mation is easier for strongly scattering objects than for weakly scattering objects.

References

[1] A. J. Devaney and G. A. Tsihrintzis, "Maximum likelihood estimation of object location in diffraction tomography," *Trans. Signal Processing*, vol. 39, no. 3, pp. 672-682, Mar. 1991.

[2] G. A. Tsihrintzis and A. J. Devaney, "Estimation of object location from diffraction tomographic intensity data," submitted for publication.

[3] R. G. Newton, *Scattering Theory of Waves and Particles*. New York: McGraw-Hill, 1966.

[4] A. J. Devaney, "Reconstructive tomography with diffracting wavefields," *Inverse Problems*, vol. 2, pp. 161-183, 1986.

[5] A. J. Devaney, "A computer simulation study of diffraction tomography," *IEEE Trans. Biomed. Eng.*, vol. BME-30, pp. 377-386, 1982.

[6] A. J. Devaney, "The limited view problem in diffraction tomography," *Inverse Problems*, vol. 5, pp. 501-521, 1989.

[7] D. J. Rossi and A. S. Willsky, "Reconstruction from projections based on detection and estimation of objects, Parts I and II: Performance analysis and robustness analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 886-906, 1984.

[8] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics, Part I*. New York: McGraw-Hill, 1953.

[9] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*. New York: Wiley, 1968.

[10] D. Duchene, D. Lesselier, and W. Tabbara, "Acoustic imaging of 2d fluid targets buried in a half-space: A diffraction tomography approach," *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, vol. UFFC-34, pp. 540-549, 1987.

Adaptive Convergence of Linearly Constrained Beamformers Based on the Sample Covariance Matrix

Barry D. Van Veen

Abstract—**A statistical analysis of the adaptive convergence behavior of linearly constrained beamformers is given, assuming the sample covariance estimator is used to estimate the covariance matrix. The sensor data is assumed to be Gaussian distributed and independent from data vector to data vector. The output power and mean-squared error in the absence of the desired signal are shown to be multiples of chi-squared random variables. The presence of the desired signal results in an excess mean-squared error that is beta distributed and depends only on the signal power, number of data vectors, and number of adaptive degrees of freedom. The expected value of the excess mean-squared error resulting from the signal presence is directly proportional to the signal power and number of adaptive degrees of freedom, but is inversely proportional to the number of data vectors.**

I. Introduction

Linearly constrained beamforming is a powerful and versatile method of spatial filtering. The weights in a linearly constrained beamformer are a function of the data covariance matrix which is usually unknown. One common estimate of the covariance matrix is the sample covariance matrix. The sample covariance matrix is the maximum likelihood estimate given no prior constraints on the covariance matrix [1]. The beamformer output is a function of the covariance matrix estimate so it is a random variable with distribution dependent on the statistics of the covariance matrix estimate. In this correspondence, the distributions of the output power, mean-squared error (MSE) in the absence of the desired signal, and the excess MSE due to the presence of a desired signal are derived.

Several investigators have studied the convergence characteristics of adaptive beamformers that utilize the sample covariance matrix inversion algorithm (commonly referred to as the SMI algorithm). In a well-known paper, Reed *et al.* [2] derive the distribution of a normalized signal-to-noise ratio (SNR) assuming the beamformer weights are based on signal free data vectors. Capon and Goodman [3] derive the distribution of the output power for a minimum variance beamformer subject to a single linear constraint. Monzingo and Miller [4] treat adaptive convergence of SNR and MSE for various configurations of beamformers commonly used in narrow-band processing; several of these are equivalent to a minimum variance beamformer subject to a single linear constraint. Additional related work is found in [5]-[8]. This correspondence addresses the general linearly constrained beamforming problem. The results for the output power and MSE in the absence of the desired signal represent generalizations of several of the results cited above. The results for the excess MSE due to the signal presence appear completely new. These results clearly show the adaptive convergence advantages of partially adaptive beamforming [10].

