

Minimum Variance Beamforming with Soft Response Constraints

Barry D. Van Veen, *Member, IEEE*

Abstract—Soft constraints on the beamformer response to the signal are examined in the context of minimum variance beamforming. A quadratic constraint on the beamformer weights is used to control the mean-squared error between a desired response and the actual response in the signal direction. The constraint is purposely chosen to permit distortion of the signal with the goal of achieving improved interference cancellation. Under the assumptions of known signal direction and spectral shape the signal-to-noise ratio is shown to be a nondecreasing function of the mean-square distortion. The distortion presented to the signal is easily computed from the beamformer weights and can be equalized after beamforming if desired. Properties of this beamforming method and its relationship to linearly constrained beamforming are discussed. Simulations verify analytic results and illustrate the utility of the soft constrained approach.

I. INTRODUCTION

A COMMON criterion for choosing the weights in a beamformer is minimization of the output power or variance subject to a constraint on the beamformer's response to the signal. Linear constraints have been studied in this context by several investigators, e.g., [1], [2], and specify the response to the signal exactly, independent of the interference environment. Recent work with small arrays operating at broad bandwidths has demonstrated that relaxation of these hard constraints on the beamformer response can result in improved interference cancellation at the expense of signal distortion [3]–[5].

Several methods have been suggested for controlling signal distortion while using relaxed or "soft" constraints. Kaneda and Ohga [3] artificially introduce a white noise signal into each sensor channel which is designed to represent the signal of interest. The distortion presented to the signal of interest is estimated and adjusted by measuring the distortion of the white noise signal. This approach requires significant computational effort since three beamforming systems must be implemented simultaneously in addition to the white noise signal generation procedure. Sondhi and Elko [4] utilize a nonlinear (fourth-order) constraint where the beamformer response is constrained to approximate an allpass filter for speech applications. Errors in the phase response are not deemed im-

portant since speech is not sensitive to phase. Chazan *et al.* [5] consider a two stage optimization approach where the signal is assumed to be absent in the first stage optimization. Requiring knowledge of the desired signal presence/absence is a significant limitation of this technique.

The approach presented here has the same goal, improved interference cancellation through relaxation of constraints, but takes a substantially different approach. It does not require the absence of the desired signal during adaptation or generation of an artificial signal which models the desired signal. In this paper, the mean-squared error between the desired and actual response is constrained. This results in a quadratic constraint on the weights. Quadratic constraints have been employed for a variety of purposes in adaptive beamforming [6]–[9]. In contrast with previous applications of quadratic constraints, here the constraint is purposely chosen to permit signal distortion with the goal of achieving improved interference cancellation. This soft constrained minimum variance (SCMV) philosophy represents a trade of "bias" (signal distortion) for reduced "variance" (interference power). Other examples of bias variance tradeoffs in signal processing are given in [10].

A key result of this paper is a proof guaranteeing that the SNR is a nondecreasing function of the bound on the mean-squared response error, assuming the signal direction and spectral shape are known. This implies that by allowing the signal distortion to increase the beamformer can provide much better interference cancellation, such that the SNR improves (or remains constant). The potential SNR improvement resulting from the use of soft constraints is greatest for systems operating at broad bandwidths; it is shown that no SNR improvement is obtained in narrow-band beamforming. The signal distortion can be computed from the beamformer weights and be equalized at the beamformer output if desired. In general, equalization changes the SNR.

This paper is organized as follows. Section II introduces notation and formulates the SCMV problem. Analysis of the SCMV beamformer is provided in Section III, including the proof of SNR improvement for increasing levels of distortion. Section IV discusses computational structures and algorithms for SCMV beamformer implementation. Equalization of the desired signal distortion is addressed in Section V. Simulations illustrating the effectiveness and characteristics of this method are provided in Section VI, followed by a summary in Section VII.

Manuscript received May 1, 1989; revised August 11, 1990. This work was supported in part by the Army Research Office under Award DAAL03-89-K-0141 and the National Science Foundation under Grant MIP-8958559.

The author is with the Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI 53706.
IEEE Log Number 9101339.

II. SCMV BEAMFORMER DERIVATION

Let the beamformer output y be expressed as the inner product of a weight vector \mathbf{w} and the total collection of data in the beamformer¹ \mathbf{x}

$$y = \mathbf{w}^H \mathbf{x}. \quad (1)$$

\mathbf{w} and \mathbf{x} are assumed to be N -dimensional vectors. Lowercase boldface letters denote vectors, uppercase boldface denotes matrices, and superscript H denotes complex conjugate transpose. Assuming the data are zero mean and wide-sense stationary, the output power is given by

$$P_o = E\{|y|^2\} = \mathbf{w}^H \mathbf{R}_x \mathbf{w} \quad (2)$$

where $\mathbf{R}_x = E\{\mathbf{x}\mathbf{x}^H\}$ is the N -by- N data covariance matrix.

Define $\mathbf{d}(\theta, \omega)$ as the array response vector,² that is, $\mathbf{d}(\theta, \omega)$ describes the amplitude and phase relationship between a specified reference point and each element of \mathbf{x} when the signal arriving at the array is of frequency ω and from direction θ . The response of the beamformer at frequency ω and direction θ is expressed as

$$r(\theta, \omega) = \mathbf{w}^H \mathbf{d}(\theta, \omega). \quad (3)$$

Let \mathbf{w}_d be a fixed set of beamformer weights which implements the desired response $r_d(\theta, \omega)$ to signals located within a band of frequencies Ω and range of directions Θ , i.e.,

$$r_d(\theta, \omega) = \mathbf{w}_d^H \mathbf{d}(\theta, \omega); \quad \theta \in \Theta, \omega \in \Omega. \quad (4)$$

The weighted mean-square response error between the desired and actual response over Θ and Ω is expressed as

$$e = \int_{\theta \in \Theta} \int_{\omega \in \Omega} \rho(\theta, \omega) |r_d(\theta, \omega) - \mathbf{w}^H \mathbf{d}(\theta, \omega)|^2 d\omega d\theta \quad (5)$$

where $\rho(\theta, \omega)$ is a nonnegative weighting function. Decompose \mathbf{w} as the sum of two components: $\mathbf{w} = \mathbf{w}_d - \mathbf{w}_a$. It is straightforward to show using (4) that

$$e = \mathbf{w}_a^H \mathbf{Q} \mathbf{w}_a \quad (6a)$$

where

$$\mathbf{Q} = \int_{\theta \in \Theta} \int_{\omega \in \Omega} \rho(\theta, \omega) \mathbf{d}(\theta, \omega) \mathbf{d}^H(\theta, \omega) d\omega d\theta. \quad (6b)$$

Thus, the weights which minimize the variance subject to a constraint on the mean-square response error satisfy

$$\min_{\mathbf{w}_a} (\mathbf{w}_d - \mathbf{w}_a)^H \mathbf{R}_x (\mathbf{w}_d - \mathbf{w}_a) \quad \text{subject to } \mathbf{w}_a^H \mathbf{Q} \mathbf{w}_a \leq e_0 \quad (7)$$

with e_0 representing the maximum tolerable mean-square response error.

