repeat Steps 2 to 6 (e.g., until \( q_k = q_{k+1} \), or some other stopping criterion is reached).

This algorithm exploits the fact that given the sequence of (synchronized) transmitted symbols, it is possible to perfectly reconstruct a noise-free replica of the original signal. The estimated symbol stream will of course differ from the original, so we use the reconstructed version as a reference signal in computing an approximation to the column of \( R_{xs} \) associated with the SOI. At iteration \( k \), \( \hat{s}_k(t) \) will differ from the original signal \( s(t) \) only in places corresponding to the symbols that have been demodulated incorrectly. The number of incorrectly demodulated symbols will depend, among other variables, on the quality of the initial estimate, the noise power, and the degree to which symbol synchronization has been accurately carried out. These factors also determine whether or not the algorithm will converge to a reasonable estimate. Before moving on to an analysis of the algorithm's performance, we note the following:

- The number of samples per symbol assumed for the reconstructed signal in Step 5 is arbitrary (provided it is no greater than at the receiver), although if it is more than one, the pulse shape of the transmit filter must be known. In either case, we are implicitly neglecting any dispersive effects of the channel (or at least we are assuming a prior equalization step has occurred to mitigate such effects).

The algorithm can be modified to perform a simultaneous temporal and spatial equalization by simply replacing \( X \) in (8) and (9) with

\[
\hat{X} = \begin{bmatrix} X(1) & \vdots & X(p) \end{bmatrix}
\]

where \( X(i) = [x(i), \ldots, x(N+i-1)] \) and \( p \) is the desired number of (temporal) filter taps. The resulting
weight vector estimate is also partitioned as in (10)

\[
\hat{w} = \begin{bmatrix}
W(1) \\
\vdots \\
W(p)
\end{bmatrix}
\]

and would in this case correspond to a 2-D spatio-temporal filter.

- Given an initial estimate of \( \mathbf{S} \), a variety of options exist for computing the columns of the weight matrix \( \hat{W} \) besides (9). For example, to use the LS signal copy weights in (7), an estimate of the steering matrix \( \hat{A} \) is required. Given the signal waveforms (or an estimate thereof), an estimate of \( \hat{A} \) can be obtained without first finding the DOA’s by minimizing (6) with respect to \( \hat{A} \) (e.g., see [17]):

\[
\hat{A} = \mathbf{X} \mathbf{S}^* (\mathbf{S} \mathbf{S}^*)^{-1}.
\]

Thus, given the initial estimate \( \hat{S}_0 \), Step 2 in the algorithm above could be replaced with the following procedure:

1) Compute the LS signal copy vector \( \hat{W} = \hat{A}^* \), where

\[
\hat{A} = \mathbf{X} \mathbf{S}^* (\mathbf{S} \mathbf{S}^*)^{-1}.
\]

This is the approach taken in the ILSP method of [13] and [14]. However, it will be shown in Section IV-C that even if \( \hat{A} \) is known exactly, this approach will, in general, give rise to a higher symbol error rate than if the MMSE weights are used. Thus, the step shown in (9) is preferred.

- Use of the MMSE weighting in (9) makes the algorithm robust to carrier synchronization errors, relatively small symbol timing errors, or any other type of error that causes a spurious phase shift in the reconstructed signal estimate. For example, if the SOI is phase shifted by \( e^{j\phi} \) relative to the reference carrier at the receiver (the carrier used in forming the reference signal \( s(t) \), and denoting which symbol was transmitted at time \( n \)), the phase term in (9) will be phase shifted by \( e^{-j\phi} \) from its nominal value, and the phase of the updated signal estimate in Step 3 will automatically be aligned with the phase of the carrier at the receiver.

The performance degradation due to more significant symbol synchronization mismatch can be calculated exactly as in the single channel case, and thus will not be explicitly addressed in our analysis. Achieving SOI symbol synchronization in the presence of co-channel interference is not a trivial task, and probably requires that some degree of spatial discrimination be used in obtaining the initial signal estimates. As \( \hat{w} \) is updated at each iteration of the algorithm, the ability to spatially separate the signals will improve, and consequently so will the ability to attain symbol synchronization. Thus, in practice, Steps 3 and 4 above would probably include an additional step where symbol timing is reacquired.