Section II reviews the linearly constrained beamforming problem and the generalized sidelobe canceller (GSC) implementation. Expressions for the output power and the MSE are derived in Section III using the sample covariance matrix estimate. The corresponding probability distributions are derived in Section IV. A brief discussion of the results is given in Section V.

II. Linearly Constrained Beamforming and the Generalized Sidelobe Canceller

Let the beamformer output at time n, $y(n)$, be the inner product of an N-dimensional weight vector $\boldsymbol{w}$ with an N-dimensional data vector $\boldsymbol{x}(n)$

$$y(n) = \boldsymbol{w}^H \boldsymbol{x}(n). \tag{1}$$

If tap delay lines are used in the sensor channels, then we assume the data at the taps and the corresponding weights are represented in $\boldsymbol{x}(n)$ and $\boldsymbol{w}$. For notation, we use boldface lowercase and uppercase symbols to represent vectors and matrices, respectively. Superscript ^{H} denotes complex conjugate transpose. The linearly constrained minimum variance technique chooses $\boldsymbol{w}$ to minimize the output variance (power) $E\{|y(n)|^2\} = \boldsymbol{w}^H \boldsymbol{R} \boldsymbol{w}$ where $\boldsymbol{R} = E\{\boldsymbol{x}\boldsymbol{x}^H\}$ subject to a set of L linear constraints $\boldsymbol{C}^H \boldsymbol{w} = \boldsymbol{f}$

$$\min_{\boldsymbol{w}} \boldsymbol{w}^H \boldsymbol{R} \boldsymbol{w} \quad \text{subject to } \boldsymbol{C}^H \boldsymbol{w} = \boldsymbol{f}. \tag{2}$$

The N by L constraint matrix $\boldsymbol{C}$, and L by one response vector $\boldsymbol{f}$, are designed to control the beamformer response over direction and/or frequency [11].

The GSC is an alternate but equivalent formulation of the linearly constrained minimum variance beamformer [12]. In the GSC,

Manuscript received June 25, 1990; revised December 19, 1990. This work was supported in part by the Army Research Office under Grant DAAL03-89-K-0141 and by the National Science Foundation under Award MIP-8958559.

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

IEEE Log Number 9143785.

1053-587X/91/0600-1470$01.00

w is decomposed into two orthogonal components, one that lies in the space spanned by columns of C, denoted as w_q, and another that is orthogonal to the space spanned by the columns of C, represented as $C_n w_n$. The N by $N - L$ matrix C_n satisfies det $[C\ C_n] \neq 0$, and $C^H C_n = 0$. It is straightforward to show that $w_q = C(C^H C)^{-1} f$ to satisfy the constraint. The orthogonality of C and C_n implies that the constraint is satisfied independent of the value of the $N - L$ dimensional weight vector w_n. Now $w = w_q - C_n w_n$ and the GSC equivalent of (2) is

$$\min_{w_n} (w_q - C_n w_n)^H R (w_q - C_n w_n). \tag{3}$$

The solution to (3) is given by

$$w_n = (C_n^H R C_n)^{-1} C_n^H R w_q. \tag{4}$$

In partially adaptive beamforming the number of adaptive degrees of freedom is reduced. A partially adaptive GSC has the representation $w = w_q - C_n T_n w_n$ where T_n is a rank K, $N - L$ by K $(K < N - L)$ dimensioned matrix that determines which of the available degrees of freedom are adaptive [10]. w_n is now of dimension K. Note that reducing the adaptive degrees of freedom is equivalent to adding constraints. The analysis below illustrates the adaptive convergence advantages associated with reducing the number of adaptive degrees of freedom. Let $T_a = C_n T_n$ for notational convenience. Thus, $w = w_q - T_a w_n$.

