WIDEBAND ARRAY SIGNAL PROCESSING USING MCMC METHODS

William Ng, James P. Reilly, Thia Kirubarajan, and Jean-René Larocque*

Department of Electrical and Computer Engineering,
McMaster University, 1280 Main St. W.,
Hamilton, Ontario, Canada L8S 4K1

ABSTRACT

This paper proposes a novel wideband structure for array signal processing. A new wideband model is formed where the source amplitudes are linear functions of the observations, but nonlinear in the direction of arrival (DOA) parameters. The method lends itself well to a Bayesian approach for jointly estimating the model order and the DOAs through a reversible jump Markov chain Monte Carlo (MCMC) procedure. The source amplitudes are estimated through a *maximum a posteriori* (MAP) procedure. The DOA estimation performance of the proposed method is compared with the theoretical Cramér-Rao lower bound (CRLB) for this problem. Simulation results demonstrate the effectiveness and robustness of the method.

1. INTRODUCTION

Array signal processing, which has found in use in radar, sonar, communications, geophysical exploration, astrophysical exploration, biomedical signal processing, and acoustics [1], has to do with 1) detection the number of incident sources, 2) estimation of parameters, like direction-of-arrival (DOA) or time-of-arrival (TOA) of the sources impinging onto the array, and 3) recovery of the incident source waveforms. Methods for each of the above objectives can be classified as either narrowband or wideband. For the narrowband scenario, there exist many algorithms to solve this detection and estimation problem [1] [2] [3] [4] [5] [6]. [3] [4] can perform the determination of model order and the estimation of desired signal parameters jointly rather than independently. However, for the wideband scenario, no existing methods can attain the objective of joint detection and estimation simultaneously due to the cumbersome nature of the problem.

In this paper, we propose a novel model structure which applies equally well to both narrowband and wideband cases, that detects model order, estimates DOA, and recovers the source waveforms in a computationally efficient manner. The approach proposed in this paper is an extension of the method of [3] to seamlessly perform joint detection of the number of sources and estimation of TOAs (DOAs), and recovery of the sources, for both narrowband and wideband models.

This paper is organized as follows. Section 2 presents a general model to represent wideband signals and describes the derivation of the necessary probability distributions. A description of the reversible jump MCMC algorithm used for model order detection is given in Section 3 , followed by the simulation results and discussion in Section 4 . Conclusions are given in Section 5.

J. Reilly, corresponding author: ph: 905 525 9140 x22895, fax: 905 521 2922, email: reillyj@mcmaster.ca

2. THE DATA MODEL

The signal model we consider consists of a set of data vector $\boldsymbol{y}(n) \in \mathcal{R}^M$, which represents the data received by a linear array of M sensors at the nth snapshot. The data vector is composed of incident wideband plane wave signals, each of which impinges on the array of sensors at an angle $\theta_k, k = 0, 1, ..., K-1$, and is bandlimited to $|f| \in \left[f_k^l, f_k^u\right]$, where $f_k^u = f_k^l + \Delta f_k$, f_k^l and f_k^u are the lower and upper frequencies, and Δf_k is the bandwidth of the kth source.

It can be shown [7] that the inter-sensor delay of source k, τ_k, is bounded as follows $|\tau_k| \leq \frac{1}{2f_k^u}$, where $\tau_k \triangleq \frac{\Delta}{C}\sin\theta_k$, Δ is the interspacing of the sensors, and C is the speed of propagation. Denoting the maximum allowable inter-senor delay by T_{max}, we have

$$T_{max} = \min_{k=0,\ldots,K-1}\left\{\frac{1}{2f_k^u}\right\}. \quad (1)$$

The received vector at the nth snapshot can be written as [7]

$$\boldsymbol{y}(n) = \sum_{k=0}^{K-1} \boldsymbol{s}_k(t-\tau_k) + \sigma_w \boldsymbol{w}(n), \;\; n = 1, ..., N \quad (2)$$

$$\approx \sum_{k=0}^{K-1} \tilde{\boldsymbol{H}}(\tau_k)\boldsymbol{s}_k(n) + \sigma_w \boldsymbol{w}(n), \quad (3)$$

