WIDEBAND ARRAY SIGNAL PROCESSING USING MCMC METHODS

William Ng, James P. Reilly and Thia Kirubarajan

Department of Electrical and Computer Engineering,
McMaster University, 1280 Main St. W.,
Hamilton, Ontario, Canada L8S 4K1
ng@reverb.crl.mcmaster.ca, (reilly)(kiruba)@mail.ece.mcmaster.ca

ABSTRACT

This paper proposes a novel wideband structure for array signal processing. The method lends itself well to a Bayesian approach for jointly estimating the model order (number of sources) 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. Advantages of the proposed method include joint detection of model order and estimation of the DOA parameters, and the fact that meaningful results can be obtained using fewer observations than previous methods. 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]. Methods such as [3] [4] [5] can perform the determination of model order and the estimation of desired signal parameters jointly. However, for the wideband scenario, no existing methods can attain the objective of joint detection and estimation simultaneously due to the difficult nature of the problem.

As an extension of the method of [3], in this paper, we propose a novel *wideband* model structure which applies equally well to both the narrowband and wideband cases, that detects model order, estimates DOA, and recovers the source waveforms.

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. Simulation results are shown in Section 3 and the Conclusions are in Section 4.

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 K 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 is readily verified that the inter-sensor delay τ_k of source k, is bounded by $|\tau_k| \leq \frac{1}{2f_k^u}$, 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 then be written as [6]

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

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

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

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

interpolation matrix [2] and is defined as [6]

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

where $\boldsymbol{E}_M \in \mathcal{R}^{M\times M}$ is an *exchange matrix* (i.e., all zeros except for ones along the anti-diagonal) and $\boldsymbol{s}_k(t)$, the kth signal, and $\boldsymbol{s}_k(n)$, the corresponding discrete–time version, are defined respectively as

$$\boldsymbol{s}_k(t-\tau_k) = [s_k(t), s_k(t-\tau_k), \ldots, s_k(t-(M-1)\tau_k)]^T, \tag{5}$$

$$\boldsymbol{s}_k(n) = [s_k(n), s_k(n-1), \ldots, s_k(n-L+1)]^T. \tag{6}$$

The matrix $\boldsymbol{H}(\tau_k)$ interpolates the kth discrete–time sequence $\boldsymbol{s}_k(n)$ to give the desired sequence $\boldsymbol{s}_k(t-m\tau), m = 0, \ldots, M-1$. We now re–order (3) into a more convenient form as follows. We define $\tilde{\boldsymbol{H}}_l(\boldsymbol{\tau}) \in \mathcal{R}^{M\times K}$ as

$$\tilde{\boldsymbol{H}}_l(\boldsymbol{\tau}) = \left[\tilde{\boldsymbol{h}}_l(\tau_0), \tilde{\boldsymbol{h}}_l(\tau_1), \ldots, \tilde{\boldsymbol{h}}_l(\tau_{K-1})\right], \tag{7}$$

where $\tilde{\boldsymbol{h}}_l(\tau_k)$ is the lth column in the interpolation matrix $\tilde{\boldsymbol{H}}(\tau_k)$, and a signal vector $\boldsymbol{a}(n) \in \mathcal{R}^{K\times 1}$ as

$$\boldsymbol{a}(n) \triangleq [s_0(n), s_1(n), \ldots, s_{K-1}(n)]^T.$$

Then, (3) can be expressed as

$$\boldsymbol{y}(n) = \sum_{l=0}^{L-1} \tilde{\boldsymbol{H}}_l(\boldsymbol{\tau})\boldsymbol{a}(n-l) + \sigma_w \boldsymbol{w}(n), \tag{8}$$

where the signal vector $\boldsymbol{a}(n)$ for $l = 1, ..., L-1$ can be considered known since it consists only of past values of the sources, $s_k(n)$ for $k = 0, ..., K-1$. Accordingly, we define a vector $\boldsymbol{z}(n)$ as

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

and hence we may rewrite (8) in the following form:

$$\boldsymbol{z}(n) = \tilde{\boldsymbol{H}}_0(\boldsymbol{\tau})\boldsymbol{a}(n) + \sigma_w \boldsymbol{w}(n), \tag{10}$$

which represents the desired form of the model. This model can accommodate either narrowband or wideband sources, without change of structure or parameters [6]. Furthermore, all quantities in (10), including the data, are pure real, which leads to significant savings in computations and in hardware.

