ADAPTIVE BAYESIAN BEAMFORMING FOR STEERING VECTOR UNCERTAINTIES WITH ORDER RECURSIVE IMPLEMENTATION

Chunwei Jethro Lam, Andrew C. Singer

University of Illinois at Urbana Champaign
Coordinated Science Laboratory
1308 West Main St., Urbana, IL 61801
217-244-9263
clam@uiuc.edu, acsinger@uiuc.edu

ABSTRACT

An order recursive algorithm for minimum mean square error (MMSE) estimation of signals under a Bayesian model defined on the steering vector is introduced. The MMSE estimate can be viewed as a mixture of conditional MMSE estimates weighted by the posterior probability density function (PDF) of the random steering vector given the observed data. This paper derives an adaptive closed form Kalman-filter implementation that updates the weight vector by successive incorporations of data collected from additional array elements in the steering vector. The performance of the Bayesian beamformer is compared against several robust beamformers in terms of mean square error (MSE) and output signal-to-interference-plus-noise ratio (SINR).

1. BACKGROUND

The received data vector of an N-element sensor array at sample time k has the form

$$\mathbf{x}(k) = \mathbf{a}\, s^*(k) + \mathbf{i}(k) + \mathbf{n}(k), \tag{1}$$

where $s(k)$ is the desired signal with known power σ_s^2, $\mathbf{a} \in \mathbb{C}^N$ is the steering vector, $\mathbf{i}(k)$ is the interence and $\mathbf{n}(k)$ is the noise. Let $R_{i+n} \triangleq E[(\mathbf{i}(k)+\mathbf{n}(k))(\mathbf{i}(k)+\mathbf{n}(k))^H]$ be the interference-plus-noise covariance. Let $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ be the complex conjugate, transpose and Hermitian transpose, respectively. Assume that $s(k)$, $\mathbf{i}(k)$ and $\mathbf{n}(k)$ are zero mean, temporally white, complex Gaussian processes that are mutually independent to each other.

In practice, the true steering vector often deviates from its presumed value for various reasons such as improper array modeling, asynchronous sampling, pointing error, miscalibration, or source motion. It is often reasonable to model these errors collectively as a random error vector associated some prior informations that are often available in statistical form. This modeling technique has been used in robust beamforming for direction-of-arrival (DOA) uncertainties in [1, 2] and steering vector uncertainties in [3] and [4].

Using the Bayesian approach, the steering vector is modeled as an $N \times 1$ complex Gaussian random vector

$$\mathbf{a} \backsim \mathcal{CN}(\bar{\mathbf{a}}, C), \tag{2}$$

with prior PDF

$$p(\mathbf{a}) \propto \exp\{-(\mathbf{a}-\bar{\mathbf{a}})^H C^{-1}(\mathbf{a}-\bar{\mathbf{a}})\} \tag{3}$$

where $\bar{\mathbf{a}}$ is the presumed steering vector and C is the error covariance. Based on this model, the conditional data covariance matrix is given by

$$R_{|\mathbf{a}} = E[\mathbf{x}(k)\mathbf{x}(k)^H|\mathbf{a}] = \sigma_s^2 \mathbf{a}\mathbf{a}^H + R_{i+n}. \tag{4}$$

Given a data block of K samples, $\mathbf{X} = [\mathbf{x}(1)\ldots\mathbf{x}(K)]$, the MMSE estimate of the signal vector, $\mathbf{s} = [s(1)\ldots s(K)]^T$, is given by the conditional mean of $\mathbf{s}$ given $\mathbf{X}$, which can be expressed as

$$\hat{\mathbf{s}} = E[\mathbf{s}|\mathbf{X}] = \int p(\mathbf{a}|\mathbf{X}) E[\mathbf{s}|\mathbf{X}, \mathbf{a}] d\mathbf{a}. \tag{5}$$