¹When FIR filters are utilized at the sensor outputs, \mathbf{x} represents the data at the taps of the filters and \mathbf{w} the corresponding tap weights.

² $\mathbf{d}(\theta, \omega)$ is also known as the steering vector or direction vector. It characterizes propagation effects and the individual sensor response.

The constraint in (7) is essentially identical to that utilized in [7], [18]. In [7], the constraint is used for incorporating robustness to directional errors into the beamformer, while in [18] it is used for steering the beamformer. In both cases e_0 is assumed to be a "small number" so the desired response is approximately achieved over the sectors Θ and Ω of interest. Here we are specifically interested in considering relatively "large" values of e_0 to obtain improved interference cancellation.

The linearly constrained minimum variance beamforming problem

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{R}_x \mathbf{w} \quad \text{subject to } \mathbf{C}^H \mathbf{w} = \mathbf{f} \quad (8)$$

is easily derived from (7) when \mathbf{Q} is an approximately low rank matrix. In (8) the linear constraints are used to guarantee that \mathbf{w} implements the desired response on Θ and Ω . Decompose \mathbf{w} as before, $\mathbf{w} = \mathbf{w}_d - \mathbf{w}_a$, where \mathbf{w}_d satisfies (4) or, equivalently, $\mathbf{C}^H \mathbf{w}_d = \mathbf{f}$. This implies that $\mathbf{C}^H \mathbf{w}_a = \mathbf{0}$, i.e. \mathbf{w}_a must lie in a subspace orthogonal to the space spanned by \mathbf{C} . If \mathbf{Q} is approximately low rank, then we obtain $e_0 \approx 0$ when \mathbf{w}_a lies in the nullspace of \mathbf{Q} . Thus, the space spanned by the columns of \mathbf{C} corresponds to the space spanned by the columns of \mathbf{Q} . Further discussion of the relationship between the quadratically and linearly constrained beamformers of (7) and (8) is given in [7].

The solution to (7) is obtained via the method of Lagrange multipliers. Define

$$L(\mathbf{w}_a, \lambda) = (\mathbf{w}_d - \mathbf{w}_a)^H \mathbf{R}_x (\mathbf{w}_d - \mathbf{w}_a) + \lambda (\mathbf{w}_a^H \mathbf{Q} \mathbf{w}_a - e_0). \quad (9)$$

Setting the gradient of (9) with respect to \mathbf{w}_a equal to zero yields

$$\mathbf{w}_a = (\mathbf{R}_x + \lambda \mathbf{Q})^{-1} \mathbf{R}_x \mathbf{w}_d \quad (10)$$

where the Lagrange multiplier λ is chosen as large as possible subject to $\mathbf{w}_a^H \mathbf{Q} \mathbf{w}_a \leq e_0$ [14]. Note that the cost function in (9) and the solution (10) correspond to the ridge regression or regularization [11]–[13] problem and its solution. λ and e_0 are inversely related [14]. A distortionless beamformer ($e_0 = 0$) has $\lambda = \infty$ while an unconstrained beamformer has $\lambda = 0$.

Since the relationship between λ and e_0 is one to one, λ can also be used to define a unique set of SCMV weights. Parameterizing the SCMV weights in terms of λ may result in significant computational savings for some implementations because an iterative algorithm is usually required to solve for λ given e_0 . The implementation tradeoffs between parameterizing the soft constraint in terms of λ instead of e_0 are discussed in greater detail in Section IV. The disadvantage of parameterizing the weights with λ is a loss of control over the mean-square error, since if λ is fixed, e_0 will vary as the signal and interference environment varies. In the following section the dependence of e_0 , the SNR, and the signal power on λ is examined. This provides general guidelines for choosing λ .

III. SNR IMPROVEMENT AND RELATED PROPERTIES

This section analyzes several SCMV beamformer properties assuming the signal power spectrum shape is known. This assumption may be reasonable in some cases, e.g., in speech applications the average spectral shape of speech can be used. Even if this assumption is not satisfied exactly, the analysis does provide useful insight into SCMV beamformer performance characteristics. We begin the section by proving that the SCMV system SNR must improve or remain constant as increasing levels of distortion are permitted and conclude with examination of several performance characteristics as a function of λ .

\mathbf{Q} as given in (6b) corresponds to the data covariance matrix due to a source with power spectral density $\rho(\theta, \omega)$. If Θ corresponds to a single direction, θ_s , then \mathbf{Q} represents a broad-band spatial point source. Assume that $\mathbf{R}_x = \mathbf{R}_s + \mathbf{R}_I$ where \mathbf{R}_s represents the covariance of the signal and \mathbf{R}_I the covariance of the interference and noise, assumed uncorrelated with the signal. Further assume that $\rho(\theta_s, \omega)$ represents the shape and σ_s the level of the signal power spectral density. Thus, the power spectral density of the signal is given by $\sigma_s \rho(\theta_s, \omega)$ and $\mathbf{R}_s = \sigma_s \mathbf{Q}$. Rewrite the equation for \mathbf{w}_a given in (10) as

$$\mathbf{w}_a = [\mathbf{R}_I + (\sigma_s + \lambda)\mathbf{Q}]^{-1}(\mathbf{R}_I + \sigma_s \mathbf{Q})\mathbf{w}_d. \quad (11)$$

The SNR is the ratio of signal P_s to interference P_I powers at the beamformer output

$$\text{SNR} = \frac{P_s}{P_I} = \frac{(\mathbf{w}_d - \mathbf{w}_a)^H \sigma_s \mathbf{Q} (\mathbf{w}_d - \mathbf{w}_a)}{(\mathbf{w}_d - \mathbf{w}_a)^H \mathbf{R}_I (\mathbf{w}_d - \mathbf{w}_a)}. \quad (12)$$

A generalized eigendecomposition of \mathbf{R}_I and \mathbf{Q} is useful for analysis of (11), (12). Let $\mathbf{R}_I = \mathbf{G}\mathbf{G}^H$ with \mathbf{G} a nonsingular matrix. Furthermore, let $\mathbf{V}\mathbf{\Gamma}\mathbf{V}^H$ be the eigendecomposition of $\mathbf{G}^{-1}\mathbf{Q}[\mathbf{G}^H]^{-1}$ with the columns of \mathbf{V} representing the eigenvectors and $\mathbf{\Gamma}$ a diagonal matrix with the corresponding eigenvalues on the diagonal. Denote the i th diagonal element of $\mathbf{\Gamma}$ (eigenvalue) as γ_i . The nonnegative definiteness of \mathbf{Q} guarantees that $\gamma_i \geq 0$. Substituting into (11) we derive

$$\mathbf{w}_a = \mathbf{G}^{-H} \mathbf{V} \mathbf{\Sigma} \mathbf{V}^H \mathbf{G}^H \mathbf{w}_d \quad (13a)$$

where $\mathbf{\Sigma}$ is the diagonal matrix

$$\mathbf{\Sigma} = [\mathbf{I} + (\sigma_s + \lambda)\mathbf{\Gamma}]^{-1}(\mathbf{I} + \sigma_s \mathbf{\Gamma}). \quad (13b)$$