We assume that a portion of the constraints are used to control the beamformer's response to the desired signal. This implies that the signal lies in the space spanned by the columns of C and that C_n is orthogonal to the signal; hence, C_n is sometimes termed the "signal blocking matrix" [12] and satisfies $C_n^H R_s = 0$, where R_s is the signal covariance matrix. Thus, $T_a^H R_s = 0$. Let $R = R_s + R_n$ be the sum of signal R_s and noise R_n covariance matrices where the signal and noise are assumed to be independent. The orthogonality of T_a and R_s implies

$$w_n = (T_a^H R_n T_a)^{-1} T_a^H R_n w_q. \tag{5}$$

The adaptive weight vector w_n is independent of the signal. The output power or variance is given by

$$P = w_q^H R w_q - w_q^H R T_a (T_a^H R T_a)^{-1} T_a^H R w_q \tag{6a}$$

or equivalently, $P = P_s + P_n$ where

$$P_s = w_q^H R_s w_q \tag{6b}$$

$$P_n = w_q^H R_n w_q - w_q^H R_n T_a (T_a^H R_n T_a)^{-1} T_a^H R_n w_q \tag{6c}$$

are the ouput powers due to the signal and the noise, respectively.

In order to derive the MSE we assume that the constraints are chosen to pass the desired signal without distortion, e.g., with unit gain and linear phase. Let $x(n) = x_s(n) + x_n(n)$ where $x_s(n)$ and $x_n(n)$ are the components of the data due to the desired signal and noise, respectively. The constraints imply that $w_q^H x_s(n) = s(n)$ where $s(n)$ is the desired signal. Equivalently, the beamformer output is equal to the desired signal in the absence of noise $(x_n(n) = 0)$. We also have $T_a^H x_s(n) = 0$ since $x_s(n)$ must lie in the space spanned by the columns of R_s. The random nature of $x_s(n)$ only affects where a given observation lies within this column space. Define the MSE as $e = E\{|s(n) - y(n)|^2\}$. Now,

$$\begin{aligned} e &= E\{|s(n) - w^H x(n)|^2\} \\ &= E\{|w^H x_n(n)|^2\} \\ &= w^H R_n w \\ &= w_q^H R_n w_q - w_q^H R_n T_a (T_a^H R_n T_a)^{-1} T_a^H R_n w_q. \end{aligned} \tag{7}$$

w is assumed known so the expectation in (7) is taken with respect to $x_n(n)$. The last line in (7) is obtained by subsituting $w = w_q - T_a w_n$ with w_n given by (5). Note that $e = P_n$.

III. MSE and Output Power Using the Sample Covariance Matrix Estimate for R

In practice R is unknown and is estimated from the data. Assume there are M data vectors $x(n)$, $n = 1, 2, \cdots, M$ available. The sample covariance matrix estimate of R is

$$\hat{R} = \frac{1}{M} \sum_{n=1}^{M} x(n)\, x^H(n). \tag{8}$$

Define the data matrix $X = [x(1)\, x(2) \cdots x(M)]$ and let $X = X_s + X_n$ where X_s is the data due to the desired signal and X_n the data due to the noise. Now

$$\begin{aligned} \hat{R} &= \frac{1}{M} X X^H \\ &= \frac{1}{M} [X_s X_s^H + X_s X_n^H + X_n X_s^H + X_n X_n^H]. \end{aligned} \tag{9}$$

Substitute (9) in (4) after replacing C_n by T_a. Use the orthogonality between T_a and the signal, i.e., $T_a^H X_s = 0$ to obtain

$$\begin{aligned} w_n &= (T_a^H X X^H T_a)^{-1} T_a^H X X^H w_q \\ &= (T_a^H X_n X_n^H T_a)^{-1}\, T_a^H X_n (X_s^H + X_n^H) w_q. \end{aligned} \tag{10}$$

Let $s = [s(1)\, s(2) \cdots s(M)]^H$ be the column vector containing samples of the complex conjugate of the desired signal at times $n = 1, 2, \cdots, M$. The constraints imply that $w_q^H X_s = s^H$. Let $y = [y(1)\, y(2) \cdots y(M)]^H$ be a column vector representing the complex conjugate of the beamformer output. Now $y^H = w^H X = (w_q - T_a w_n)^H X$. Define the sample MSE $\hat{e}$ as

$$\begin{aligned} \hat{e} &= \frac{1}{M} |s^H - y^H|^2 \\ &= \frac{1}{M} |s^H - w^H X|^2. \end{aligned} \tag{11}$$

