The Bayesian ABEL Bound on the Mean Square Error

Alexandre Renaux, Philippe Forster, Pascal Larzabal, Christ Richmond

THE BAYESIAN ABEL BOUND ON THE MEAN SQUARE ERROR

Alexandre Renaux[(1)], *Philippe Forster*[(2)], *Pascal Larzabal*[(1)] *and Christ Richmond*[(3)]

[(1)]ENS Cachan / SATIE(CNRS UMR8029)
61 av. du Président Wilson, 94235 Cachan cedex, France
[(2)]University Paris 10 / GEA
1, chemin Desvallieres, 92410 Ville d'Avray, France
[(3)] MIT / Lincoln Laboratory
244 Wood Street, Lexington, MA02420, USA

ABSTRACT

This paper deals with lower bound on the Mean Square Error (MSE). In the Bayesian framework, we present a new bound which is derived from a constrained optimization problem. This bound is found to be tighter than the Bayesian Bhattacharyya bound, the Reuven-Messer bound, the Bobrovsky-Zakaï bound, and the Bayesian Cramér-Rao bound.

1. INTRODUCTION

The Cramér-Rao bound (CRB) is known to be the most used tool to study the performance of any estimator. Unfortunately, even if the CRB is a good benchmark in the asymptotic areas, i.e. for a large number of observations and/or a large Signal-to-Noise Ratio (SNR), it becomes too weak and too optimistic when the number of observation and/or the SNR decrease. To fill this lack, alternative bounds have been introduce in the literature. These bounds are tighter than the CRB and provide better baselines to study estimator performance. When the parameters are modelled as unknown deterministic quantities, the so-called deterministic bounds are: the CRB, the Bhattacharyya [1], the Chapman-Robbins [2], the Barankin [3] and the Abel bound [4]. The CRB and the Bhattacharyya bound account for the *small estimation error* (near the true value of the parameters). The Chapman-Robbins bound and the Barankin bound account for the *large estimation error* generally due to the appearance of outliers which creates a performance breakdown of the estimator. In [4], Abel combines the two kinds of bounds in order to obtain a bound which accounts for both local and large errors. The obtained deterministic Abel bound leads to a generalization of the Cramér-Rao, the Bhattacharyya, the Chapman-Robbins, and the Barankin bounds. On the other hand, when the parameters are modelled as random quantities, the so-called Bayesian bounds are divided in two families [5]: the Ziv-Zakaï family which derives from a binary hypothesis test [6] [7] and

the Weiss-Weinstein family which derives, as the deterministic bounds, from a covariance inequality principle [8]. This paper focus on the second family. The main bounds of the Weiss-Weinstein family are: the Bayesian Cramér-Rao [9], the Bayesian Bhattacharyya [9], the Bobrovsky-Zakaï [10], the Reuven-Messer (all parameters random) [11] and the Weiss-Weinstein bounds [12]. As the deterministic bounds, the Bayesian Cramér-Rao bound and the Bayesian Bhattacharyya bound are *small error* bounds contrary to the Bobrovsky-Zakaï bound and the Reuven-Messer bound which are *large error* bounds. The purpose of this paper is to apply the idea of Abel in the Bayesian context, i.e. to derive a bound which combines the Bayesian small and large error bounds. This will be done by the way of a constrained optimization problem introduced in [13] in which the Bayesian bounds of the Weiss-Weinstein family are seen to be the result of the MSE minimization subject to a set of constraints. Each kind of constraints defines a bound of the family. In the proposed paper, a bound is derived by mixing the constraints of the Reuven-Messer bound and the Bhattacharyya bound and, thus, represents a generalization of these bounds. A comparison with bounds of the Weiss-Weinstein family is given. The proposed bound is found to be a better tool to exhibit the well known SNR threshold phenomena.

2. BACKGROUND

2.1. Constrained optimization problem

The best Bayesian bound, given by the conditional mean estimator, can be reformulated as the solution of the following constrained optimization problem:

$$
\begin{cases}
\min_{v} \int\!\!\int_{\Omega} \int_{\Theta} v^2(\mathbf{x}, \theta) \, p(\mathbf{x}, \theta) \, d\theta d\mathbf{x} \\
\text{subject to } v(\mathbf{x}, \theta) = \hat{\theta}(\mathbf{x}) - \theta
\end{cases}
\tag{1}
$$

where $\hat{\theta}$ is an estimator, $p(\mathbf{x}, \theta)$ is the joint probability density function, $\mathbf{x} \in \Omega$ is an observation vector, and $\theta \in \Theta$ is a scalar

parameter that we want to estimate.