After straightforward manipulation we obtain

$$P_s = \sigma_s \mathbf{w}_d^H \mathbf{G} \mathbf{V} (\mathbf{I} - \mathbf{\Sigma})^2 \mathbf{V}^H \mathbf{G}^H \mathbf{w}_d \quad (14a)$$

$$P_I = \mathbf{w}_d^H \mathbf{G} \mathbf{V} (\mathbf{I} - \mathbf{\Sigma})^2 \mathbf{V}^H \mathbf{G}^H \mathbf{w}_d. \quad (14b)$$

Define

$$g_i = |[V^H \mathbf{G}^H \mathbf{w}_d]_i|^2 \quad (15a)$$

and

$$h_i = (\mathbf{I} - \mathbf{\Sigma})_{ii}^2 = \frac{\gamma_i^2 \lambda^2}{[1 + (\sigma_s + \lambda)\gamma_i]^2} \quad (15b)$$

to simplify notation. The constraint is expressed as

$$e_0 \geq \mathbf{w}_d^H \mathbf{Q} \mathbf{w}_d = \mathbf{w}_d^H \mathbf{G} \mathbf{V} \mathbf{\Sigma}^2 \mathbf{V}^H \mathbf{G}^H \mathbf{w}_d$$

or

$$e_0 \geq \sum_{i=1}^N g_i \gamma_i \left[\frac{1 + \sigma_s \gamma_i}{1 + (\lambda + \sigma_s) \gamma_i} \right]^2. \quad (16)$$

Equations of this form are termed secular equations [14]. The constraint will be satisfied with equality except for degenerate cases. The SNR is expressed as

$$\text{SNR} = \frac{\sigma_s \sum_{i=1}^N g_i \gamma_i h_i}{\sum_{i=1}^N g_i h_i}. \quad (17)$$

Note that there are at least two conditions for which the SNR is independent of λ : 1) if \mathbf{Q} is rank one, or 2) if all nonzero γ_i are equal. In both cases all nonzero h_i are equal and (17) simplifies to $\text{SNR} = \sigma_s \gamma_1$.

Theorem: The SNR as defined in (12) and (17) is a nonincreasing function of λ on the interval $0 \leq \lambda \leq \infty$.

The proof is given in the Appendix. Recall that λ and e_0 are inversely related. The theorem indicates that as e_0 increases, i.e., as greater levels of distortion are tolerated, the SNR improves or at least remains constant. This represents a trade of bias (distortion) for reduced variance (noise power).

Note that as λ approaches zero (maximum SNR) the array shuts down, i.e., $\mathbf{\Sigma}$ approaches the identity matrix so \mathbf{w}_a in (13a) approaches \mathbf{w}_d which implies that \mathbf{w} goes to zero. As λ approaches infinity (minimum SNR) (13a) indicates that the components of \mathbf{w}_a corresponding to $\gamma_i > 0$ are zero. This is consistent with the requirement in the linearly constrained problem (8) that \mathbf{w}_a lie in the space orthogonal to the constraint space. If \mathbf{Q} is full rank and λ approaches infinity, then $\mathbf{w}_a = \mathbf{0}$, corresponding to the linearly constrained weights under N constraints (zero degrees of freedom).

The SNR is also a function of σ_s , the signal power. Suppose λ is held fixed and σ_s increases. Differentiation of (16) with respect to σ_s indicates that e_0 increases as σ_s increases. Using an approach similar to that described in the Appendix, it can be shown that $\sigma_s \text{SNR}$ is a nondecreasing function of σ_s and $\sigma_s^{-1} \text{SNR}$ is a nonincreasing function of σ_s . Thus, as σ_s increases, the SNR cannot decrease faster than σ_s^{-1} or increase faster than σ_s . Now suppose e_0 is held fixed and σ_s increases. This implies from (16) that λ must also increase so there are two parameters changing in the SNR expression. Increasing λ tends to decrease the SNR while increasing σ_s may increase or decrease the SNR. It is not evident in general how these individual effects combine. However, the above theorem indicates that the soft constrained ($\lambda < \infty$) SNR is always greater than the hard constrained ($\lambda = \infty$) SNR, independent of σ_s , because the hard constrained SNR is independent of σ_s .

A. Narrow-Band Beamforming

Assume the response is constrained at a single point of direction θ_0 and frequency ω_0 so that $\mathbf{Q} = \mathbf{d}(\theta_0, \omega_0) \mathbf{d}^H(\theta_0, \omega_0)$. \mathbf{Q} is rank one so $\gamma_i = 0$ for $i \geq 2$. Substituting for \mathbf{Q} in (10) and applying the matrix inversion lemma gives

$$\mathbf{w} = \frac{\lambda \mathbf{d}^H(\theta_0, \omega_0) \mathbf{w}_d}{1 + \lambda \mathbf{d}^H(\theta_0, \omega_0) \mathbf{R}_x^{-1} \mathbf{d}(\theta_0, \omega_0)} \mathbf{R}_x^{-1} \mathbf{d}(\theta_0, \omega_0). \quad (18)$$

The minimum variance distortionless response (MVDR) beamformer weights for the narrow-band case are obtained by solving (8) for $\mathbf{C} = \mathbf{d}(\theta_0, \omega_0)$ and $\mathbf{f} = 1$, yielding

$$\mathbf{w} = \frac{1}{\mathbf{d}^H(\theta_0, \omega_0) \mathbf{R}_x^{-1} \mathbf{d}(\theta_0, \omega_0)} \mathbf{R}_x^{-1} \mathbf{d}(\theta_0, \omega_0). \quad (19)$$

The weight vectors in (18) and (19) differ only by a scale factor. The SNR in (17) becomes

$$\text{SNR}_{\text{NB}} = \sigma_s \gamma_1 \quad (20)$$

which is independent of λ for all interference environments. Thus, the distortionless response ($\lambda = \infty$) and soft constrained ($0 < \lambda < \infty$) beamformers provide identical SNR. γ_1 is the only nonzero eigenvalue of $\mathbf{G}^{-1} \mathbf{Q} [\mathbf{G}^H]^{-1}$ which is easily shown to be

$$\gamma_1 = \mathbf{d}^H(\theta_0, \omega_0) \mathbf{R}_I^{-1} \mathbf{d}(\theta_0, \omega_0). \quad (21)$$

B. Bounds Involving λ

As noted in Section II, under some conditions it may be desirable to avoid solving for λ in real time given e_0 . Instead, λ can be fixed to a prespecified value at the expense of a loss of control over e_0 . The SNR and signal distortion are a function of λ so it is important that λ not be chosen in a haphazard fashion. An offline, brute force approach to choosing λ can be used if partial information about likely signal and interference conditions is available. e_0 and the SNR could be plotted as a function of λ for specific examples of expected signal/interference scenarios. λ would then be chosen as a value which provided acceptable performance over these examples. Here we approach this problem in more general sense and present several inequalities concerning the behavior of e_0 , the SNR, and P_s as a function of λ .