where[1] N is the number of snapshots, $\boldsymbol{w}(n)$ is an iid Gaussian variable with zero mean and unit variance, σ_w^2 is the noise variance in the observation, $\tilde{\boldsymbol{H}}(\tau_k) \in \mathcal{R}^{M\times L}$ is an *interpolation matrix* [2] for τ_k with L taps, defined as [7]

$$\tilde{\boldsymbol{H}}(\tau_k) = \begin{cases} \boldsymbol{H}(\tau_k), & \text{if} \quad \theta_k \leq \pi/2 \\ \boldsymbol{E}_M \boldsymbol{H}(\tau_k), & \text{if} \quad \theta_k > \pi/2 \end{cases}, \tag{4}$$

where

$$\boldsymbol{E}_M = \begin{bmatrix} 0 & 0 & \dots & 1 \\ 0 & \dots & 1 & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 1 & 0 & 0 & 0 \end{bmatrix}, \tag{5}$$

and $\boldsymbol{s}_k(t)$, the kth signal, and $\boldsymbol{s}_k(n)$, the corresponding discrete sequences, are respectively defined as

$$\boldsymbol{s}_k(t-\tau_k) = \left[s_k(t), s_k(t-\tau_k), \dots, s_k(t-(M-1)\tau_k)\right]^T, \tag{6}$$

$$\boldsymbol{s}_k(n) = \left[s_k(n), s_k(n-1), \dots, s_k(n-(L-1)\tau_k)\right]^T. \tag{7}$$

Defining $\tilde{\boldsymbol{H}}_l(\boldsymbol{\tau}) \in \mathcal{R}^{M\times K}$ as follows

$$\tilde{\boldsymbol{H}}_l(\boldsymbol{\tau}) = \left[\tilde{\boldsymbol{H}}_l(\tau_0), \tilde{\boldsymbol{H}}_l(\tau_1), \dots, \tilde{\boldsymbol{H}}_l(\tau_{K-1})\right], \tag{8}$$

we can express (3) as follows

$$\boldsymbol{y}(n) = \sum_{l=0}^{L-1} \tilde{\boldsymbol{H}}_l(\boldsymbol{\tau})\boldsymbol{a}(n-l) + \sigma_w \boldsymbol{w}(n), \tag{9}$$

where $\boldsymbol{a}(n) \triangleq [s_0(n), s_1(n), \dots, s_{K-1}(n)]^T$. The signal vector $\boldsymbol{a}(n)$ for $l = 1, ..., L-1$ in can be considered known since it consists of past values of the sources, $s_k(n)$ for $k = 0, ..., K-1$. We define a vector $\boldsymbol{z}(n)$ as follow

$$\boldsymbol{z}(n) \triangleq \boldsymbol{y}(n) - \sum_{l=1}^{L-1} \tilde{\boldsymbol{H}}_l(\boldsymbol{\tau})\boldsymbol{a}(n-l), \tag{10}$$

which can be interpreted as the error between the snapshot $\boldsymbol{y}(n)$ and its approximation, based on the past, known values of the signals from $\boldsymbol{a}(n-1)$ to $\boldsymbol{a}(n-L+1)$ and the associated column vectors in the interpolation matrix. Accordingly, we may rewrite (9) as follows

$$\boldsymbol{z}(n) = \tilde{\boldsymbol{H}}_0(\boldsymbol{\tau})\boldsymbol{a}(n) + \sigma_w \boldsymbol{w}(n), \tag{11}$$

which represents the desired form of the model. This model can accomodate either narrowband or wideband sources, without change of structure or parameters.

[1]Note that for notational convenience, from this point onwards we replace the approximation with an equality.

[2]For example, in the case of the uniform linear array, the interpolation matrix can be computed using a windowed *sinc*$(\cdot)$ function.