Using $T_a^H X_s = 0$ and $w_q^H X_s = s^H$ gives

$$\begin{aligned} \hat{e} &= \frac{1}{M} |w^H X_n|^2 \\ &= \frac{1}{M} w^H X_n X_n^H w. \end{aligned} \tag{12}$$

Substitute the GSC representation for w into (12) with w_n given in (10) and simplify to obtain

$$\begin{aligned} \hat{e} = \frac{1}{M} [&w_q^H X_n X_n^H w_q - w_q^H X_n X_n^H T_a (T_a^H X_n X_n^H T_a)^{-1} T_a^H X_n X_n^H w_q \\ &+ w_q^H X_s X_n^H T_a (T_a^H X_n X_n^H T_a)^{-1} T_a^H X_n X_s^H w_q]. \end{aligned} \tag{13}$$

Define

$$\hat{e}_n = \frac{1}{M} [w_q^H X_n X_n^H w_q - w_q^H X_n X_n^H T_a (T_a^H X_n X_n^H T_a)^{-1} T_a^H X_n X_n^H w_q] \tag{14a}$$

$$\hat{e}_s = \frac{1}{M} [w_q^H X_n X_n^H T_a (T_a^H X_n X_n^H T_a)^{-1} T_a^H X_n X_n^H w_q]$$

$$= \frac{1}{M} [s^H X_n^H T_a (T_a^H X_n X_n^H T_a)^{-1} T_a^H X_n s] \tag{14b}$$

so that $\hat{e} = \hat{e}_n + \hat{e}_s$. Note that $\hat{e}_n$ is the MSE due to the noise, that is $\hat{e}_n$ is the MSE in the absence of the signal ($X_s = 0$) and corresponds to e (7) with R_n replaced by $M^{-1}X_nX_n^H$. $\hat{e}_s$ is an additional MSE that results because $X_sX_n^H$ and $X_nX_s^H$ are nonzero.

The expression for the output power is obtained by substituting $M^{-1}XX^H$ for R in (6a)

$$\hat{P} = \frac{1}{M}[w_q^H XX^H w_q - w_q^H XX^H T_a(T_a^H XX^H T_a)^{-1} \cdot T_a^H XX^H w_q]. \tag{15}$$

Note that this expression is not equivalent to $\hat{P}_s + \hat{P}_n$ with $\hat{P}_s$ and $\hat{P}_n$ obtained from (6b) and (6c) by substituting $R_s = M^{-1}X_sX_s^H$ and $R_n = M^{-1}X_nX_n^H$ because of the nonzero cross terms $X_sX_n^H$ and $X_nX_s^H$.

IV. Probability Density Functions of $\hat{e}_n$, $\hat{e}_s$, and $\hat{P}$

The columns of X and X_n are assumed to be independent and identically Gaussian distributed with zero mean. Under these assumptions XX^H and X_nX^H are complex Wishart distributed [1], with distribution denoted as $W(M, N; R)$ and $W(M, N; R_n)$, respectively. Here M is the number of columns in $X(X_n)$, N is the number of rows, and $R(R_n)$ is the covariance matrix associated with the columns of $X(X_n)$. Muirhead [9] is used as a reference for most of the properties of the Wishart distribution needed below. Muirhead [9] only considers real random variables; however, it is straightforward to extend these properties to the complex case (e.g., see [7, appendix 2]).