In [13], it is shown that the constraint in (1) is equivalent to a continuum of appropriate linear constraints and that the relaxation of this continuum leads to lower bounds on the MSE as the solution of the following optimization problem

$$
\begin{cases}
\min_{v} \iint\limits_{\Omega\ \Theta} v^2\left(\mathbf{x},\theta\right) p\left(\mathbf{x},\theta\right) d\theta d\mathbf{x} \\
\text{s. t.} \iint\limits_{\Omega\ \Theta} v\left(\mathbf{x},\theta\right) g_i\left(\mathbf{x},\theta\right)\sqrt{p\left(\mathbf{x},\theta\right)}d\theta d\mathbf{x} = c_i
\end{cases}
\tag{2}
$$

where $\{g_i\left(\mathbf{x},\theta\right)\}$ $i=0...K$ are a set of functions of $IR^N \to IR$ and c_i $i=0...K$ are $K+1$ real numbers.

The solution of (2) is given by [13]:

$$
MSE_{\min} = \mathbf{c}^{\mathrm{T}}\mathbf{G}^{-1}\mathbf{c},
\tag{3}
$$

where $\mathbf{c} = \left[c_0\ c_1...c_K\right]^{\mathrm{T}}$ and the element $G_{m,n}$ of the matrix $\mathbf{G}$ is given by

$$
G_{m,n} = \iint\limits_{\Theta\ \Omega} g_m\left(\mathbf{x},\theta\right), g_n\left(\mathbf{x},\theta\right) d\mathbf{x}d\theta.
\tag{4}
$$

Particular choices of the functions $g_i\left(\mathbf{x},\theta\right)$ and of the numbers c_i in the problem (2) lead to the bounds of the Weiss-Weinstein family. In the following, the matrix $\mathbf{G}$ is assumed to be non-singular.

2.2. Bayesian Bhattacharyya bound

Let us set $\mathbf{g} = \left[g_0\left(\mathbf{x},\theta\right), g_1\left(\mathbf{x},\theta\right),...g_K\left(\mathbf{x},\theta\right)\right]^{\mathrm{T}}$. By setting the following set of $K+1$ constraints

$$
\mathbf{g} = \frac{1}{\sqrt{p\left(\mathbf{x},\theta\right)}}\left[p\left(\mathbf{x},\theta\right)\ \frac{\partial p\left(\mathbf{x},\theta\right)}{\partial\theta}...\frac{\partial^K p\left(\mathbf{x},\theta\right)}{\partial\theta^K}\right]^{\mathrm{T}},
\tag{5}
$$

and

$$
\mathbf{c} = \left[0\ 1\ 0\cdots 0\right]^{\mathrm{T}} \stackrel{def}{=} \left[0\ \boldsymbol{\alpha}^{\mathrm{T}}\right]^{\mathrm{T}},
\tag{6}
$$

the solution given by eqn. (3) is

$$
MSE_{\min} = \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{B}^{-1}\boldsymbol{\alpha},
\tag{7}
$$

where the element $B_{i,j}$ of the matrix $\mathbf{B}$ is given by

$$
B_{i,j} = \iint\limits_{\Omega\ \Theta} \frac{1}{p\left(\mathbf{x},\theta\right)}\frac{\partial^i p\left(\mathbf{x},\theta\right)}{\partial\theta^i}\frac{\partial^j p\left(\mathbf{x},\theta\right)}{\partial\theta^j}d\theta d\mathbf{x}.
\tag{8}
$$

This is the Bayesian Bhattacharyya bound of order K [9]. In the particular case where $K=1$,

$$
MSE_{\min} = \left(\iint\limits_{\Omega\ \Theta} \frac{1}{p\left(\mathbf{x},\theta\right)}\left(\frac{\partial p\left(\mathbf{x},\theta\right)}{\partial\theta}\right)^2 d\theta d\mathbf{x}\right)^{-1},
\tag{9}
$$

which is the Bayesian Cramér-Rao bound [9].