We assume the noise vectors $\boldsymbol{w}(n)$ are *iid*, and that all the parameters describing the received signal are stationary throughout the entire observation interval. Consider a linear array of M sensors. Using (11), we may define a set of N snapshots $\boldsymbol{Z} = [\boldsymbol{z}(1), \dots, \boldsymbol{z}(N)]$, and hence the desired posterior distribution function of the parameters as follows

$$\begin{aligned} \pi(\boldsymbol{a}, \boldsymbol{\tau}, \sigma_w^2, k|\boldsymbol{Z}) \propto\, & p(\boldsymbol{Z}|\boldsymbol{a}, \boldsymbol{\tau}, \sigma_w^2, k) \\ & \times p(\boldsymbol{a}|k, \boldsymbol{\tau}, \delta^2\sigma_w^2) \\ & \times p(\boldsymbol{\tau}|k)p(\sigma_w^2)p(k), \end{aligned} \tag{12}$$

where $\boldsymbol{a} = [\boldsymbol{a}(1), \dots, \boldsymbol{a}(N)]$, and k represents an estimate of the true number of sources, K. Assuming that the observations are *iid*, the total likelihood function is

$$\ell(\boldsymbol{Z}|\boldsymbol{a}, \boldsymbol{\tau}, \sigma_w^2, k) = \prod_{n=1}^{N} \mathcal{N}\left(\tilde{\boldsymbol{H}}_0(\boldsymbol{\tau})\boldsymbol{a}(n), \sigma_w^2 \boldsymbol{I}\right). \tag{13}$$

To complete the model, prior distributions of the parameters are required. The amplitudes are chosen *iid* with covariance matrix corresponding to the *maximum entropy* prior as follows [3]

$$p(\boldsymbol{a}|k, \boldsymbol{\tau}, \delta^2\sigma_w^2) = \prod_{n=1}^{N} \mathcal{N}\left(0, \delta^2\sigma^2\left[\tilde{\boldsymbol{H}}_0^T(\boldsymbol{\tau})\tilde{\boldsymbol{H}}_0(\boldsymbol{\tau})\right]^{-1}\right), \tag{14}$$

where δ^2 is a hyperparameter equal to the signal-to-noise ratio. The prior distribution of $\boldsymbol{\tau}$ is chosen to be uniform:

$$p(\boldsymbol{\tau}|k) = \mathcal{U}\left[-T_{max}, T_{max}\right]^k. \tag{15}$$

The prior for the parameter σ_w^2 is chosen as the inverse-Gamma distribution, which is the conjugate prior corresponding to a Gaussian likelihood function. It is defined as

$$p(\sigma_w^2) = \mathcal{IG}\left(\frac{\nu_0}{2}, \frac{\gamma_0}{2}\right). \tag{16}$$

Finally, the prior distribution on k is chosen to be Poisson with expected number of sources Λ as follows

$$p(k) = \frac{\Lambda^k}{k!}\exp(-\Lambda). \tag{17}$$

This choice of prior is not strictly noninformative, but it contributes to a more efficient MCMC sampling routine.

We can simplify the estimation of the parameters in the posterior distribution function (12) by considering $\boldsymbol{a}$ and σ_w to be nuisance parameters, and analytically integrating them out. The only quantities of interest at this stage are $\boldsymbol{\tau}$ and k. We recover the $\boldsymbol{a}$ later. By following the similar procedures in [3] [8], it can be shown that (12) can be simplified as

$$\pi(\boldsymbol{\tau}, k|\boldsymbol{Z}) \propto \int_0^\infty \int_{-\infty}^\infty \pi(\boldsymbol{a}, \boldsymbol{\tau}, \sigma_w^2, k|\boldsymbol{Z}) d\boldsymbol{a} d\sigma_w^2. \tag{18}$$

To restore the amplitudes $\boldsymbol{a}(n)$, a *maximum aposteriori* estimate of $\boldsymbol{a}(n)$ can be computed from (12), knowing other parameters is readily available.

3. REVERSIBLE JUMP MCMC

MCMC methods [9] are numerical techniques for performing Bayesian estimation, inference, or detection, corresponding to an arbitrary distribution of interest. The basic idea behind MCMC methods is to draw samples from a posterior distribution of interest, in effect forming a histogram, and then form sample averages to approximate expectations, thus facilitating parameter estimation.MCMC methods operate by setting up a Markov chain, whose invariant distribution is the posterior distribution $\pi(\boldsymbol{x})$ of interest. Thus, each state of the chain corresponds to a potential sample from $\pi(\boldsymbol{x})$.