The posterior distribution $\pi(\cdot|\boldsymbol{Z})$ of the parameters given the data is now developed. 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. In the case of a uniform linear array of M sensors, we may define a set of N snapshots from (10) as $\boldsymbol{Z} = [\boldsymbol{z}(1), \ldots, \boldsymbol{z}(N)]$. Hence the desired posterior distribution of the parameters is given as

$$\pi(\boldsymbol{a}, \boldsymbol{\tau}, \sigma_w^2, k|\boldsymbol{Z}) \propto p(\boldsymbol{Z}|\boldsymbol{a}, \boldsymbol{\tau}, \sigma_w^2, k) p(\boldsymbol{a}|k, \boldsymbol{\tau}, \delta^2\sigma_w^2) \times p(\boldsymbol{\tau}|k) p(\sigma_w^2) p(k), \tag{11}$$

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

$$\ell(\boldsymbol{Z}|\boldsymbol{a}, \boldsymbol{\tau}, \sigma_w^2, k) = \prod_{n=1}^{N} \mathcal{N}\left(\tilde{\boldsymbol{H}}_0(\boldsymbol{\tau})\boldsymbol{a}(n), \sigma_w^2 \boldsymbol{I}\right). \tag{12}$$

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]

$$p(\boldsymbol{a}|k, \boldsymbol{\tau}, \delta^2\sigma_w^2) = \prod_{n=1}^{N} \mathcal{N}\left(0, \delta^2\sigma_w^2 \left[\tilde{\boldsymbol{H}}_0^T(\boldsymbol{\tau})\tilde{\boldsymbol{H}}_0(\boldsymbol{\tau})\right]^{-1}\right), \tag{13}$$

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

$$p(\boldsymbol{\tau}|k) = \mathcal{U}\left[-T_{max}, T_{max}\right]^k. \tag{14}$$

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

$$p(\sigma_w^2) = \mathcal{IG}\left(\frac{\nu_0}{2}, \frac{\gamma_0}{2}\right). \tag{15}$$

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

$$p(k) = \frac{\Lambda^k}{k!} \exp(-\Lambda). \tag{16}$$

In this problem, the only quantities of interest are $\boldsymbol{\tau}$ and k, and the others, $\boldsymbol{a}$ and σ_w^2, can be treated as nuisance parameters which are integrated out analytically. The resulting desired posterior distribution can then be expressed as [3][6]

$$\pi(\boldsymbol{\tau}, k|\boldsymbol{Z}) \propto \frac{1}{(1+\delta^2)^{Nk/2}} \left(\frac{\Lambda}{2T_{max}}\right)^k \frac{\exp(-\Lambda)}{k!} \left(\gamma_0 + \mathrm{tr}\left(\tilde{\boldsymbol{P}}_{\mathrm{H}_0}^{\perp}(\boldsymbol{\tau})\hat{\boldsymbol{R}}_{\mathrm{zz}}\right)\right)^{-\left(\frac{MN+\nu_0}{2}\right)}, \tag{17}$$

where $\hat{\boldsymbol{R}}_{zz} = \sum_{n=1}^{N} \boldsymbol{z}(n)\boldsymbol{z}^T(n)$, the sample covariance matrix of $\boldsymbol{z}(n)$. The $(\boldsymbol{\tau}, k)$ are estimated from this distribution using the reversible jump MCMC procedure [7], which

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

has been discussed in the narrowband context in [3] and [8]. It is discussed in further detail in the wideband context in [6]. The reversible jump MCMC procedure is an extension of the Metropolis Hastings MCMC algorithm [9][10], and uses proposal distributions corresponding to different dimensions to yield an approximation to the joint posterior distribution of the model order and the parameters. Once the $k, \boldsymbol{\tau}$ are estimated, the amplitude vector $\boldsymbol{a}(n)$ can be estimated for $n = 1, \ldots, N$ according to the method in [6].