Using (13b) we determine that

$$\left[\frac{\sigma_s}{\sigma_s + \lambda} \right]^2 < \Sigma_{ii}^2 < 1 \quad (22a)$$

from which it follows using (16) and $\mathbf{G} \mathbf{V} \mathbf{V}^H \mathbf{G} = \mathbf{Q}$ that

$$\left[\frac{\sigma_s}{\sigma_s + \lambda} \right]^2 \mathbf{w}_d^H \mathbf{Q} \mathbf{w}_d < e_0 < \mathbf{w}_d^H \mathbf{Q} \mathbf{w}_d. \quad (22b)$$

$\mathbf{w}_d^H \mathbf{Q} \mathbf{w}_d$ is the maximum e_0 and is approached as the array shuts down (λ approaches zero). If an upper bound on e_0 is specified, then the lower bound in (22b) determines the minimum permissible value for λ relative to σ_s . For example, if we require e_0 to be less than one fourth the maximum error, then we cannot choose $\lambda \leq \sigma_s$ since if $\lambda = \sigma_s$, then $e_0 > 1/4 \mathbf{w}_d^H \mathbf{Q} \mathbf{w}_d$.

A similar bound on the minimum possible λ is obtained in terms of P_s (14a) by noting

$$(\mathbf{I} - \Sigma)_{ii}^2 < \left[\frac{\lambda}{\lambda + \sigma_s} \right]^2 \quad (23a)$$

and $\mathbf{R}_s = \sigma_s \mathbf{Q}$ so

$$P_s < \left[\frac{\lambda}{\lambda + \sigma_s} \right]^2 \mathbf{w}_d^H \mathbf{R}_s \mathbf{w}_d. \quad (23b)$$

$\mathbf{w}_d^H \mathbf{R}_s \mathbf{w}_d$ is the distortionless response signal output power. If a lower bound on P_s is specified, then (23b) determines the minimum value for λ relative to σ_s . For example, if we require P_s to be greater than one fourth the maximum signal output power, then we cannot choose $\lambda \leq \sigma_s$.

It was established earlier that the SNR is a nonincreasing function of λ . Using a similar approach it can be shown that λ^2 SNR is a nondecreasing function of λ . This places an upper bound on the rate the SNR can increase as λ decreases. Suppose the SNR is SNR_1 at λ_1 and SNR_2 at $\lambda_2 < \lambda_1$. We have $\text{SNR}_2 \leq (\lambda_1^2 / \lambda_2^2) \text{SNR}_1$. Similarly, it can be shown that P_s is a nondecreasing function of λ , while $\lambda^{-2} P_s$ is a nonincreasing function. This implies that if λ is reduced from λ_1 to λ_2 , then we know $P_{s2} > (\lambda_2^2 / \lambda_1^2) P_{s1}$. Finally, differentiation of $\lambda^2 e_0$ indicates that $\lambda^2 e_0$ is a nondecreasing function of λ if $\lambda + \sigma_s \geq 1$. Again reducing λ from λ_1 to λ_2 , we have $e_{02} < (\lambda_1^2 / \lambda_2^2) e_{01}$. These results indicate the maximum SNR benefit, signal power loss, and increase in mean-squared error resulting from a reduction in λ . For example, reducing λ by a factor of two will give at most a factor of four: 1) increase in SNR, 2) loss in signal power, and 3) increase in mean-squared error.

These are not only helpful in choosing a good λ *a priori*, but are also helpful in applications where λ is adjusted in real time.

IV. ALGORITHMS AND ARCHITECTURES FOR SCMV BEAMFORMER IMPLEMENTATION

A variety of adaptive algorithms and computational architectures can be utilized to implement the SCMV beamformer. Several are discussed in this section.

A. A Sidelobe Cancelling Structure

Define the eigendecomposition of \mathbf{Q} as $\mathbf{F} \mathbf{D} \mathbf{F}^H$ where the columns of \mathbf{F} are eigenvectors corresponding to the eigenvalues located on the diagonal of the diagonal matrix \mathbf{D} . Transform \mathbf{w}_d so that $\mathbf{w}_a = \mathbf{F} \mathbf{w}_f$. Equation (7) is now written

$$\begin{aligned} \min_{\mathbf{w}_f} (\mathbf{w}_d - \mathbf{F} \mathbf{w}_f)^H \mathbf{R}_x (\mathbf{w}_d - \mathbf{F} \mathbf{w}_f) \\ \text{subject to } \mathbf{w}_f^H \mathbf{D} \mathbf{w}_f \leq e_0. \end{aligned} \quad (24)$$

In many practical cases \mathbf{Q} is approximately low rank. If \mathbf{Q} is rank p , then only p diagonal elements of \mathbf{D} are nonzero and the constraint in (24) only affects p of the elements of \mathbf{w}_f . The remaining $N-p$ are unconstrained. De-

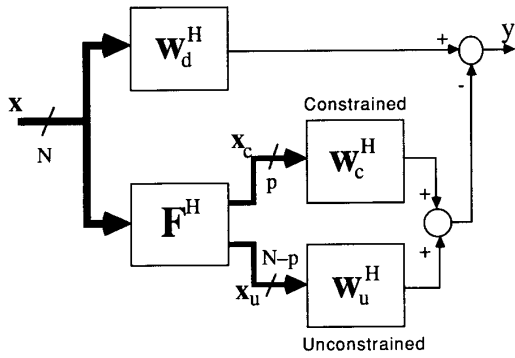


Fig. 1. Decomposition of soft constrained beamformer into constrained and unconstrained adaptive components.

fine w_c and w_u as the p constrained and $N-p$ unconstrained components of w_f , respectively, and \tilde{D} as the p -by- p diagonal matrix constructed from the p nonzero elements of D . Using these definitions, the constraint is written $w_c^H \tilde{D} w_c \leq e_0$; a block diagram depicting this beamformer structure is given in Fig. 1.