Metropolis-Hastings (M-H) algorithm [10] [9] is used to draw samples from an arbitrary distribution. At each iteration step i of the Markov chain, the next state $\boldsymbol{\tau}^{(i+1)}$ is chosen by first sampling a *candidate* state $\boldsymbol{\tau}^{\star}$ from a *proposal* distribution $q(\cdot|\boldsymbol{\tau})$. The proposal distribution may depend on the current state of the Markov chain $\boldsymbol{\tau}$. The candidate state $\boldsymbol{\tau}^{\star}$ is then *accepted* with probability $\alpha(\boldsymbol{\tau}, \boldsymbol{\tau}^{\star}) = \min\{1, r(\boldsymbol{\tau}, \boldsymbol{\tau}^{\star})\}$, where $r(\boldsymbol{\tau}, \boldsymbol{\tau}^{\star})$ is the *acceptance ratio*, defined as

$$r(\boldsymbol{\tau}, \boldsymbol{\tau}^{\star}) = \frac{\pi(\boldsymbol{\tau}^{\star})q(\boldsymbol{\tau}|\boldsymbol{\tau}^{\star})}{\pi(\boldsymbol{\tau})q(\boldsymbol{\tau}^{\star}|\boldsymbol{\tau})}. \quad (19)$$

The proposal function $q(\cdot)$ is chosen to be easy to sample from, and must be non-null and positive over the support of the distribution $\pi(\cdot)$. If the candidate is accepted, the next state of the chain will be $\boldsymbol{\tau}^{(i+1)} = \boldsymbol{\tau}^{\star}$. If the candidate is rejected, the chain remains in state $\boldsymbol{\tau}^{(i+1)} = \boldsymbol{\tau}^{(i)}$.

The reversible jump MCMC algorithm [11] for joint detection of the model order and estimation of the other parameters from the posterior distribution is similar to the M-H algorithm, but it allows the sampling process to jump between subspaces of different dimensions, which facilitates the detection of model order. Denote the whole parameter space by $\cup_{k=0}^{k_{max}} k \times \Phi_k$, where Φ_k is the space of the parameters of the model of order k. At each iteration, candidate samples are chosen from a set of proposal distributions, which correspond to three moves [7]: update, birth, and death.

3.1. Update Move

Here, we assume that the current state of the algorithm is in $\{\boldsymbol{\Phi}_k, k\}$. In this move, the algorithm samples only on the space of $\boldsymbol{\Phi}_k$ for k fixed. The acceptance function for an update move is defined using (19) as

$$r_{update} = \frac{\pi(\boldsymbol{\tau}_k^{\star}|\boldsymbol{Z}^{\star})q(\boldsymbol{\tau}_k|\boldsymbol{\tau}_k^{\star})}{\pi(\boldsymbol{\tau}_k|\boldsymbol{Z})q(\boldsymbol{\tau}_k^{\star}|\boldsymbol{\tau}_k)}, \quad (20)$$

The candidate $\boldsymbol{\tau}_k^{\star}$ is then accepted as the current state $\boldsymbol{\tau}_k^{(i+1)} = \boldsymbol{\tau}_k^{\star}$, with probability $\alpha_{update}(\boldsymbol{\tau}_k, \boldsymbol{\tau}_k^{\star}) = \min\{1, r_{update}\}$.

3.2. Birth Move

In the birth move, we assume the state of the algorithm is in $\{\boldsymbol{\Phi}_k, k\}$ at the present ith iteration, and we wish to determine whether the state is in $\{\boldsymbol{\Phi}_{k+1}, k+1\}$ at the next iteration. The acceptance function of the birth move is therefore defined as

$$r_{birth} = \frac{\pi(\boldsymbol{\tau}_{k+1}^{\star}, k+1|\boldsymbol{Z})q(\boldsymbol{\tau}_k, k|\boldsymbol{\tau}_{k+1}^{\star}, k+1)}{\pi(\boldsymbol{\tau}_k, k|\boldsymbol{Z})q(\boldsymbol{\tau}_{k+1}^{\star}, k+1|\boldsymbol{\tau}_k, k)}. \quad (21)$$

The probability of accepting a birth move is therefore defined as $\alpha_{birth} = \min\{1, r_{birth}\}$.