IV. PERFORMANCE ANALYSIS FOR PSK SIGNALS

Assume that the SOI is the \( d \)th element of \( s(t) \), and that it is an \( M \)-ary PSK signal with an arbitrary, unit energy pulse shaping waveform \( p(t) \). After the transmission of \( n \) symbols,

\[
s_d(t) = \sigma_d p(t - nT)c_z(n-1)e^{-j\omega t\phi},
\]

\[
nT \leq t \leq (n+1)T
\]

where \( \sigma_d \) is the (real-valued) amplitude of the SOI (\( \sigma_d^2 \) is the \( d \)th diagonal element of \( \mathbf{R}_{ss} \)), \( q(n) \) is an integer from 1 to \( M \) denoting which symbol was transmitted at time \( n \), \( \omega_c \) is the carrier frequency, \( \phi \) is an arbitrary phase factor, and \( T \) is the symbol period. Our goal in this section will be to determine the symbol error rate (SER) of the decision directed (DD) algorithm described in Section III for PSK signals of the form (12). We will make the following assumptions in our analysis:

1. The number of data samples \( N \) used in the DD algorithm is large enough so that sample averages may be replaced by their limiting values.

2. The symbol sequence \( q(n) \) is white.

3. Each of the elements of the symbol set are equally likely

4. The interfering signals \( s_k(t) \), \( k < d \) may be correlated with \( s_d(t) \) and will thus be decomposed into two parts—one correlated with the SOI and the other not:

\[
s(t) = \begin{bmatrix}
\tilde{s}(t) \\
s_d(t)
\end{bmatrix} = \begin{bmatrix}
s_{-1}(t) \\
0
\end{bmatrix} + \frac{1}{\sigma_d^2} \mathbf{R}_{ss} s_d(t)
\]

where \( \tilde{s}(t) = [s_1(t), \ldots, s_{d-1}(t)]^T \) and \( s_{-1}(t) \) is the part of \( \tilde{s}(t) \) that is uncorrelated with the SOI. The “pure” interference component \( s_{-1}(t) \) will be modeled as a stationary, zero-mean, Gaussian random process.

A. Effect of the Decision Directed Update

Suppose the initial signal estimate is demodulated to obtain an estimate of the symbol sequence \( q(n) \), and let \( \hat{s}_d(t) \) denote the reconstructed signal using this as the modulating symbol stream (we drop the 0 subscript in Step 1 of the DD algorithm for convenience). The beamformer weight vector for the first iteration (Step 2 of the algorithm) is given by

\[
\hat{w}_d = \hat{R}_{dd} \hat{R}_{dd}.
\]

For large \( N \), \( \hat{R}_{xx} \rightarrow \mathbf{R}_{xx} \) and \( \hat{R}_{xx_d} \rightarrow \mathbf{R}_{xx_d} \), where

\[
\mathbf{R}_{xx_d} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^{N} x(t) \tilde{s}_d(t)^* \]

\[
= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^{N} [A s(t) + n(t)] [\hat{s}_d(t) + \tilde{s}_d(t)]^* \]

\[
= A \mathbf{R}_{ss} + A \mathbf{R}_{ss_d} + \mathbf{H}
\]

and

\[
\tilde{s}_d(t) = \hat{s}_d(t) - s_d(t)
\]

\[
\mathbf{H} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^{N} n(t) \tilde{s}_d(t).
\]

2Since the beamforming presumably takes place prior to making any symbol decisions, \( N \) will be larger than the total number of symbol decisions if the array is sampled faster than the symbol rate.