B. The Frost Case

The SCMV beamformer has a particularly simple implementation for the situation considered by Frost [1]. Frost assumed the array is steered to the signal direction θ_s so the signal is aligned in phase across the sensors. If there are L sensors and M tap FIR filters in each sensor channel, then $d(\theta_s, \omega) = d(\omega) \otimes \mathbf{1}$ where \otimes denotes Kronecker product,³ $d(\omega)$ is an M -dimensional vector with k th element $e^{-j\omega(k-1)}$, and $\mathbf{1}$ is the L -dimensional vector of ones. Frost considers $\Theta = \theta_s$ and Ω as the entire frequency band $-\pi \leq \omega \leq \pi$. Setting $\rho(\theta, \omega) = (2\pi)^{-1}$ in (6b) we obtain

$$Q = CC^H \quad (25a)$$

where

$$C = I_M \otimes \mathbf{1}. \quad (25b)$$

I_M is the M -by- M identity matrix. Note that C is the constraint matrix of [1]. A valid w_d is obtained from the minimum norm solution to the linear constraints $C^H w_d = f$. Substituting (25b) and simplifying yields

$$w_d = L^{-1}(f \otimes \mathbf{1}). \quad (26)$$

Using (25), the eigendecomposition of Q , FDF^H , is easily verified to be $F = I_M \otimes P$ and $D = I_M \otimes E$ where PEP^H is the eigendecomposition of $\mathbf{1}\mathbf{1}^H$. As a result of the special structure of $\mathbf{1}\mathbf{1}^H$, P and E are easily obtained. E is a matrix of all zeros except for the value L in the 1,1 position. The first column of P is $(L)^{-1/2} \mathbf{1}$ and the remaining columns are arbitrary provided they sum to zero (are orthogonal to $\mathbf{1}$).

Consider the transformed data $F^H x$. Let x_s represent

the data at the L sensor outputs and ψ the M vector of unit delay operators $\psi = [1 \ z^{-1} \ z^{-2} \ \dots \ z^{-(M-1)}]^T$. Thus, $x = \psi \otimes x_s$, and

$$\begin{aligned} F^H x &= (I_M \otimes P)^H (\psi \otimes x_s) \\ &= \psi \otimes (P^H x_s). \end{aligned} \quad (27)$$

Equation (27) indicates that transforming x by F^H is equivalent to transforming x_s by P^H and then placing tap delay lines in each transformed channel. Similarly, using (26)

$$\begin{aligned} w_d^H x &= \frac{1}{L} (f \otimes \mathbf{1})^H (\psi \otimes x_s) \\ &= \frac{1}{L} (f \otimes \psi) \otimes (\mathbf{1}^H x_s) \\ &= \left(\frac{1}{L} f^H \psi \right) (\mathbf{1}^H x_s). \end{aligned} \quad (28)$$

Thus, the output of the beamformer with weight vector w_d is the output of an FIR filter whose input is the sum of the sensor data ($\mathbf{1}^H x_s$) and coefficients are given by the elements of $L^{-1} f^H \psi$. Noting that the sum of the sensor data (scaled by $L^{-1/2}$) is the first element of $P^H x_s$, we obtain the structure depicted in Fig. 2. In Fig. 2 P is partitioned as $P = [L^{-1/2} \mathbf{1} \ \tilde{P}]$. \tilde{P} represents the $L-1$ eigenvectors of $\mathbf{1}\mathbf{1}^H$ associated with zero eigenvalues, hence each column of \tilde{P} sums to zero and "blocks" the desired signal. This structure is the SCMV equivalent of the Griffiths and Jim beamformer [16]. The constraint on the FIR filter with coefficients w_c is $L w_c^H w_c \leq e_0$. If e_0 is zero, then $w_c = \mathbf{0}$ and structure in Fig. 2 reduces to that of Griffiths and Jim.

C. Adaptive Algorithms

A gradient based scaled projection algorithm is given in [6] for minimizing output power subject to linear and quadratic constraints on the weights. The projection step in the algorithm of [6] is not necessary here since there are no linear constraints active. In the scaled projection algorithm the weights are explicitly dependent on e_0 and not λ . Thus, there is no computational advantage to parameterizing the weights in terms of λ if this type of algorithm is used.

The SCMV beamforming problem is easily formulated as a regularized least squares problem. This permits application of the algorithms in [13] and systolic architectures of [15] to solve for the beamformer weights. Solutions to the least squares problem are explicitly dependent on λ and not e_0 . Determining the value for λ which satisfies the constraint $w_f^H D w_f \leq e_0$ (or $w_a^H Q w_a \leq e_0$) generally requires solving the least squares problem multiple times. Thus, in contrast to the gradient based scaled projection algorithm discussed above, the computational complexity is greatly reduced in the least squares approach by fixing λ to some prespecified value and tolerating the resulting mean-square error.

³The Kronecker product identities used in this subsection are given in [17, table II].

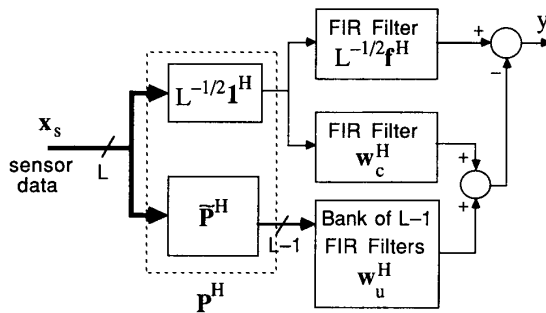


Fig. 2. SCMV beamformer structure for case considered by Frost [1]. $\bar{\mathbf{P}}$ is a matrix whose columns sum to zero.

V. EQUALIZATION

Once the weights \mathbf{w} are determined in the SCMV beamformer, the response is defined by (3). Thus, the distortion resulting from the soft constraints can be computed using (3) and be equalized by temporal filtering at the beamformer output. Suppose the signal arrives from direction θ_s . The distortion is equalized by processing the beamformer output with a filter having frequency response

$$H(\omega) = \begin{cases} [\mathbf{w}^H \mathbf{d}(\theta_s, \omega)]^{-1}, & \omega \in \Omega \\ 0, & \omega \notin \Omega \end{cases} \quad (29)$$

If the magnitude of $\mathbf{w}^H \mathbf{d}(\theta_s, \omega)$ is zero or very small at some ω , then the inverse should be replaced by a pseudoinverse since there is no point in attempting to equalize a frequency at which the signal has already been zeroed. This prevents amplification of any noise which may leak into the postbeamforming system.

Equalization may or may not be desirable, depending on the application. In cases where the SCMV weights are parameterized by λ , equalization diminishes the significance of picking good *a priori* values for λ because it removes the effect of the distortion. Equalization also provides the possibility of obtaining increased SNR without signal distortion. The SNR at the output of an equalized SCMV beamformer can be significantly larger than the SNR at the output of a hard constrained beamformer.

In general, the equalization process will change the SNR at the SCMV beamformer output. Furthermore, the theorem in Section III which established that the SNR is a nonincreasing function of λ does not apply to an equalized SCMV beamformer. In order to obtain insight into equalized SCMV beamformer performance, let the unequalized beamformer output y consist of a signal term s and an interference/noise term n : $y = s + n$. The theorem of Section III applies to a broad-band SNR, reexpressed here as

$$\text{SNR} = \frac{E\{|s|^2\}}{E\{|n|^2\}} = \frac{\int_{-\pi}^{\pi} S(\omega) d\omega}{\int_{-\pi}^{\pi} N(\omega) d\omega} \quad (30)$$

$S(\omega)$ and $N(\omega)$ are the power spectral densities of s and n . The theorem does not apply to the SNR at each frequency, $\text{SNR}(\omega) = S(\omega)/N(\omega)$, which can be an increasing or decreasing function of the mean-square distortion. The SNR of (30) can increase due to a large improvement in $\text{SNR}(\omega_1)$ even though $\text{SNR}(\omega_2)$ decreases. However, if the SCMV approach improves the SNR in (30), then it is reasonable to expect that $\text{SNR}(\omega)$ improves over most values of ω .