3.3. Death Move

In order to maintain the invariant distribution of the reversible jump MCMC algorithm with respect to model order, the Markov chain must be *reversible* with respect to moves across subspaces of different model orders. That is, the probability of moving from model order k to $k+1$ must be equal to that of moving from $k+1$ to k. Therefore we propose a death move in which a source in the current state $(\boldsymbol{\tau}_{k+1}, k+1)$ is randomly selected to be removed such that the next state becomes $(\boldsymbol{\tau}_k, k)$ at the next iteration. A sufficient condition for reversibility with respect to model order [11] is that the acceptance ratio for the death move be defined as

$$r_{death} = \frac{1}{r_{birth}}, \quad (22)$$

and the new candidate of dimension k is accepted with probability $\alpha_{death} = \min\{1, r_{death}\}$.

4. SIMULATION RESULTS

The proposed algorithm is now applied to a wideband scenerio to demonstrate the capabilities of joint detection and estimation of $(\boldsymbol{\tau}, k)$ and the source amplitudes $\boldsymbol{s}_k(n)$. In this experiment, the model order k and the delay parameters $\boldsymbol{\tau}$ are kept constant throughout the entire observation period, and the hyper-parameters γ_0 and ν_0 are set to zero.

Parameter	Value
L	8
σ_w^2	0.0169
F_s (Hz)	1,000
$\boldsymbol{\theta}$ (deg)	$[-3.44, 3.44]$
$\boldsymbol{\tau}$ (sec)	$[-7.5, 7.5] \times 10^{-5}$

Table 4.1. Parameters for the experiments.

In this experiment, we generate $K = 2$ Gaussian processes for the sources that are zero mean and variance, δ^2/σ_w^2, and bandlimited to $f \in [100, 400]\,\mathrm{Hz}$. The incident angles $\boldsymbol{\theta}$ are separated by an angle less than a half standard beamwidth given as [1]. According to the other parameters in Table 4.1, an array of $M = 8$ sensors is used to generate $N = 50$ using (9) with $SNR = 14$dB. The hyper-parameter $\delta^2 = 25.12$ is assumed known and constant.

The proposed algorithm randomly initializes all unknown parameters, and randomly assigns the initial model order k uniformly in $[1, k_{max}]$, where $k_{max} = M - 1 = 7$, is the maximum allowable model order. As shown in Fig. 1, the algorithm takes about 25 iterations to converge to the correct order and about 2,000 iterations for a burn-in before the chain centres on the true delay values.

Fig. 2 depicts a comparison between the estimate of the delays of the sources and the true values after the burn-in stage, versus iteration number. It is clear that the chain centers on the true delay values. Table 4.2 summarizes a comparison between the true and estimated values of the incident angles and the corresponding delay parameters. Given the MAP estimate of the delays $\boldsymbol{\tau}$, the algorithm can now restore the source signals, as shown in Fig. 3. It is clear that the signal amplitudes are well separated and restored by MCMC. The mean-squared errors of the restored signals relative to the true signal amplitudes are -16.19dB and -15.97dB, respectively.