2.3. Reuven-Messer bound

By setting the following set of $K+1$ constraints

$$
\mathbf{g} = \frac{1}{\sqrt{p\left(\mathbf{x},\theta\right)}}\left[p\left(\mathbf{x},\theta\right)\ p\left(\mathbf{x},\theta+h_1\right)...p\left(\mathbf{x},\theta+h_K\right)\right]^{\mathrm{T}},
\tag{10}
$$

and

$$
\mathbf{c} = \left[0\ h_1\ h_2\cdots h_K\right]^{\mathrm{T}} \stackrel{def}{=} \left[0\ \mathbf{h}^{\mathrm{T}}\right]^{\mathrm{T}},
\tag{11}
$$

with $h_i \in IR$ for $i=1,...,K$, the solution given by eqn. (3) is

$$
MSE_{\min} = \mathbf{h}^{\mathrm{T}}\left(\mathbf{D} - \mathbf{1}\mathbf{1}^{\mathrm{T}}\right)^{-1}\mathbf{h},
\tag{12}
$$

where the element $D_{i,j}$ of the matrix $\mathbf{D}$ is given by

$$
D_{i,j} = \iint\limits_{\Omega\ \Theta} \frac{p\left(\mathbf{x},\theta+h_i\right) p\left(\mathbf{x},\theta+h_j\right)}{p\left(\mathbf{x},\theta\right)}d\theta d\mathbf{x}.
\tag{13}
$$

This is the Reuven-Messer bound of order K where all parameters are random [11] which can be seen as the Bayesian version of the Barankin bound. In the particular case where $K=1$, one obtains

$$
MSE_{\min} = h^2\left(\iint\limits_{\Omega\ \Theta} \frac{p^2\left(\mathbf{x},\theta+h\right)}{p\left(\mathbf{x},\theta\right)}d\theta d\mathbf{x} - 1\right)^{-1},
\tag{14}
$$

which is the Bobrovsky-Zakaï bound [10].

3. THE BAYESIAN ABEL BOUND

Here, we combine the Bayesian Bhattacharyya constraints and the Reuven-Messer constraints. The goal is to provide a bound which takes into account the small and large error bounds.

By concatenating both vectors $\mathbf{g}$ and $\mathbf{c}$ from the Bayesian Bhattacharyya bound of order m and from the Reuven-Messer bound of order r we obtain the following set of $m+r+1$ constraints[1]

$$
\mathbf{g} = \frac{1}{\sqrt{p\left(\mathbf{x},\theta\right)}}\begin{bmatrix} p\left(\mathbf{x},\theta\right) \\ \frac{\partial p\left(\mathbf{x},\theta\right)}{\partial\theta} \\ \vdots \\ \frac{\partial^m p\left(\mathbf{x},\theta\right)}{\partial\theta^m} \\ ----- \\ p\left(\mathbf{x},\theta+h_1\right) \\ \vdots \\ p\left(\mathbf{x},\theta+h_r\right) \end{bmatrix} \text{ and } \mathbf{c} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -- \\ h_1 \\ \vdots \\ h_r \end{bmatrix}.
\tag{15}
$$

We must now calculate the quadratic form (3). Since

$$
\iint\limits_{\Omega\ \Theta} p\left(\mathbf{x},\theta+h_i\right) d\theta d\mathbf{x} = 1 \quad \forall h_i \in IR,
\tag{16}
$$

---
[1]The first constraint of the two bounds is the same.

and[2]

$$\iint_{\Omega\ \Theta} \frac{\partial^i p\left(\mathbf{x},\theta\right)}{\partial \theta^i} d\theta d\mathbf{x} = 0 \quad \forall i \geq 1, \qquad (17)$$

the matrix $\mathbf{G} = \iint_{\Omega\ \Theta} \mathbf{g}\mathbf{g}^{\mathrm{T}} d\theta d\mathbf{x}$ can now be written as the following partitioned matrix

$$\mathbf{G} = \begin{pmatrix} 1 & \mathbf{0}^{\mathrm{T}} & \mathbf{1}^{\mathrm{T}} \\ \mathbf{0} & \mathbf{B} & \boldsymbol{\Gamma}^{\mathrm{T}} \\ \mathbf{1} & \boldsymbol{\Gamma} & \mathbf{D} \end{pmatrix}, \qquad (18)$$