The MMSE estimate can be viewed as a combination of conditional MMSE estimates weighted according to the posterior PDF $p(\mathbf{a}|\mathbf{X})$. Since the signals are jointly, conditionally Gaussian, the conditional MMSE estimate can be obtained by a linear processor

$$E[\mathbf{s}|\mathbf{X}, \mathbf{a}] = \sigma_s^2\, \mathbf{X}^H R_{|\mathbf{a}}^{-1}\, \mathbf{a} \;\triangleq \mathbf{X}^H \mathbf{w}_{|\mathbf{a}}. \tag{6}$$

The processor $\mathbf{w}_{|\mathbf{a}}$ is also the Wiener filter for the conditional estimation problem. Substituting (6) into (5), the MMSE estimate is

$$\hat{\mathbf{s}} = \int p(\mathbf{a}|\mathbf{X}) \sigma_s^2 \mathbf{X}^H R_{|\mathbf{a}}^{-1}\, \mathbf{a}\, d\mathbf{a} \;\triangleq \mathbf{X}^H \mathbf{w}_B. \tag{7}$$

This work was supported in part by the National Science Foundation under grant number CCR-0092598 (career) and by NASA under grant number NAG2-2840.

where the weight vector

$$\mathbf{w}_B \triangleq \int p(\mathbf{a}|\mathbf{X})\sigma_s^2 R_{|\mathbf{a}}^{-1}\,\mathbf{a}\,d\mathbf{a} \tag{8}$$

is known as the Bayesian beamformer. The posterior PDF can be expressed as [1]

$$p(\mathbf{a}|\mathbf{X}) \propto \frac{p(\mathbf{a})}{(1+\sigma_s^2\mathbf{a}^H R_{i+n}^{-1}\mathbf{a})^K}\exp\left\{\frac{\sigma_s^2\|\mathbf{X}^H R_{i+n}^{-1}\mathbf{a}\|^2}{1+\sigma_s^2\mathbf{a}^H R_{i+n}^{-1}\mathbf{a}}\right\}. \tag{9}$$

2. APPROXIMATE BAYESIAN BEAMFORMER

The integral in the Bayesian beamformer is difficult to evaluate directly. In [1] and [2], the authors parameterize the steering vector error by a single random variable and develop several efficient algorithms. However, these algorithms do not apply to vector uncertainties. In [4], the authors express the integral in closed form by approximating the posterior PDF with the prior PDF and simplifying the integral with algebraic approximations. This method discards information in $p(\mathbf{a}|\mathbf{X})$, which leads to a significant loss of MMSE optimality. Moreover, the method depends on R_{i+n}, which is often not available.

The purpose of this paper is to provide a feasible adaptive algorithm for computing the Bayesian beamformer in (8). The main idea is to approximate the posterior PDF in (9) with some minor statistical assumptions such that the integral can be evaluated in closed form. First, the conditional data covariance $R_{|\mathbf{a}}$ is approximated by the sample covariance matrix

$$\hat{R} = \frac{1}{K}\mathbf{X}\mathbf{X}^H. \tag{10}$$

In practice, the sample covariance is often loaded with a small diagonal term to mitigate the approximation error [5]. With this approximation, the Bayesian beamformer is approximated as

$$\mathbf{w}_B = \sigma_s^2\hat{R}^{-1}\int p(\mathbf{a}|\mathbf{X})\ \mathbf{a}\ d\mathbf{a}. \tag{11}$$

Second, the variance of the random functional $\mathbf{a}^H R_{i+n}^{-1}\mathbf{a}$ is assumed to be sufficiently small such that it can be approximated by a constant. This condition is valid as soon as the expected projection of the steering vector onto the interference subspace is small, or equivalently, the interferers are located far away from the main lobe of the prior Gaussian distribution. With this approximation, the posterior PDF can be simplified as [2]

$$p(\mathbf{a}|\mathbf{X}) \propto p(\mathbf{a})\exp\{-K\sigma_s^2\mathbf{a}^H\hat{R}^{-1}\mathbf{a}\}. \tag{12}$$