Consider $\hat{P}$. Define the $K + 1$ by $K + 1$ matrices A and $\hat{A}$ as

$$A = [w_q \quad T_a]^H R [w_q \quad T_a]$$
$$\hat{A} = [w_q \quad T_a]^H XX^H [w_q \quad T_a] = \begin{bmatrix} w_q^H XX^H w_q & w_q^H XX^H T_a \\ T_a^H XX^H w_q & T_a^H XX^H T_a \end{bmatrix}. \tag{16}$$

Use the identity for the inverse of a partitioned matrix [13] to express the element in the first row and column of $\hat{A}^{-1}$ as

$$[\hat{A}^{-1}]_{1,1} = [w_q^H XX^H w_q - w_q^H XX^H T_a(T_a^H XX^H T_a)^{-1} T_a^H XX^H w_q]^{-1}. \tag{17}$$

This requires $M \geq K + 1$. Thus, (15) is expressed as

$$\hat{P}_n = \frac{1}{M}[[\hat{A}^{-1}]_{1,1}]^{-1} = \frac{1}{M}[u_1^H \hat{A}^{-1} u_1]^{-1} \tag{18}$$

where $u_1 = [1\ 0\ 0 \cdots 0]^H$. Similarly, (6a) is written as $P = [u_1^H A^{-1} u_1]^{-1}$.

Using [9, theorem 3.2.5] we see that $\hat{A}$ is distributed as $W(M, K + 1; A)$. Now apply [9, theorem 3.2.11] to show that $M\hat{P}$ is distributed by $W(M - K, 1; [u_1^H A^{-1} u_1]^{-1}) = W(M - K, 1; P)$. A random variable distributed as $W(M - K, 1; 1)$ is one half a chi-squared random variable with $(M - K)$ complex degrees of freedom ($2(M - K)$ real degrees of freedom). Thus, the mean of $\hat{P}$ is given by

$$E\{\hat{P}\} = \frac{M - K}{M} P. \tag{19}$$

The factor $(M - K)/M$ represents the loss due to estimation of R and determines the adaptive convergence behavior of the mean output power when viewed as a function of M.

Capon's ML estimator [14] is equivalent to the output power of a beamformer using a single linear constraint to ensure distortionless response at a specified frequency and direction. The distribution of $\hat{P}$ derived here is the multiple constraint equivalent of the distribution for Capon's ML estimator given in [3].

Next, consider the distribution of $\hat{e}_n$, the MSE due to the noise. The expression for $\hat{e}_n$ (14a) is identical to the expression for $\hat{P}$ (15), if X_n is replaced by X. Thus, $M\hat{e}_n$ is distributed as $W(M - K, 1; e)$ where e is given in (7). The mean of $\hat{e}_n$ is

$$E\{\hat{e}_n\} = \frac{M - K}{M} e. \tag{20}$$

Again $(M - K)/M$ determines the adaptive convergence of the mean when viewed as a function of M.

Lastly, consider the distribution of $\hat{e}_s$. First, the conditional distribution of $\hat{e}_s$ given s is obtained. Define $T_a^H X_n = V$ so

$$\hat{e}_s = \frac{1}{M}[s^H V^H (VV^H)^{-1} Vs]. \tag{21}$$

Let $E\{VV^H\} = MT_a^H R_n T_a$ have Cholesky factorization TT^H and rewrite $\hat{e}_s$ as

$$\hat{e}_s = \frac{1}{M}[s^H V^H T^{-H} T^H (VV^H)^{-1} TT^{-1} Vs]$$
$$= \frac{1}{M}[s^H V^H T^{-H} (T^{-1} VV^H T^{-H})^{-1} T^{-1} Vs]$$
$$= \frac{1}{M}[s^H U^H (UU^H)^{-1} Us]. \tag{22}$$

The columns of $U = T^{-1}V$ are i.i.d. Gaussian distributed random vectors with $E\{UU^H\} = T^{-1}E\{VV^H\}T^{-H} = I$. This implies that $\hat{e}_s$ is independent of R_n and the elements of T_a. Let $s = as_0$ where $s_0^H s_0 = 1$ and define $q = Us_0$. Given s, q is a Guassian random vector with mean $E\{q\} = 0$ and covariance $E\{qq^H\} = I$. Substituting in (22) we obtain

$$\hat{e}_s = \frac{a^2}{M}[q^H (UU^H)^{-1} q] = \frac{a^2}{M}\rho \tag{23}$$

where $\rho = q^H (UU^H)^{-1} q$.