The equalized SNR is given by

$$\text{SNR}_{\text{eq}} = \frac{\int_{-\pi}^{\pi} |H(\omega)|^2 S(\omega) d\omega}{\int_{-\pi}^{\pi} |H(\omega)|^2 N(\omega) d\omega} \quad (31)$$

In (31), the signal and noise power spectral densities are weighed by $|H(\omega)|^2$ before integration. If $|H(\omega)|^2$ is approximately constant, then equalization does not have a large effect on the broad-band SNR. The greatest potential for a loss in broad-band SNR due to equalization occurs when $|H(\omega)|^2$ is large at values of ω for which $\text{SNR}(\omega)$ is small and small at ω for which $\text{SNR}(\omega)$ is large. Note, however, that $\text{SNR}(\omega)$ is not changed by the equalization process.

If one only desires to equalize the overall gain, i.e., multiply y by a constant, then the SNR is clearly unchanged. The weighted average gain to the signal is one possible basis for an equalization constant. That is, the output y is scaled by

$$\beta = \left[\int_{\omega \in \Omega} \nu(\omega) |\mathbf{w}^H \mathbf{d}(\theta_s, \omega)|^2 d\omega \right]^{-1/2} \quad (32)$$

If the weighting factor $\nu(\omega)$ integrates to unity over Ω and has the same shape as the signal power spectrum, then $S(\omega) = \sigma_s \nu(\omega) |\mathbf{w}^H \mathbf{d}(\theta_s, \omega)|^2$ and it follows that the correct signal power level is obtained, i.e., $E\{|\beta s|^2\} = \sigma_s$.

VI. SIMULATIONS

The effectiveness and features of the SCMV approach to beamforming are demonstrated in this section. In all simulations the true covariance matrices are assumed known in order to isolate theoretical SCMV performance from finite data effects. Analysis of the effects of covariance estimation errors is beyond the scope of the present work.

The array used in all simulations consists of five elements in a linear configuration spaced at one half wavelength at one half the temporal sampling frequency. Five tap FIR filters are used in each sensor channel resulting in a total of twenty-five adaptive weights. With one exception, the signals and interference are real and assumed to have constant spectral level on $0.4 \leq |\omega| \leq 3$ where the sampling rate is chosen so that ω is normalized to $-\pi \leq \omega \leq \pi$. The desired response weights \mathbf{w}_d are chosen to approximate unit magnitude and linear phase response in a least squares sense over the band $0.4 \leq |\omega| \leq 3$ in the

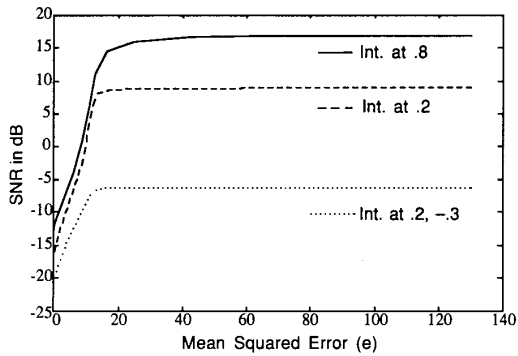


Fig. 3. SNR as a function of mean-squared error for three interference scenarios.

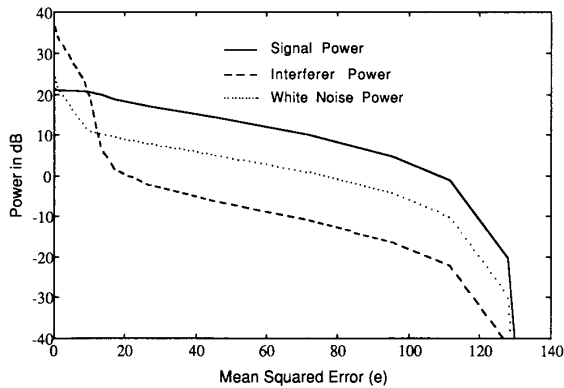


Fig. 4. Signal, white noise, and interferer output powers as a function of mean-squared error when the interferer is located at a direction sine of 0.2.

signal direction. All directions are given as the sine of the angle between source location and the perpendicular to the array (broadside).

Fig. 3 depicts the SNR as a function of mean-squared error (e) for three different interference scenarios when the signal arrives at the array from the broadside direction. In one case the interferer is at 0.8, in another at 0.2, and in the third there are two interferers, one at 0.2 and the other at -0.3 . The interferers have power levels of 30 dB relative to the signal power. White noise of relative power -10 dB is also present. The maximum value for e , $w_d^H Q w_d$, is 130. The SCMV approach provides SNR improvements of 30, 25, and 14 dB over distortionless response beamforming ($e = 0$) by choosing moderate values for e . The majority of the improvement occurs for small values of e and very little improvement is obtained by allowing larger values of distortion.

Fig. 4 illustrates the signal, white noise, and interferer output powers as a function of the mean-squared error for the case of a single interferer located at 0.2. Comparison to Fig. 3 indicates that a 25 dB gain in SNR is obtained at the expense of about 4 dB loss in signal power. Note that for small values of mean-squared error the white noise gain decreases much more rapidly than the signal power.

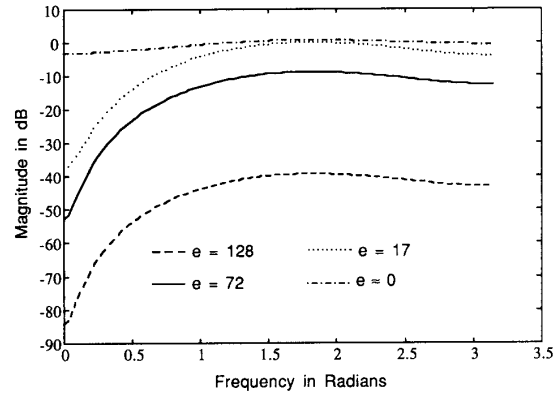


Fig. 5. Gain in the signal direction as a function of frequency for four values of mean-squared error (e) when an interferer is located at direction sine 0.2.

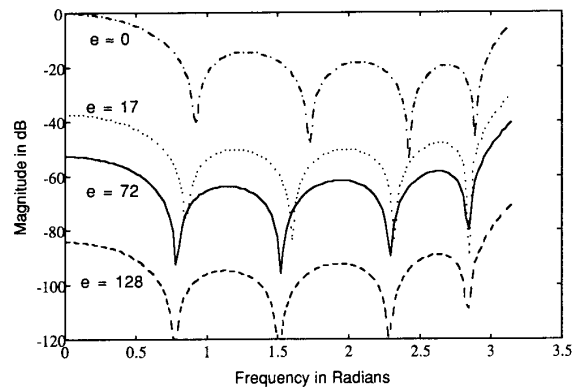


Fig. 6. Gain in the interference direction (0.2) as a function of frequency for four values of mean-squared error (e).