3Although technically this assumption is necessary to derive an analytical expression for the SER, we will see in Section V that the resulting expression can still be quite accurate when the interference is non-Gaussian (e.g., digital communications signals).
With proper symbol synchronization, the error signal will be given by
\[ \tilde{s}_d(t) = \left( \frac{1}{\sigma_d} - 1 \right) s_d(t) \]
when a correct decision is made for the symbol at time \( t \) (since the amplitude of the received SOI is unknown and a unit amplitude reconstructed signal is used, the error is nonzero even for a correct decision). When a symbol error occurs, with high probability it will be because the symbol was associated with an immediately adjacent point on the signal constellation, in which case
\[ \tilde{s}_d(t) = \frac{e^{\pm j2\pi/M}}{\sigma_d} s_d(t) \]
for \( M \)-ary PSK. Thus, if we let \( b \) denote the probability of a symbol error in \( \tilde{s}_d(t) \) and assume that each of the two "possible" demodulation errors is equally likely, we may write
\[ \tilde{s}_d(t) = \delta(t) s_d(t) \]
\[ \delta(t) = \begin{cases} \frac{1}{\sigma_d} - 1 & \text{w.p. } 1 - b \\ \frac{e^{\pm j2\pi/M}}{\sigma_d} - 1 & \text{w.p. } b/2 \end{cases} \]
and together with (14) we have
\[ R_{\tilde{s}d} = \frac{1}{\sigma_d} \left[ 1 + b \cos(2\pi/M) - b - \sigma_d \right] R_{ssd} \]
and
\[ R_{\tilde{s}d} = (1 + \alpha) R_{ssd} \]
where we have defined \( \alpha = [1 + b \cos(2\pi/M) - b - \sigma_d] / \sigma_d \).
Thus, as \( N \to \infty \), the beamformer weights converge to
\[ \tilde{\mathbf{w}} = (1 + \alpha) R_{\tilde{s}d}^{-1} \mathbf{A} R_{ssd} \]
which is just a scaled version of the weight vector that would be obtained if \( s_d(t) \) were known exactly. A real-valued scaling of the weights will have no effect on the SER performance of the algorithm, and thus our analysis implies that for \( N \to \infty \), the SER of the DD algorithm will converge in a single iteration to that of the optimal MMSE beamformer independent of the SER of the initial signal estimate. The simulation examples of Section V demonstrate that this is approximately true even for relatively small values of \( N \).