Substituting the prior PDF in (3) into (12) and completing the square, the posterior PDF becomes a complex Gaussian PDF with posterior mean given as

$$\int p(\mathbf{a}|\mathbf{X})\ \mathbf{a}\ d\mathbf{a} \propto (K\sigma_s^2\hat{R}^{-1}+C^{-1})^{-1}C^{-1}\bar{\mathbf{a}}. \tag{13}$$

When K is zero, the posterior PDF is equal to the prior PDF, and both sides equal to $\bar{\mathbf{a}}$. This eliminates the scaling factor. Substituting the posterior mean into (11) gives

$$\mathbf{w}_B = \sigma_s^2(\hat{R}+K\sigma_s^2C)^{-1}\bar{\mathbf{a}}. \tag{14}$$

Thus, the approximate Bayesian beamformer belongs to the class of generalized covariance loading algorithms and it shares the same basic structure as the robust beamformers introduced in [6, 7, 8].

3. ORDER RECURSIVE IMPLEMENTATION

An order recursive algorithm is derived to compute the Bayesian beamformer in (14). For $n = 1, \ldots N$, let $\mathbf{e}_n$ be the nth column of the $N\times N$ identity matrix, and Φ_n be an $N\times n$ matrix such that

$$\Phi_n = [\mathbf{e}_1\ \mathbf{e}_2\ \ldots\ \mathbf{e}_n]. \tag{15}$$

For notational simplicity, define

$$\tilde{R} \triangleq \hat{R}+K\sigma_s^2C \tag{16}$$

$$Q_n \triangleq \Phi_n(\Phi_n^H\tilde{R}\Phi_n)^{-1}\Phi_n^H \tag{17}$$

$$\mathbf{u}_n \triangleq \sigma_s^2 Q_n\bar{\mathbf{a}}. \tag{18}$$

This yields a sequence of $N\times 1$ filters $\{\mathbf{u}_n\}_{n=1}^N$, of which the last filter is equal to the desired Bayesian beamformer, i.e., $\mathbf{u}_N = \mathbf{w}_B$. Note that $\mathbf{u}_n$ only operates on the reduced (lower dimensional) data block $\mathbf{X}_{1:n}$ and the presumed steering vector $\bar{\mathbf{a}}_{1:n}$ defined respectively by $\mathbf{X}_{1:n} \triangleq \Phi_n^H\mathbf{X}$ and $\bar{\mathbf{a}}_{1:n} \triangleq \Phi_n^H\bar{\mathbf{a}}$. The matrix Φ_n can be partitioned as

$$\Phi_n = [\Phi_{n-1}\ \mathbf{e}_n]. \tag{19}$$

Applying the block matrix inversion formula, the matrix Q_n can be expressed recursively as

$$Q_n = [\Phi_{n-1}\ \mathbf{e}_n]\begin{bmatrix}\Phi_{n-1}^H\tilde{R}\Phi_{n-1} & \Phi_{n-1}^H\tilde{R}\mathbf{e}_n\\ \mathbf{e}_n^H\tilde{R}\Phi_{n-1} & \mathbf{e}_n^H\tilde{R}\mathbf{e}_n\end{bmatrix}^{-1}\begin{bmatrix}\Phi_{n-1}^H\\ \mathbf{e}_n^H\end{bmatrix} \tag{20}$$

$$= Q_{n-1} + \frac{(\mathbf{I}-Q_{n-1}\tilde{R})\mathbf{e}_n\mathbf{e}_n^H(\mathbf{I}-\tilde{R}Q_{n-1})}{\mathbf{e}_n^H\tilde{R}(\mathbf{I}-Q_{n-1}\tilde{R})\mathbf{e}_n} \tag{21}$$