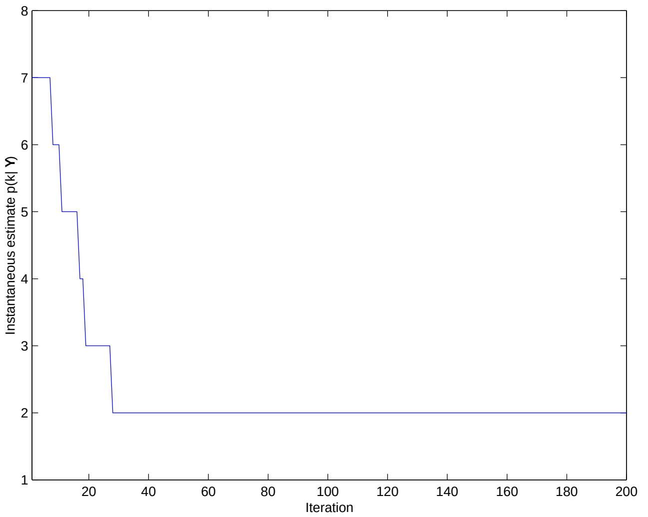

Fig. 1. Instantaneous estimate of $p(k|\boldsymbol{Y})$, for the first 200 iterations of the chain.

We now present comparisions with the theoretical CRLB derived in [7] for this problem. The variances of the estimated $\boldsymbol{\tau}$ are plotted in Fig. 4 along with the respective theoretical CRLBs. This evaluation is obtained by applying the algorithm to 100 independent trials on all the parameters but $\boldsymbol{\tau}$ listed in Table 4.1 over a range of SNR, from -5dB to 18dB. The delay parameters used in this evaluation are $[-7.5, 7.5] \times 10^{-5}$. As shown in Fig. 4, while the algorithm starts to break down for SNR levels lower than -2dB, the variances approach the CRLB closely. The reasons why the variances do not come closer to the theoretical CRLB are [7]: 1) interpolation errors when a non-ideal interpolation function is used and 2) the suboptimal procedure for estimating the source amplitudes. This procedure has an impact on the estimation accuracy of the DOA parameters. Further simulation results show [7] that the probability of an error in detection of the model order tends to diminish toward zero with increasing number of snapshots, N, with moderate SNR values.

Parameter	True	Estimated	Relative Difference (%)
τ_0	$-7.5e^{-5}$	$-7.95e^{-5}$	6.00
τ_1	$7.5e^{-5}$	$7.25e^{-5}$	3.36
θ_0	-3.44	-3.65	6.00
θ_1	3.44	3.32	3.37

Table 4.2. Comparison between the true and estimated parameters.

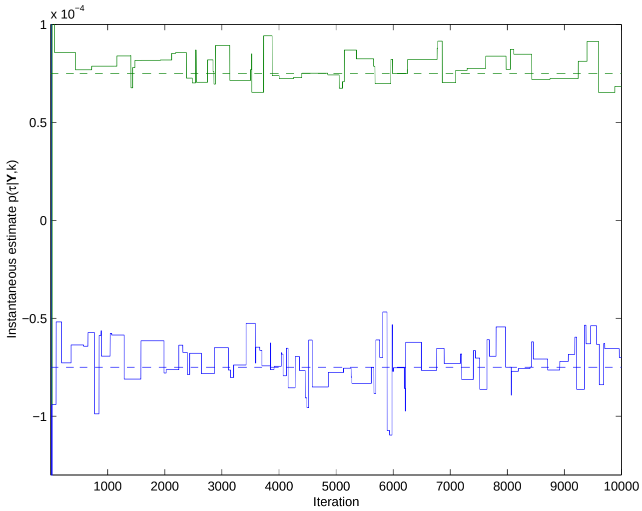

Fig. 2. Instantaneous estimate of the delays $\boldsymbol{\tau}$ for 2 sources: the solid lines are the estimates and the dashed lines are the ture values

5. CONCLUSION

A novel structure for wideband array signal processing is proposed. It has been demonstrated the method applies equally well to the narrowband case. A Bayesian approach is used, where a posterior density function which has the nuisance parameters integrated out is formulated. The desired model order and DOA estimation parameters are determined through a reversible jump MCMC procedure. The source amplitudes