The poor SNR (-16 dB) and large white noise power (24 dB) indicate that the hard constrained system ($e = 0$) is severely stressed in this interference environment. A small relaxation of the hard constraint provides the beamformer flexibility to attenuate the interference using a weight vector with a much smaller norm, resulting reduced white noise power. Beyond a mean-squared error of 20 little improvement is obtained as the ratios of signal to white noise and signal-to-interference powers are approximately constant.

The magnitude response of the beamformer in the signal and interference directions is illustrated in Figs. 5, 6 for several values of mean-squared error. The interferer is located at 0.2. The majority of the signal distortion occurs at low frequencies (recall that the lowest frequency in the signal is 0.4). The basic response shapes do not change appreciably for larger values of e .

Fig. 7 depicts the equalized SNR as a function of mean-squared error for the same signal/interference scenarios used in Fig. 3. After equalization the overall improvement in SNR is 25, 21, and 10 dB, respectively. Thus, equalization results in a loss of about 4–5 dB in maximum SNR gain for these three examples. Fig. 5 depicts

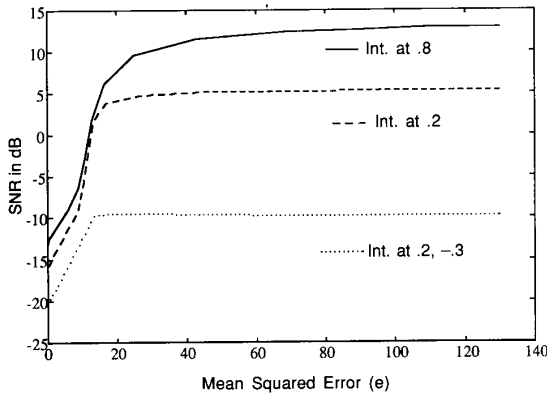


Fig. 7. Equalized SNR as a function of mean-squared error for three interference scenarios.

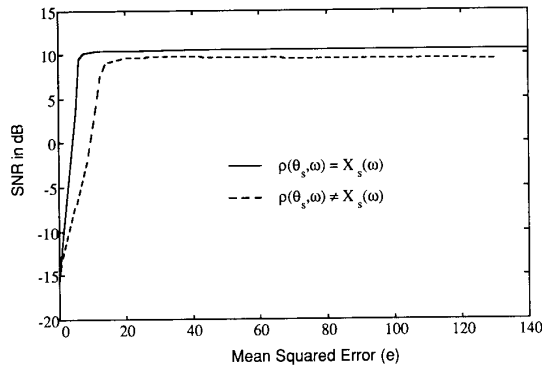


Fig. 8. SNR as a function of mean-squared error for a single interferer at 0.2. The signal has spectrum $X_s(\omega) = 1 + 0.9 \cos(\omega - 3)$, $0.4 \leq \omega \leq 3$, $X_s(-\omega) = X_s(\omega)$.

$|w^H d(\theta_s, \omega)|$ when the interferer is at 0.2, so $|H(\omega)|$ is the inverse of each curve. For each e in Fig. 5, almost all the variation in $|H(\omega)|$ occurs at the low frequencies; $|H(\omega)|$ is relatively constant above 1.2 rad.

The effect of a difference between the signal power spectral density and $\rho(\theta_s, \omega)$ is depicted in Fig. 8. The signal power spectral density has a raised cosine shape: $X_s(\omega) = 1 + 0.9 \cos(\omega - 3)$, $0.4 \leq \omega \leq 3$, $X_s(-\omega) = X_s(\omega)$. The dashed line represents the SNR obtained assuming $\rho(\theta_s, \omega)$ is constant on $4 \leq |\omega| \leq 3$ while the solid line assumes $\rho(\theta_s, \omega) = X_s(\omega)$. There is only a slight performance loss if $\rho(\theta_s, \omega) \neq X_s(\omega)$. Similar results were obtained using several other signal spectra and interference scenarios. The SNR improvement theorem does not hold when $\rho(\theta_s, \omega) \neq X_s(\omega)$ as evidenced in this example; at larger values of e the SNR decreases slightly.

An example illustrating the use of the SCM beamformer for spatial spectrum estimation is given in Fig. 9. The power in each direction is computed as the power at the output of an SCM beamformer steered to that direction. For these examples gain equalization is used assuming $\nu(\omega)$ in (32) is constant. The hard constrained estimates are computed using $\lambda = 30\,000$ while the soft

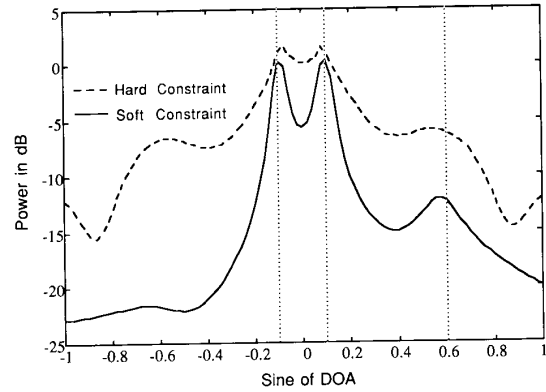


Fig. 9. Spatial spectrum estimates for three sources in white noise of -30 dB. The sources at ± 0.1 are each 0 dB, the source at 0.6 has relative power -20 dB. Dotted lines indicate true source locations.

constrained estimates assume $\lambda = 0.05$. Three sources are present: two unit power sources at ± 0.1 and a -20 dB source at 0.6 . The background white noise level is -30 dB. The presence of the weak source is difficult, if not impossible, to detect in the hard constrained estimate. The soft constrained estimate clearly indicates the presence of a third source. The source energy is smeared throughout all directions in the hard constrained estimate because the hard constrained beamformer cannot adequately attenuate the “interference” which leaks through its sidelobes.

VII. SUMMARY

A soft constrained approach to minimum variance beamforming is presented and analyzed. The constraint acts on the mean-squared error between the desired and actual beamformer response and is chosen to permit signal distortion for the purpose of obtaining improved interference cancellation. The SNR is established to be a non-decreasing function of the chosen distortion level assuming the signal direction and spectral shape are known. Significant improvements in SNR can result from permitting nonzero levels of distortion. SNR improvement only occurs for broad-band beamforming; the SNR is shown to be constant for all distortion levels in narrow-band beamforming. Several performance bounds are derived.

A sidelobe cancelling structure is derived for the SCM beamformer and a brief discussion of adaptive algorithm issues is given. The distortion presented to desired signals is easily computed from the beamformer weights and can be equalized after beamforming. In general, equalization changes the SNR.