With (21), $\mathbf{u}_n$ can be expressed recursively as

$$\mathbf{u}_n = \mathbf{u}_{n-1} + \frac{\sigma_s^2\bar{a}_n - \mathbf{e}_n^H\tilde{R}\mathbf{u}_n{-}1}{\mathbf{e}_n^H\tilde{R}P_{n-1}\mathbf{e}_n}P_{n-1}\mathbf{e}_n, \tag{22}$$

where $\bar{a}_n \triangleq \mathbf{e}_n^H \bar{\mathbf{a}}$, $P_n \triangleq \mathbf{I} - Q_n \tilde{R}$ and $\mathbf{u}_0 \triangleq \mathbf{0}$. Similarly, the recursion of the oblique projector P_n is given by

$$P_n = P_{n-1} - \frac{P_{n-1}\mathbf{e}_n\mathbf{e}_n^H \tilde{R} P_{n-1}}{\mathbf{e}_n^H \tilde{R} P_{n-1}\mathbf{e}_n}. \tag{23}$$

with $P_0 \triangleq \mathbf{I}$. It can be seen from (22) and (23) that the beamformer has the form of a Kalman filter that is recursive in order instead of time. A survey of similar Kalman filter realizations for adaptive filters are presented in [9]. The proposed Kalman filter can be viewed as a multistage realization of the beamformer, in which the beamformer is decomposed into the span of a set of $\tilde{R}$-conjugate basis vectors generated from $\mathbf{e}_n$'s through successive oblique projections. In particular, the basis vectors are

$$\{\mathbf{e}_1,\ P_1\mathbf{e}_2,\ \ldots,\ P_{N-1}\mathbf{e}_N\}. \tag{24}$$

To realize the $\tilde{R}$-conjugacy, note that

$$\Phi_{n-1}\tilde{R}P_{n-1}\mathbf{e}_n = \mathbf{0} \tag{25}$$

for all n. The basis vector $P_{n-1}\mathbf{e}_n$ is thus $\tilde{R}$-conjugate to $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}$. Since all the previous basis vectors are linear combinations of $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}$, the basis vector $P_{n-1}\mathbf{e}_n$ is $\tilde{R}$-conjugate to the previous basis vectors. As this holds for all n, all the basis vectors are $\tilde{R}$-conjugate to each other.

Assuming the error covariance is diagonal, i.e.,

$$C = \begin{bmatrix} \delta_1 & & \\ & \ddots & \\ & & \delta_N \end{bmatrix}, \tag{26}$$

and replacing $\tilde{R}$ with $\hat{R} + K\sigma_s^2 C$, the recursion can be simplified into

$$\mathbf{u}_n = \mathbf{u}_{n-1} + \frac{\sigma_s^2 \bar{a}_n - \mathbf{e}_n^H \hat{R}\mathbf{u}_n - 1}{\mathbf{e}_n^H \hat{R} P_{n-1}\mathbf{e}_n + K\sigma_s^2\delta_n} P_{n-1}\mathbf{e}_n \tag{27}$$

$$P_n = P_{n-1} - \frac{P_{n-1}\mathbf{e}_n\mathbf{e}_n^H \hat{R} P_{n-1} + K\sigma_s^2\delta_n P_{n-1}\mathbf{e}_n\mathbf{e}_n^H}{\mathbf{e}_n^H \hat{R} P_{n-1}\mathbf{e}_n + K\sigma_s^2\delta_n}. \tag{28}$$

When δ_n is small, i.e., the uncertainty on a_n is small, the estimator puts more emphasis on the presumed value $\bar{a}_n$. When δ_n is large, the estimate depends more on the past recursions. The algorithm terminates after N iterations. Each iteration requires $O(n^3)$ flops, and can be modified to $O(N^2)$ by introducing an additional recursion for $\hat{R}P_n$. Note that even the notation $\hat{R}$, which represents $\frac{1}{K}\mathbf{X}\mathbf{X}^H$, is present in the formulas, the nth iteration only depends on the lower dimensional data block $\mathbf{X}_{1:n}$ instead of the whole data block $\mathbf{X}$.