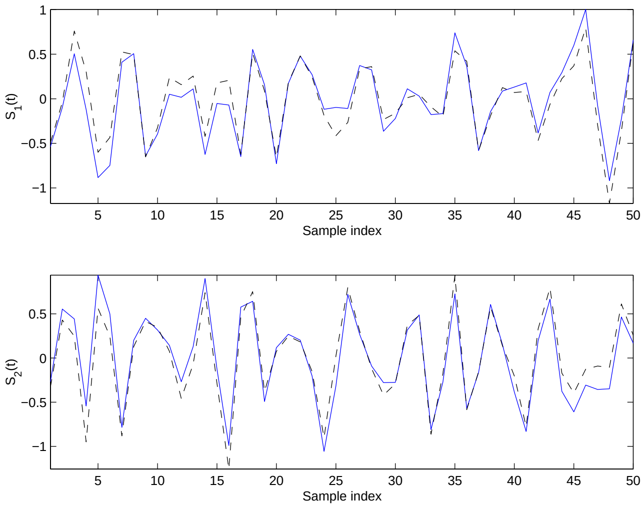

Fig. 3. A comparison between the true and the restored amplitudes using MCMC in one realization: solid lines correspond the restored amplitudes using MCMC and dashed lines correspond the true amplitudes.

are given using a MAP estimate. Simulation results support the effectiveness of the method, and demonstrate reliable detection of the number of sources and estimation of their times of arrival in white noise environment with a single linear array. As a result, the source signals are reliably restored.

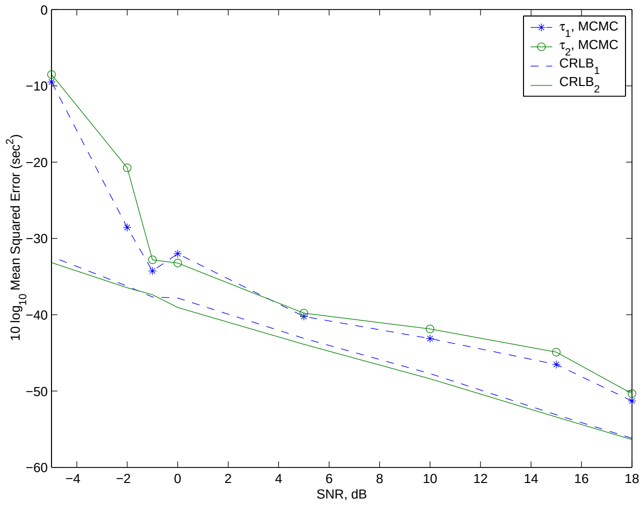

Fig. 4. Mean squared error of $\boldsymbol{\tau}$ versus the CRLB.

6. REFERENCES

[1] D.H. Johnson and D.E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, PTR Prentice Hall, New Jersey, 1993.

[2] R.O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, pp. 284–280, Mar. 1986.

[3] C. Andrieu and A. Doucet, "Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC," Oct. 1999, vol. 47, pp. 2667–2676.

[4] Q. Wu and D.R. Fuhrmann, "A parametric method for determining the number of signals in narrow-band direction finding," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 39, pp. 1848–1857, Aug. 1991.

[5] Simon Haykin and Jim Reilly, "Maximum likelihood reciever for low-angle tracking radar; Part I: The symmetric case," *IEE Proceedings*, vol. 129 Pt.F, no. 4, pp. 261–272, Aug. 1982.

[6] Simon Haykin and Jim Reilly, "Maximum likelihood reciever for low-angle tracking radar; Part II: The non-symmetric case," *IEE Proceedings*, vol. 129 Pt.F, no. 5, pp. 2667–2676, Oct. 1982.

[7] William Ng, James. P. Reilly, Thia Kirubarajan, and Jean-René Larocque, "Wideband array signal processing using mcmc methods," *IEEE Transactions on Signal Processing*, 2002, Submitted for publication, also available at http://www.ece.mcmaster.ca/∼reilly.

[8] C. Andrieu, A. Doucet, W.J. Fitzgerald, and S.J. Godsill, "An introduction to the theory and applications of simulation based computational methods in Bayesian signal processing," in *Tutorial from Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Seattle, WA, 1998.

[9] Joseph O. Ruanaidh and William Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing.*, Springer-Verlag, New York, 1996.

[10] W.K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," in *Biometrika*, 1970, vol. 57, pp. 97–109.

[11] P.J. Green, "Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination," in *Biometrika*, 1995, vol. 82, pp. 711–732.