where $\mathbf{0}$ is a column vector of m zeros, $\mathbf{1}$ is a column vector of r ones, the elements $B_{i,j}$ and $D_{i,j}$ of the matrices $\mathbf{B}$ $(m \times m)$ and $\mathbf{D}$ $(r \times r)$ are given by relation (8) and (13), respectively, and the element $\Gamma_{i,j}$ of the matrix $\boldsymbol{\Gamma}$ $(r \times m)$ is given by

$$\Gamma_{i,j} = \iint_{\Omega\ \Theta} \frac{p\left(\mathbf{x},\theta+h_i\right)}{p\left(\mathbf{x},\theta\right)} \frac{\partial^j p\left(\mathbf{x},\theta\right)}{\partial \theta^j} d\theta d\mathbf{x}. \qquad (19)$$

Let us set $\boldsymbol{\beta} = \left[\mathbf{0}^{\mathrm{T}}\ \mathbf{1}^{\mathrm{T}}\right]^{\mathrm{T}}$, $\tilde{\mathbf{G}} = \begin{pmatrix} \mathbf{B} & \boldsymbol{\Gamma}^{\mathrm{T}} \\ \boldsymbol{\Gamma} & \mathbf{D} \end{pmatrix}$ and $\mathbf{c} = \left[0\ \boldsymbol{\alpha}^{\mathrm{T}}\ \mathbf{h}^{\mathrm{T}}\right]^{\mathrm{T}}$ where $\boldsymbol{\alpha}$ and $\mathbf{h}$ are defined by the relations (6) and (11), respectively. Since the first element of $\mathbf{c}$ is null, only the bottom right corner $\mathbf{R}$ (size $(m+r) \times (m+r)$) of $\mathbf{G}^{-1}$ is of interest in the relation (3). $\mathbf{R}$ is given straightforwardly (see e.g. [14] eqn. A.68) by

$$\mathbf{R} = \left(\tilde{\mathbf{G}} - \boldsymbol{\beta}\boldsymbol{\beta}^{\mathrm{T}}\right)^{-1} = \begin{pmatrix} \mathbf{B} & \boldsymbol{\Gamma}^{\mathrm{T}} \\ \boldsymbol{\Gamma} & \mathbf{D} - \mathbf{1}\mathbf{1}^{\mathrm{T}} \end{pmatrix}^{-1}, \qquad (20)$$

where we recognize the term $\mathbf{D} - \mathbf{1}\mathbf{1}^{\mathrm{T}} \overset{def}{=} \mathbf{J}$ of the Reuven-Messer bound. The Bayesian Abel bound denoted $BAB_{m,r}$ is then given by reducing equation (3) to

$$BAB_{m,r} = \left[\boldsymbol{\alpha}^{\mathrm{T}}\ \mathbf{h}^{\mathrm{T}}\right] \begin{pmatrix} \mathbf{B} & \boldsymbol{\Gamma}^{\mathrm{T}} \\ \boldsymbol{\Gamma} & \mathbf{J} \end{pmatrix}^{-1} \begin{bmatrix} \boldsymbol{\alpha} \\ \mathbf{h} \end{bmatrix}. \qquad (21)$$

After some algebraic effort, one obtains the final form

$$\boxed{BAB_{m,r} = \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{B}^{-1}\boldsymbol{\alpha} + \mathbf{u}^{\mathrm{T}}\tilde{\mathbf{J}}^{-1}\mathbf{u},} \qquad (22)$$

with

$$\begin{cases} \mathbf{u} = \boldsymbol{\Gamma}\mathbf{B}^{-1}\boldsymbol{\alpha} - \mathbf{h} \\ \tilde{\mathbf{J}} = \mathbf{J} - \boldsymbol{\Gamma}\mathbf{B}^{-1}\boldsymbol{\Gamma}^{\mathrm{T}} \end{cases} \qquad (23)$$

Let's note that the first term of right hand side of (22) is equal to $BAB_{m,0}$ which is the Bayesian Bhattacharyya bound of order m and that $BAB_{0,r}$ is the Reuven-Messer bound of order r. In [13], it is shown that the problem (2) leads to the conditional mean estimator MSE, which is the Bayesian optimal estimator, if the number of constraints tends to infinity. Here, from the increase of constraints, it follows that

---

[2] We assume that $\lim_{\theta \to \pm\infty} \frac{\partial^i p(\mathbf{x},\theta)}{\partial \theta^i} = 0\ \forall i \geq 1$.