One advantage of the proposed algorithm is its ability to reduce computational power while maintaining an accurate estimate. Based on the fact that the elements in a steering vector are often not equally perturbed, the algorithm can process the data from the sensors that are known *a priori* to have a small uncertainty level, and terminate the iteration prior to processing all the available data. The order of iteration may be chosen according to the error statistics about the steering vector. Note that the proposed algorithm does not necessarily outperform the cross-spectral reduced-rank Wiener filter or the multistage Wiener filter presented in [10]. It, however, provides an iteration algorithm to compute the MMSE estimate through successive incorporation of data from additional sensors without knowing *a priori* the size of the whole steering vector, N.

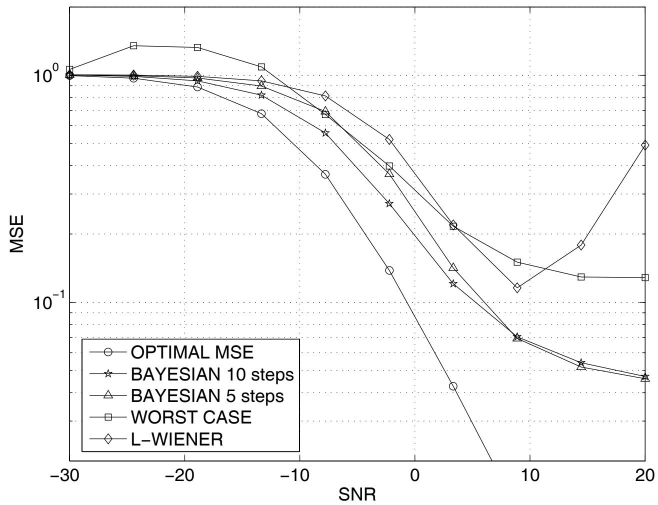

Fig. 1. MSE versus SNR.

4. SIMULATIONS

The performance of the Bayesian beamformer is compared against two robust beamformers, namely the loaded Wiener filter [5] and the robust beamformer with worst case optimization [6], in terms of both MSE and output SINR. A uniform linear array with 10 array elements is considered. The desired signal power is $\sigma_s^2 = 1$. The presumed pointing direction is 0^0 such that $\bar{\mathbf{a}} = [1, \ldots, 1]^T$. An interferer with power 30dB above the noise level and 30^0 away from the desired signal is present. The error covariance matrix C is assumed to be diagonal as in (26) with

$$\delta_n = \alpha e^n, \quad n = 1, \ldots, N, \tag{29}$$

where α is the scaling factor that ensures $\mathrm{tr}(C) = 1$. The optimal MSE and SINR are achieved by the Wiener filter in (6) with perfect knowledge of the steering vector. The loaded Wiener filter is given by

$$\mathbf{w} = \sigma_s^2(\hat{R} + \xi\mathbf{I})^{-1}\bar{\mathbf{a}} \tag{30}$$

where the loading level ξ is 10 times the noise level in a single sensor. The upper bound on uncertainty set of the worst

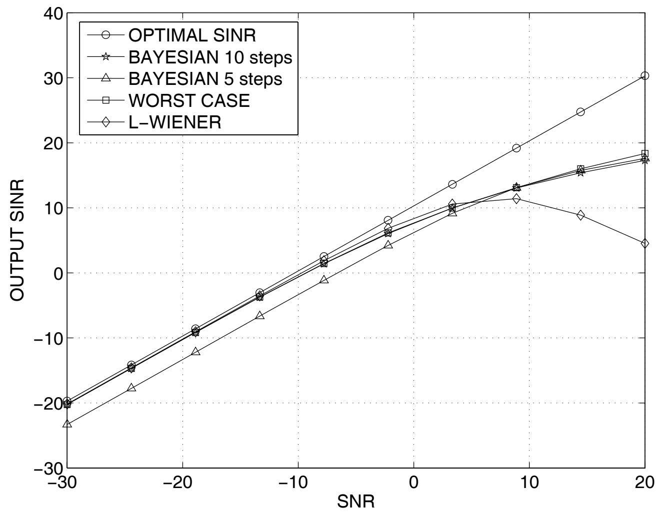

Fig. 2. Output SINR versus SNR.