APPENDIX

Theorem: The SNR as defined in (12) and (17) is a nonincreasing function of λ on the interval $0 \leq \lambda \leq \infty$.

Proof: Let SNR_1 and SNR_2 correspond to λ_1 and λ_2 , respectively, and assume $\lambda_1 < \lambda_2$. The theorem is true iff for all λ_1 and λ_2 we have

$$\rho = \frac{\text{SNR}_1}{\text{SNR}_2} = \frac{\sum_{i=1}^N g_i \gamma_i h_i^1 \sum_{j=1}^N g_j h_j^2}{\sum_{i=1}^N g_i h_i^1 \sum_{j=1}^N g_j \gamma_j h_j^2} \geq 1 \quad (\text{A1})$$

where superscripts 1 and 2 on h_i indicate (15b) evaluated at λ_1 and λ_2 . Rewrite ρ as

$$\rho = \frac{\mathbf{g}^T \mathbf{\Gamma} \mathbf{h}^1 \mathbf{h}^{2T} \mathbf{g}}{\mathbf{g}^T \mathbf{\Gamma} \mathbf{h}^2 \mathbf{h}^{1T} \mathbf{g}} \quad (\text{A2})$$

with \mathbf{g} and \mathbf{h}^k vectors having elements g_i and h_i^k . Superscript T denotes matrix transpose. Now $\rho \geq 1$ iff $\mathbf{g}^T \mathbf{A} \mathbf{g} \geq 0$ where

$$\mathbf{A} = \mathbf{\Gamma} \mathbf{h}^1 \mathbf{h}^{2T} - \mathbf{\Gamma} \mathbf{h}^2 \mathbf{h}^{1T}. \quad (\text{A3})$$

However, $\mathbf{g}^T \mathbf{A} \mathbf{g} \geq 0$ is equivalent to $\mathbf{g}^T (\mathbf{A} + \mathbf{A}^T) \mathbf{g} \geq 0$. We now complete the proof by establishing that the elements of $\mathbf{A} + \mathbf{A}^T$ are nonnegative. Note that this is sufficient to show $\mathbf{g}^T (\mathbf{A} + \mathbf{A}^T) \mathbf{g} \geq 0$ since the elements of \mathbf{g} are nonnegative.

After some simple algebra we obtain

$$[\mathbf{A} + \mathbf{A}^T]_{ij} = (\gamma_i - \gamma_j)(h_i^1 h_j^2 - h_i^2 h_j^1). \quad (\text{A4})$$

The signs of both terms in parentheses must be the same for the elements of $\mathbf{A} + \mathbf{A}^T$ to be nonnegative. Again, straightforward manipulation indicates that the sign of $h_i^1 h_j^2 - h_i^2 h_j^1$ is identical to the sign of

$$[1 + (\sigma_s + \lambda_2) \gamma_i]^2 [1 + (\sigma_s + \lambda_1) \gamma_j]^2 - [1 + (\sigma_s + \lambda_1) \gamma_i]^2 [1 + (\sigma_s + \lambda_2) \gamma_j]^2. \quad (\text{A5})$$

Equation (A5) is the form of $a^2 - b^2 = (a + b)(a - b)$. Since a and b are both positive we need only evaluate the term $a - b$ which simplifies to

$$(\lambda_2 - \lambda_1)(\gamma_i - \gamma_j). \quad (\text{A6})$$

Thus, the sign of $[\mathbf{A} + \mathbf{A}^T]_{ij}$ (see A4)) is equivalent to the sign of

$$(\lambda_2 - \lambda_1)(\gamma_i - \gamma_j)^2. \quad (\text{A7})$$

Since $\lambda_2 > \lambda_1$, we have established that $\mathbf{A} + \mathbf{A}^T$ is a matrix of nonnegative elements. ■

REFERENCES

- [1] O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, pp. 926-935, Aug. 1972.
- [2] K. M. Buckley, "Spatial/spectral filtering with linearly constrained minimum variance beamformers," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 249-266, Mar. 1987.
- [3] Y. Kaneda and J. Ohga, "Adaptive microphone-array system for noise reduction," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1391-1400, Dec. 1986.
- [4] M. M. Sondhi and G. W. Elko, "Adaptive optimization of microphone arrays under a nonlinear constraint," in *Proc. ICASSP 86*, vol. 2, 1986, pp. 981-984.
- [5] D. Chazan, Y. Medan, and U. Shvadron, "Noise cancellation for hearing aids," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, pp. 1697-1705, Nov. 1988.
- [6] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1365-1376, Oct. 1987.
- [7] M. H. Er and A. Cantoni, "An alternative formulation for an optimum beamformer with robustness capability," *Proc. Inst. Elec. Eng.*, vol. 132, pt. F, pp. 447-460, Oct. 1985.
- [8] N. L. Owsley, "A recent trend in adaptive spatial processing for sensor arrays: Constrained adaptation," in *Signal Processing*, J. W. R. Griffiths *et al.*, Eds. New York: Academic, 1973, pp. 591-604.
- [9] L. P. Winkler and M. Schwartz, "Constrained array optimization by penalty function techniques," *J. Acoust. Soc. Amer.*, vol. 55, pp. 1042-1048, May 1974.
- [10] L. L. Scharf and D. W. Tufts, "Rank reduction for modeling stationary signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 350-355, Mar. 1987.
- [11] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*. Englewood Cliffs, NJ: Prentice-Hall, 1974.
- [12] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: John Hopkins University Press, 1983.
- [13] L. Elden, "Algorithms for the regularization of ill-conditioned least squares problems," *Bit*, vol. 17, pp. 134-145, 1977.
- [14] W. Gander, "Least squares with a quadratic constraint," *Numer. Math.*, vol. 36, pp. 291-307, 1981.
- [15] L. Elden and R. Schriber, "An application of systolic arrays to linear discrete ill-posed problems," *SIAM J. Sci. Stat. Comput.*, vol. 7, pp. 892-903, July 1986.
- [16] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 27-34, Jan. 1982.
- [17] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 772-781, Sept. 1978.
- [18] M. H. Er and A. Cantoni, "A new approach to the design of broadband element space antenna array processors," *IEEE J. Oceanic Eng.*, vol. OE-10, pp. 231-240, July 1985.



Barry D. Van Veen (S'81-M'83) was born in Green Bay, WI, on June 20, 1961. He received the B.S. degree from Michigan Technological University in 1983 and the Ph.D. degree from the University of Colorado in 1986, both in electrical engineering.

In spring 1987 he was with the Department of Electrical and Computer Engineering, University of Colorado-Boulder. Since August 1987 he has been with the Department of Electrical and Computer Engineering at the University of Wisconsin-

Madison and is currently an Associate Professor. His research interests include sensor array processing, spectrum estimation, and adaptive algorithms.

Dr. Van Veen was a recipient of the 1989 Presidential Young Investigator Award and the 1990 Signal Processing Society Paper Award.