We shall now show that ρ is a beta distributed random variable and is independent of s. Define an M by M dimensional unitary matrix $H = [G\ s_0]$, i.e., $G^H G = I$ and $G^H s_0 = 0$. Let the matrix $B = UH = [Z\ q]$ where $Z = UG$ is a K by $M - 1$ dimensional matrix. Now, $UU^H = BB^H = ZZ^H + qq^H$ so

$$\rho = [(q^H (ZZ^H + qq^H)^{-1} q]. \tag{24}$$

Assuming $M > K$, ZZ^H is nonsingular with probability one. Application of the matrix inversion lemma to (24) yields

$$\rho = q^H (ZZ^H)^{-1} q - \frac{q^H (ZZ^H)^{-1} qq^H (ZZ^H)^{-1} q}{1 + q^H (ZZ^H)^{-1} q} = \frac{1}{1 + \frac{1}{q^H (ZZ^H)^{-1} q}}. \tag{25}$$

Define $\alpha = q^H q$ and $\gamma = q^H q / q^H (ZZ^H)^{-1} q$ so that

$$\rho = \frac{1}{1 + \frac{\gamma}{\alpha}}. \tag{26}$$

α is a complex chi-squared random variable with K complex degrees of freedom. γ is also a complex chi-squared random variable. Recall that the columns of $\boldsymbol{U}$ are independent. The columns of $\boldsymbol{B}$ are also independent and distributed identically to the columns of $\boldsymbol{U}$ since $\boldsymbol{B} = \boldsymbol{UH}$ is obtained from $\boldsymbol{U}$ via a unitary transformation (see, e.g., [7, Appendix A, part F]). The columns of $\boldsymbol{Z}$ and $\boldsymbol{q}$ are independent and Gaussian distributed with covariance $\boldsymbol{I}$; thus, [9, theorem 3.2.12] is applicable to γ. This theorem says that γ is a complex chi-squared random variable with $M - K$ degrees of freedom and is independent of $\boldsymbol{q}$. Furthermore, γ and α are independent of $\boldsymbol{s}$ so ρ is independent of $\boldsymbol{s}$. Now γ/α is a complex F distributed random variable. A simple change of variables indicates that ρ is complex beta distributed

$$p(\rho) = \frac{(M-1)!}{(M-K-1)!\,(K-1)!}\,\rho^{K-1}(1-\rho)^{M-K-1}. \tag{27}$$

The distribution of $\hat{e}_s$ given $\boldsymbol{s}$ is thus

$$p(\hat{e}_s \mid \boldsymbol{s}) = \frac{M(M-1)!}{a^2(M-K-1)!\,(K-1)!} \cdot \left(\frac{M\hat{e}_s}{a^2}\right)^{K-1}\left(1 - \frac{M\hat{e}_s}{a^2}\right)^{M-K-1}. \tag{28}$$

The dependence on $\boldsymbol{s}$ is only through $a^2 = \boldsymbol{s}^H\boldsymbol{s}$ so, in principle, the distribution of $\hat{e}_s$ is obtained by integrating the product of $p(\hat{e}_s \mid \boldsymbol{s})$ and the distribution of a^2 over a^2. However, the moments of $\hat{e}_s$ are, in general, of greater interest than the distribution itself. It is straightforward in this case to obtain the moments of $\hat{e}_s$ because ρ and a^2 are independent. For example, the mean of $\hat{e}_s$ is

$$E\{\hat{e}_s\} = M^{-1}E\{a^2\}\,E\{\rho\} = \sigma_s^2\frac{K}{M}. \tag{29}$$

where σ_s^2 is the variance or power of the signal. Equation (29) indicates that the average MSE associated with the signal presence is directly proportional to the signal power and the number of adaptive degrees of freedom, but inversely proportional to the number of data vectors. Note that the presence of a strong signal results in large MSE.