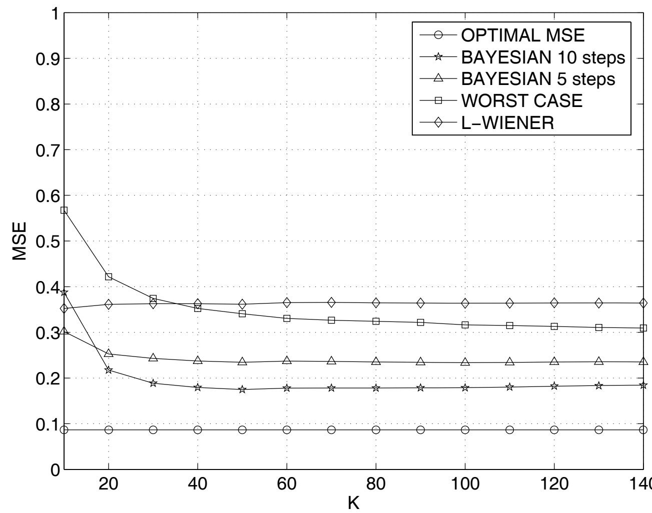

Fig. 3. MSE versus K.

case robust beamformer is set to $\epsilon = \sqrt{\mathrm{tr}(C)}$, i.e., the expected norm of the error vector. Two Bayesian beamformers are implemented, in which one terminates after 10 steps of the proposed algorithm and one terminates after 5 steps. A small diagonal loading term with 3dB above the noise level is added to all the beamformers. Each plot is obtained from 200 simulation runs.

Figure 1 and Figure 2 show the performance of the beamformers versus the signal-to-noise ratio (SNR) at $K = 50$ and Figure 3 shows the MSE of the beamformers against the data size K when the SNR is 0dB. The Bayesian beamformer gives the lowest MSE. Its output SINR is within 3dB below the worst case robust beamformer. The effectiveness of the Bayesian beamformer is mainly due to the incorporation of the prior statistics about the steering vector error and its MMSE nature. The performance of the 5-step Bayesian beamformer is close to the 10-step Bayesian beamformer in terms of both MSE and output SINR even if the computational cost is greatly reduced. This is because the data collected from the last 5 array elements are known in *a priori* to be highly defected, and thus their contributions in the 10-step beamformer do not induce any significant improvement in performance.

5. REFERENCES

[1] K. L. Bell, Y. Ephraim, and H. L. Van Trees, "A Bayesian approach to robust adaptive beamforming," *IEEE Trans. Signal Processing*, vol. 48, pp. 386–398, February 2000.

[2] C. J. Lam and A. C. Singer, "Bayesian beamforming for DOA uncertainties: Theory and implementation," *to appear in IEEE Trans. Signal Processing*.

[3] M. Viberg and A. Swindlehurst, "A Bayesian approach to auto-calibration for parametric array signal processing," *IEEE Trans. Signal Processing*, vol. 42, pp. 3495–3507, December 1994.

[4] O. Besson, A. A. Monakov, and C. Chalus, "Signal waveform estimation in the presence of uncertainties about the steering vector," *IEEE Trans. Signal Processing*, vol. 52, pp. 2432–2440, September 2004.

[5] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Processing*, pp. 1365–1376, October 1987.

[6] S. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: a solution to the signal mismatch problem," *IEEE Trans. Signal Processing*, vol. 51, pp. 313–324, February 2003.

[7] J. Li, P. Stoica, and Z. Wang, "Robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Processing*, vol. 51, pp. 172–1715, July 2003.

[8] R. G. Lorenz and P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Processing*, vol. 53, pp. 1684–1696, May 2005.

[9] A. H. Sayed and T. Kailath, "A state-space approach to adaptive RLS filtering," *IEEE Sig. Proc. Magazine*, pp. 18–60, July 1994.

[10] J. S. Goldstein, I. Reed, and L. L. Scharf, "A multistage representation of the wiener filter based on orthogonal projections," *IEEE Trans. Information Theory*, vol. 44, pp. 2943–2959, November 1998.