**B. SER Performance of the MMSE Beamformer**

To determine the asymptotic performance limit of the DD algorithm, we compute in this section the SER of the optimal weight vector \( \mathbf{w}_d = R_{\tilde{s}d}^{-1} R_{ssd} = R_{\til{s}d}^{-1} \mathbf{A} R_{ssd} \). To begin with, we note the following easily proven identity
\[ R_{\tilde{s}d}^{-1} \mathbf{A} R_{ssd} = (I - \sigma_n^2 R_{xx}^{-1}) (A_{1*})_d \]
where \( (A_{1*})_d \) denotes the \( d \)th column of \( A_{1*} \). Together with (13), this implies that the signal estimate obtained using the optimal MMSE weights may be written as
\[ \hat{s}_d = \beta s_d(t) + \tilde{n}(t) \]
where
\[ \beta = \frac{1}{\sigma_d^2} [A_1 (I - \sigma_n^2 R_{xx}^{-1}) \mathbf{A} R_{ssd}]_{dd} \]
\[ = 1 - \frac{\sigma_n^2}{\sigma_d^2} \gamma_d \]
\[ \gamma_d = [A_1 (I - \sigma_n^2 R_{xx}^{-1}) A_{1*}]_{dd} \]
\[ = [(A^* \mathbf{A} + \sigma_n^2 R_{xx}^{-1})^{-1}]_{dd} \]
\[ \tilde{n}(t) = (A_{1*})_d (I - \sigma_n^2 R_{xx}^{-1}) (A_{1*})_d(t) + n(t) \]
\[ \mathbf{A} = [a(\theta_1) \cdots a(\theta_{d-1})] \]
and \([.]_{dd}\) denotes the \( d,d \)th element of its matrix argument.
The SOI estimate is thus composed of a scaled version of the SOI plus a zero-mean white Gaussian "noise plus interference" sequence. The procedure for computing the SER of this type of signal is standard (assuming an "optimum" matched filter correlator structure), and can be found in a number of texts (e.g., [18]). The resulting SER expression depends on the ratio of the power of the signal part \((\beta^2 \sigma_n^2)\) to that of the noise and interference:
\[ \mathcal{E}[\tilde{n}(t)]^2 = \mathcal{E} \left[ (A_{1*})_d (I - \sigma_n^2 R_{xx}^{-1}) \right] \]
\[ \times \left[ \mathbf{A} s(t) - \frac{1}{\sigma_d^2} \mathbf{A} R_{ssd} s_d(t) + n(t) \right]^2 \]
\[ = (A_{1*})_d^2 (I - \sigma_n^2 R_{xx}^{-1}) \left[ R_{xx} - \frac{1}{\sigma_d^2} \mathbf{A} R_{ssd} R_{ssd}^* \mathbf{A}^* \right] \]
\[ \times (I - \sigma_n^2 R_{xx}^{-1}) (A_{1*})_d \]
\[ = [A^*(\mathbf{A} R_{ssd}^* - \sigma_n^2 I + \sigma_n^2 R_{xx}^{-1} A_{1*})]_{dd} - \sigma_n^2 \gamma_d^2 \]
\[ \sigma_n^2 \gamma_d \left[ 1 - \frac{\sigma_n^2}{\sigma_d^2} \gamma_d \right] \]
\[ \gamma_d \left[ 1 - \frac{\sigma_n^2}{\sigma_d^2} \right]. \]
If we let \( \text{SNR}_{id} = \sigma_n^2 / \sigma_d^2 \) denote the "input" signal-to-noise ratio, then the "output" SNR can be obtained using (19)–(22) and (25)–(30):
\[ \text{SNR}_{od} = \frac{\beta^2 \sigma_d^2}{\mathcal{E}[\tilde{n}(t)]^2} \]
\[ = \frac{\sigma_n^2 (1 - \text{SNR}_{id}^{-1} \gamma_d)}{\sigma_d^2 \gamma_d (1 - \text{SNR}_{id}^{-1} \gamma_d)} \]
\[ = \frac{\text{SNR}_{id}}{\gamma_d} - 1. \]
Let $N_T$ represent the number of samples per symbol, and denote the complementary error function as
\[
\Phi(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} e^{-t^2} dt.
\]
Using these definitions and the standard approach of [18], we find the SER of the MMSE beamformer to be
\[
\begin{align*}
\text{SER}_{\text{MMSE}} &= \frac{1}{2} \Phi \left( \sqrt{\frac{N_T \times \text{SNR}_{\text{od}}}{2}} \right) \\
&= \frac{1}{2} \Phi \left( \sqrt{\frac{\{(A^* A + \sigma_n^2 R_{\text{xx}})^{-1}\}_{dd}}{2}} \right) \\
&= \frac{1}{2} \Phi \left( \sqrt{\frac{N_T \times \text{SNR}_{\text{od}}}{2}} \right) \\
&= \frac{1}{4} \Phi \left( \sqrt{\frac{N_T \times \text{SNR}_{\text{od}}}{2}} \right)
\end{align*}
\]
• $M = 2$, BPSK
\[
P_{\text{dd}} = \Phi \left( \sqrt{\frac{N_T \times \text{SNR}_{\text{od}}}{2}} \right)
\]
• $M = 4$, QPSK
\[
P_{\text{dd}} = \Phi \left( \sqrt{\frac{N_T \times \text{SNR}_{\text{od}}}{2}} \right)
\]
• $M > 4$, $\text{SNR}_{\text{od}} \gg 1$
\[
P_{\text{dd}} = \Phi \left( \sqrt{\frac{N_T \times \text{SNR}_{\text{od}}}{2}} \right)
\]
where, as in (34), (21)-(22) and (32) can be used to express $\text{SNR}_{\text{od}}$ in terms of "physical" variables.