V. Discussion

The expected values of the output power and MSE due to the noise are within 3 dB of the optimum values after $M = 2K$ data vectors, while the expected value of the excess MSE due to the signal presence is down by 3 dB after $M = 2K$ data vectors. These results clearly indicate the benefits of reducing the number of adaptive degrees of freedom K: the beamformer output is defined with fewer data vectors ($M > K$), and faster convergence to the optimum is obtained. The disadvantage of reducing K is an increase in the asymptotic MSE e and noise output power P_n. This increase represents a loss in performance associated with reducing K and indicates that T_a should be designed to minimize P_n or equivalently e as suggested in [10].

References

[1] N. R. Goodman, "Statistical analysis based on a certain multivariate Gaussian distribution," *Ann. Math. Stat.*, vol. 34, pp. 152–177, Mar. 1963.
[2] I. S. Reed, J. D. Mallet, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-10, pp. 853–863, Nov. 1974.
[3] J. Capon and N. R. Goodman, "Probability distributions for estimators of the frequency-wavenumber spectrum," *Proc. IEEE*, vol. 58, pp. 1785–1786, Oct. 1970.
[4] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays.* New York: Wiley, 1980, ch. 6.
[5] A. B. Baggeroer, "Confidence intervals for regression (MEM) spectral estimates," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 534–545, Sept. 1976.
[6] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, pp. 115–127, Mar. 1986.
[7] E. J. Kelly and K. M. Forsythe, "Adaptive detection and parameter estimation for multidimensional signal model," M.I.T. Lincoln Lab., Tech. Rep. 848, Apr. 1989.
[8] M. W. Ganz, R. L. Moses, and S. L. Wilson, "Convergence of the SMI and the diagonally loaded SMI algorithms with weak interference," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 394–399, Mar. 1990.
[9] R. J. Muirhead, *Aspects of Multivariate Statistical Theory.* New York: Wiley, 1982, ch. 3.
[10] B. D. Van Veen and R. A. Roberts, "Partially adaptive beamformer design via output power minimization," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. ASSP-35, pp. 1524–1532, Nov. 1987.
[11] 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.
[12] 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.
[13] T. Kailath, *Linear Systems.* Englewood Cliffs, NJ: Prentice-Hall, 1980, Appendix.
[14] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, pp. 1408–1418, Aug. 1969.

Direction Finding Using ESPRIT with Interpolated Arrays

Anthony J. Weiss and Motti Gavish

Abstract—**The technique of interpolated arrays is applied to ESPRIT-type direction finding methods. The resulting method uses sensor arrays with an arbitrary configuration, thus eliminating the basic restrictive requirement of ESPRIT for two (or more) identical arrays. Our approach allows for resolving *D* narrow-band signals if the number of sensors is, at least, *D* + 1, while the original ESPRIT method requires at least 2*D* sensors. Moreover, it is shown that while ESPRIT performs poorly for signals propagating in parallel (or close to parallel) with the array displacement vector, the advocated technique does not exhibit such weakness. Finally, using two subarrays, ESPRIT cannot resolve azimuth and elevation even when the sensors are not collinear. However, the interpolated ESPRIT procedure resolves azimuth and elevation using only a single array. All the above mentioned advantages are obtained with a reasonable increase of computation load, thus preserving the basic and most outstanding advantage of ESPRIT. We also discuss, and illustrate numerically, the performance of the original ESPRIT when the sensor locations are perturbed. It is shown**

Manuscript received May 25, 1990; revised December 16, 1990.
A. J. Weiss is with the Department of Electronic Systems, Faculty of Engineering, Tel-Aviv University, Ramat-Aviv 69978, Israel.
M. Gavish is with Communication Systems Division, ELTA Electronics Industries Ltd., Ashdod 77102, Israel, and with the Department of Electronic Systems, Faculty of Engineering, Tel-Aviv University, Ramat-Aviv 69978, Israel.
IEEE Log Number 9143786.