---

the Bayesian Abel bound is (for r and m fixed) a better approximation of the conditional mean estimator MSE than the Bayesian Bhattacharyya bound of order m and the Reuven-Messer bound of order r.

The Bayesian Abel bound as the Reuven-Messer bound depends on r free parameters $h_1, ..., h_r$. Then, a maximization over these parameters is desired to obtain the highest bound. Therefore, the best Bayesian Abel bound is given by

$$BAB_{m,r} = \sup_{h_1,...,h_r} \left(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{B}^{-1}\boldsymbol{\alpha} + \mathbf{u}^{\mathrm{T}}\tilde{\mathbf{J}}^{-1}\mathbf{u}\right). \qquad (24)$$

This multidimensional optimization leads to a huge computational cost. A possible alternative is given by noting that the Bayesian Cramér-Rao bound is a particular case of the Bayesian Bhattacharyya bound with a single derivative and that the Bobrovsky-Zakaï bound is a particular case of the Reuven-Messer bound with a single test point. Therefore, finding a tractable form of the Bayesian Abel bound in the case where $m = 1$ and $r = 1$ could be interesting since the obtained bound will be tighter than the Bayesian Cramér-Rao bound and than the Bobrovsky-Zakaï bound with a low computational cost. In this case, equation (24) becomes straightforwardly

$$BAB_{1,1} = \sup_{h} \frac{BCRB^{-1} + BZB^{-1} - 2\phi}{BCRB^{-1}BZB^{-1} - \phi^2}. \qquad (25)$$

Where *BCRB* is the Bayesian Cramér-Rao bound, *BZB* is the Bobrovsky-Zakaï bound and

$$\phi = \frac{1}{h} \iint_{\Omega\ \Theta} \frac{\partial \ln p\left(\mathbf{x},\theta\right)}{\partial \theta} p\left(\mathbf{x},\theta+h\right) d\theta d\mathbf{x}. \qquad (26)$$

This low cost bound has been introduced in [13]. Equation (25) shows that, if the Bayesian Cramér-Rao bound and the Bobrovsky-Zakaï bound are available, the evaluation of the $BAB_{1,1}$ requires only the computation of ϕ.

4. SIMULATION RESULTS

The comparison of the Bayesian Abel bound with classical bounds of the same computational cost will be investigated in a classical spectral analysis problem. The model is the following

$$\mathbf{x} = \rho\mathbf{s} + \mathbf{n}, \qquad (27)$$

where $\mathbf{x}$ is the observation vector $(N \times 1)$, ρ^2 is the SNR, $\mathbf{s} = \left[1, e^{j\theta}, ..., e^{j(N-1)\theta}\right]^{\mathrm{T}}$ is the signal, $\mathbf{n}$ is the noise vector $(N \times 1)$ which is assumed circular, complex Gaussian with zero mean and covariance matrix $\mathbf{I}_N$, and θ is the parameter of interest. The a priori pdf of θ is Gaussian with zero mean and variance σ_θ^2. Figure 1 superimposes the Bayesian Maximum Likelihood Estimator (BMLE) MSE [15], the Maximum A Posteriori Estimator (MAPE) MSE, the Bayesian

Cramér-Rao Bound (BCRB), the Bobrovsky-Zakaï Bound (BZB) and the $BAB_{1,1}$ for $N = 10$ observations and $\sigma_\theta^2 = 0.5$ rad^2. For the BMLE, the search domain corresponds to the prior information $[-3\sigma_\theta, 3\sigma_\theta]$.
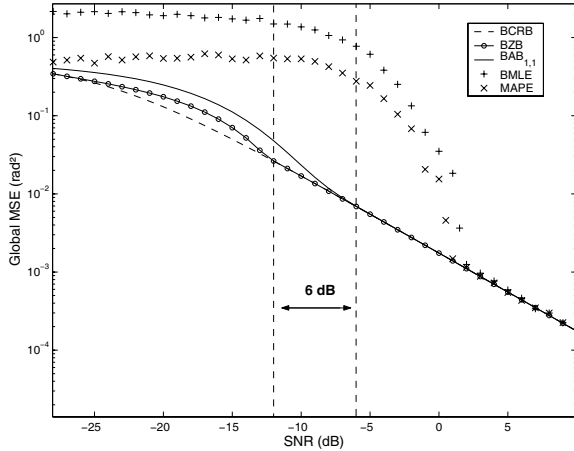

Fig. 1. Comparison of proposed bound $BAB_{1,1}$ with BMLE, MAPE, BCRB, BZB. $N = 10$, $\sigma_\theta^2 = 0.5$.