Special Cases: It is instructive to examine the behavior of the MMSE beamformer for some simple scenarios. Consider first the case where the SOI is the only signal present (i.e., no interferers). Assuming each of the $m$ array elements has unity gain in the direction of the SOI, we have
\[
\gamma = \frac{1}{m + \text{SNR}_i}
\]
and hence
\[
\text{SNR}_o = \frac{\text{SNR}_i}{\gamma} = m \text{SNR}_i.
\]
Not surprisingly, the output SNR in this case is simply the input SNR times the array gain. Using the well-known approximation
\[
\Phi(x) \approx \frac{1}{x \sqrt{\pi}} e^{-x^2}
\]
we see that the use of multiple sensors reduces the SER by a factor of approximately $\sqrt{\text{SNR}_i (m-1)} \times \text{SNR}_i$ compared with the single sensor case.

When an uncorrelated interferer is present along with the SOI, it is easy to show that (assuming equal input SNR’s for both sources)
\[
\gamma = \frac{m + \text{SNR}_i^{-1}}{(m + \text{SNR}_i)^2 - |a_i^* a_d|^2}
\]
where $a_i$ and $a_d$ are the steering vectors for the interferer and SOI, respectively. The output SNR is found to be approximately
\[
\text{SNR}_o \approx m \left[ 1 - \left( \frac{|a_i^* a_d|^2}{m \text{SNR}_i} \right)^2 \right] \text{SNR}_i
\]
when $m \text{SNR}_i \gg 1$. The array gain can thus be significantly reduced if the interferer is highly spatially coherent with the SOI, although this effect is minimized when the input SNR is relatively high.

C. SER Performance of the LS Beamformer
For purposes of comparison, we derive in this section the SER that would be achieved by the LS beamformer of (7) if the DOA’s of all signals were known exactly. The SOI estimate in this case is given by
\[
\hat{s}_d(t) = s_d(t) + (A^*)^* \hat{n}(t) = s_d(t) + \hat{n}(t)
\]
and the power of the estimation error term is easily determined to be
\[
\mathbb{E} [\hat{n}(t)] = \sigma_n^2 \left[ (A^* A)^{-1} \right]_{dd}.
\]
The resulting output SNR for the LS beamformer is thus
\[
\text{SNR}_{\text{od}}(\text{LS}) = \frac{\text{SNR}_{\text{dd}}}{\left[ (A^* A)^{-1} \right]_{dd}}.
\]
In the discussion that follows, we show that the MMSE beamformer yields a higher output SNR (and hence a lower SER) than the LS approach. In the course of this analysis, we will make use of the following theorem:

**Theorem 1:** For any $Q = Q^* > 0$ and any vector $y$, the following inequality holds
\[
(y^* y) \leq (y^* Q y)(y^* Q^{-1} y).
\]

**Proof:** See the appendix.

Using (21), (32), and
\[
R_{\text{xx}} = A^*(R_{\text{xx}} - \sigma_n^2 I) A^* \quad A^* A^{-1}\]
the output SNR’s of the MMSE and LS beamformers may be written as
\[
\begin{align*}
\text{SNR}_{\text{od}}(\text{MMSE}) &= \frac{[A^* R_{\text{xx}} A - \sigma_n^2 A^* A^*]_{dd}}{\sigma_n^2 [A^* A]^*_{dd} - 1} \quad (40) \\
\text{SNR}_{\text{od}}(\text{LS}) &= \frac{[A^* R_{\text{xx}} A^* - \sigma_n^2 A^* A^*]_{dd}}{\sigma_n^2 [A^* A^*]_{dd} - 1} \quad (41)
\end{align*}
\]
Comparing (40) and (42), we see that the inequality $\text{SNR}_{\text{od}}(\text{MMSE}) \geq \text{SNR}_{\text{od}}(\text{LS})$ will hold provided that
\[
\begin{align*}
[A^* R_{\text{xx}} A^* - \sigma_n^2 A^* A^*]_{dd} &\geq \frac{[A^* R_{\text{xx}} A^*]_{dd}}{[A^* A^*]_{dd}} \quad (43)
\end{align*}
\]
Cross-multiplying and eliminating like terms on both sides leads to
\[
-[A^* A^*]_{dd} \geq -[A^* R_{\text{xx}} A^*]_{dd} [A^* R_{\text{xx}} A^*]_{dd}
\]
or equivalently
\[
[A^* A^*]_{dd} \leq [A^* R_{\text{xx}} A^*]_{dd} [A^* R_{\text{xx}} A^*]_{dd}. \quad (44)
\]
Letting \( y = (A^*)^t_d \) and \( Q = R_{xx} \) in Theorem 1, we see that (44) is indeed true, and thus we can conclude

\[
\text{SNR}_{\text{MMSE}} \geq \text{SNR}_{\text{LS}}. \tag{45}
\]

Together with the results of Sections IV-A and IV-B, the implication of (45) is that when it converges, the decision directed algorithm described earlier (which uses no information about the array response) will yield a lower SER than an LS beamformer employing precise knowledge of the DOA’s. This fact will be illustrated by one of the simulation examples of the following section.

V. SIMULATION EXAMPLES

As a simple example of the behavior of the decision directed (DD) algorithm, a scenario involving two 25 dB SNR (baseband) BPSK signals received by a six element \( \lambda/2 \) spaced uniform linear array (ULA) was simulated. The DOA’s of the signals were 10°, 16°, and the signals were assumed to be uncorrelated, have the same baud rate (6 samples per symbol), and be symbol synchronized with one another (worst case situation for signal separation). A very crude initial estimate of each signal was obtained by a classical delay-and-sum beamformer, and the resulting signal estimates were used to initialize the DD algorithm presented herein. Figs. 1 and 2 show the resulting beampattern the algorithm converged to for each signal. In each figure, the dashed line represents the beampattern of the initial weight vector (the vertical dashed line indicates the DOA the beam was steered toward), and the solid line indicates the beampattern after convergence. The other two vertical lines indicate the DOA’s of the signals. Notice that with only a very minimal degree of spatial discrimination in the initial estimate, the algorithm was able to converge and null out the interfering signal in each case.

The results of Section IV-A imply that, after convergence, the SER of the DD algorithm should ideally be independent of the SER of the initial signal estimate. In this example, we demonstrate that this result is approximately true over a wide range of initial SER’s. The simulation variables were as follows: two independent symbol synchronized BPSK signals sampled once per symbol, 0 dB SNR, 5 element \( \lambda/2 \) ULA, and DOA’s of 0° and 20°. For simplicity, in this case the initial estimate of the SOI (the broadside signal) was obtained by taking the actual signal and artificially generating symbol errors at various SER. A total of 500 trials (500 symbols per trial) were conducted for each initial SER, and the final SER of the DD algorithm was computed. The results are plotted in Fig. 3 along with the terminal SER predicted by (34). Our theoretical SER calculation accurately matches the actual terminal SER, and correctly predicts that the final error rate is independent of the initial error rate for initial SER’s as high as 0.4. Note that the predicted SER is accurate even though the interference is BPSK, and not Gaussian.

Our final two examples serve to compare the performance of the decision directed algorithm with other blind adaptive beamformers and with the LS beamformer. The output of a four element ULA was simulated assuming a QPSK SOI arriving from 10° with 10 dB SNR, and a Gaussian interferer with 8 dB SNR. The array was sampled three times per SOI baud. In the first example, the DOA of the interferer was set at 14°, and SER performance was computed as a function of the number of bauds used to train the beamformer.
weights of several blind adaptive algorithms. In addition to the DD approach, the constant modulus array (CMA), phase SCORE, and principle components (PC) phase SCORE [19] algorithms were tested. The CMA algorithm was implemented as described in [8] with \( p = 2 \) and \( \mu = 0.0075 \) (the value of \( \mu \) was chosen to be as large as possible while still yielding convergence), and both the CMA and DD approaches were initialized using a classical delay-and-sum beamformer steered to 12°. The baud rate feature of the SOI was exploited for the SCORE algorithms, and the delay parameter \( \tau \) was set to one (since this yielded the best performance). The SER of the above algorithms after \( 2.5 \times 10^6 \) symbol decisions is plotted in Fig. 4, along with the SER predicted by (35). Note that the DD algorithm converges much more rapidly than the other methods, achieving its theoretical performance limit after only about 50 symbol decisions.