The SNR threshold prediction is improved by about 6dB (for $N = 10$) in comparison with the BZB. Extensive simulations (not reported here due to the lack of space) have shown that for critical scenarios with small number of observations the BZB tends to the BCRB without exhibiting the threshold phenomena. Such drawbacks do not appear with the proposed bound.

5. CONCLUSION

In this paper, a new bound on the mean square error has been investigated. This bound can be seen as the Bayesian version of the deterministic Abel bound. The proposed bound combines the small and large error bounds. It is a non-linear mixture of the Bayesian Bhattacharyya bound and of the Reuven-Messer bound and is found to be higher than these two bounds without increasing the computational cost. A trackable form of the bound is given for $m = r = 1$. This form is higher than the Bayesian Cramér-Rao bound and than the Bobrovsky-Zakaï bound for about the same complexity. Simulation results show the better ability to the proposed bound to predict the SNR threshold phenomena in comparison with the others.

6. REFERENCES

[1] A. Bhattacharrya, "On some analogues of the amount of information and their use in statistical estimation," *Sankhya Indian J. of Stat.*, vol. 8, pp. 1–14, 201–218, 315–328, 1946.

[2] D. G. Chapman and H. Robbins, "Minimum variance estimation without regularity assumptions," *Ann. Math. Stat.*, vol. 22, pp. 581–586, 1951.

[3] E. W. Barankin, "Locally best unbiased estimates," *Ann. Math. Stat.*, vol. 20, pp. 477–501, 1949.

[4] J. S. Abel, "A bound on mean-square-estimate error," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1675–1680, Sept. 1993.

[5] K. Bell, *Performance bounds in parameter estimation with application to bearing estimation*, Ph.D. thesis, George Mason University, Fairfax, VA, 1995.

[6] J. Ziv and M. Zakai, "Some lower bounds on signal parameter estimation," *IEEE Trans. Inform. Theory*, vol. 15, pp. 386–391, May 1969.

[7] K. Bell, Y. Steinberg, Y. Ephraim, and H. L. Van Trees, "Extended Ziv-Zakai lower bound for vector parameter estimation," *IEEE Trans. Signal Processing*, vol. 43, pp. 624–638, Mar. 1997.

[8] E. Weinstein and A. J. Weiss, "A general class of lower bounds in parameter estimation," *IEEE Trans. Inform. Theory*, vol. 34, pp. 338–342, Mar. 1988.

[9] H. L. Van Trees, *Detection, Estimation and Modulation Theory*, vol. 1, Wiley, New York, 1968.

[10] B. Z. Bobrovsky and M. Zakai, "A lower bound on the estimation error for certain diffusion processes," *IEEE Trans. Inform. Theory*, vol. 22, pp. 45–52, Jan. 1976.

[11] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of a hybrid parameter vector," *IEEE Trans. Inform. Theory*, vol. 43, no. 3, pp. 1084–1093, May 1997.

[12] A. J. Weiss and E. Weinstein, "A lower bound on the mean square error in random parameter estimation," *IEEE Trans. Inform. Theory*, vol. 31, pp. 680–682, Sept. 1985.

[13] A. Renaux, P. Forster, and P. Larzabal, "A new derivation of the Bayesian bounds for parameter estimation," in *Proc. of IEEE Statistical Signal Processing Workshop - SSP05*, Bordeaux, FR, July 2005.

[14] H. L. Van Trees, *Detection, Estimation and Modulation theory: Optimum Array Processing*, vol. 4, Wiley, New York, 2002.

[15] W. Xu, *Performance bounds on matched-field methods for source localization and estimation of ocean environmental parameters*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, 2001.