The same scenario as above was used to compare the DD and LS beamformers, except that the DOA of the interferer was varied between 13°–19°, and 150 symbols were used to adapt the DD weights. For purposes of comparison, the LS algorithm was implemented with both known and estimated DOA's (the ESPRIT algorithm [20] was used in the latter case), while the DD algorithm was initialized using the LS weights obtained with estimated DOA's. The SER was calculated after \( 2.5 \times 10^6 \) symbol decisions, and is plotted in Fig. 5. The empirical results are denoted by the symbols \( 0 \), \( x \), *, while the SER's predicted by (35) and (39) appear as solid lines. As shown in the analysis of Section 4.3, the DD approach achieves a lower SER than the LS algorithm even when the DOA's are known exactly.

VI. CONCLUSION

A decision directed approach for blind adaptive beamforming has been presented. The algorithm uses symbol decisions made on an initial signal estimate to generate a reference signal that is, in turn, used to compute an estimate of the minimum mean-squared error (Weiner) beamformer weights. These weights allow a new signal estimate to be obtained, and the process can be repeated. An asymptotic analysis was conducted, and it was shown that in principle at least, the algorithm will converge to the Weiner solution regardless of the number of symbol errors in the initial estimate, provided enough data is used in estimating the beamformer weight vector. The symbol error rate of the optimal Weiner solution was analytically determined, and was shown to be lower in general than that obtained by a standard least-squares beamformer that assumes all the signal directions of arrival are known. A number of simulation examples were also presented to validate the analysis, and to demonstrate the advantage of the decision directed approach.

APPENDIX

PROOF OF THEOREM 1

In this appendix, we show that

\[ (y^* y)^2 \leq (y^* Q y) (y^* Q^{-1} y) \]  

(46)

holds for any Hermitian, positive definite matrix \( Q \). Assume \( Q \) is \( m \times m \), and let \( Q = U \Sigma U^* \) represent its singular value decomposition, where \( \Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_m^2) \). Then, (46) may be rewritten as

\[ (z^* z)^2 \leq (z^* \Sigma z) (z^* \Sigma^{-1} z) \]  

(47)

where \( z = U^* y \). Expanding both sides of the inequality into elemental form yields

\[
(z^* x)^2 = \sum_{k=1}^{m} \sum_{l=1}^{m} |z_k|^2 |z_l|^2 \\
= \sum_{k=1}^{m} \left( |z_k|^4 + \sum_{l=k+1}^{m} 2|z_k|^2 |z_l|^2 \right) \\
(z^* \Sigma z) (z^* \Sigma^{-1} z) = \sum_{k=1}^{m} \sum_{l=1}^{m} \frac{\sigma_k^2 \sigma_l^2}{\sigma_k^2 + \sigma_l^2} |z_k|^2 |z_l|^2 \\
= \sum_{k=1}^{m} \left( |z_k|^4 + \sum_{l=k+1}^{m} \left( \frac{\sigma_k^2}{\sigma_k^2 + \sigma_l^2} + \frac{\sigma_l^2}{\sigma_k^2 + \sigma_l^2} \right) |z_l|^2 |z_k|^2 \right). 
\]
where $z_k$ represents the $k$th element of $z$. The inequality is proved by noting that

$$\frac{\sigma_x^2}{\sigma_k^2} + \frac{\sigma_y^2}{\sigma_k^2} \geq 2$$

since $x + \frac{1}{x} \geq 2$ whenever $x > 0$.

**REFERENCES**