3. 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, corresponding to a non–informative prior.

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 3.1. Parameters for the experiments.

According to the parameters in Table 3.1, we generate $K = 2$ sources which are Gaussian processes, each with zero mean and variance equal to $\delta^2\sigma_w^2$, and bandlimited to $f \in [100, 400]$ Hz (60% relative bandwidth). The incident angles $\boldsymbol{\theta}$ are separated by an angle less than a half standard beamwidth [1] at the lowest frequency of interest. An array of $M = 8$ sensors is used to generate $N = 50$ snapshots using (8), with an $SNR = 14$dB. The corresponding hyperparameter $\delta^2 = 25.12$ is assumed known and constant.

The proposed algorithm randomly initializes all unknown parameters. 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, as shown in Fig. 2. Table 3.2 summarizes a comparison between the true and estimated values of $\boldsymbol{\tau}$. Fig. 3 shows that the signal amplitudes are well separated and restored by the proposed MCMC method. The mean-squared errors of the restored signals relative to the true signal amplitudes are -16.19dB and -15.97dB, respectively.

The variances of the estimated $\boldsymbol{\tau}$ are plotted in Fig. 4 along with the respective theoretical CRLBs [11]. The algorithm is applied to 100 independent trials using the parameter values listed in Table 3.1, over a range of SNR from

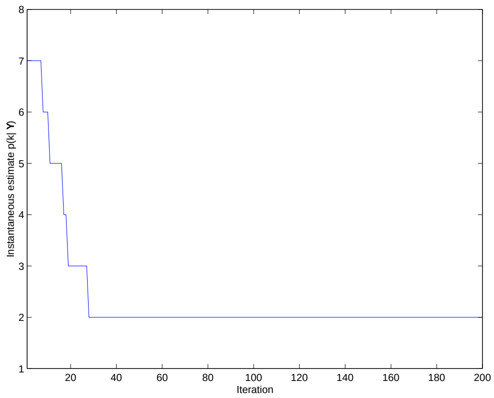

Fig. 1. Instantaneous estimate of $p(k|\boldsymbol{Y})$, for the first 200 iterations of the chain.

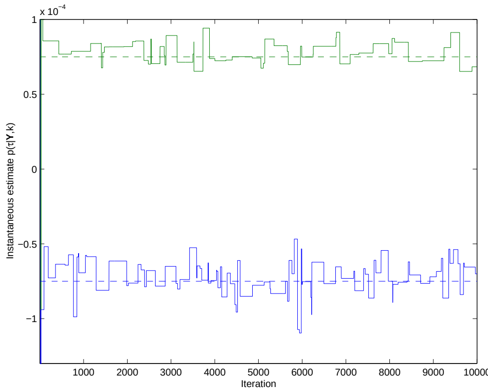

Fig. 2. Instantaneous estimate of the delays $\boldsymbol{\tau}$ for the two sources after burn–in: the solid lines are the estimates and the dashed lines are the true values

-5dB to 18dB. As shown in Fig. 4, the variances of the estimates from these trials approach the CRLB closely. The algorithm starts to break down for SNR levels lower than -2dB for this set of parameters. The reasons why the variances do not come closer to the theoretical CRLB are [6]: 1) interpolation errors when a non-ideal interpolation function is used and 2) the suboptimal procedure for estimating the source amplitudes. Further simulation results show [6] that the probability of an error in detection of the model order tends to diminish toward zero with increasing number of snapshots, N, for SNR values above threshold.

4. CONCLUSION

A novel structure for wideband array signal processing is proposed. 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 re-

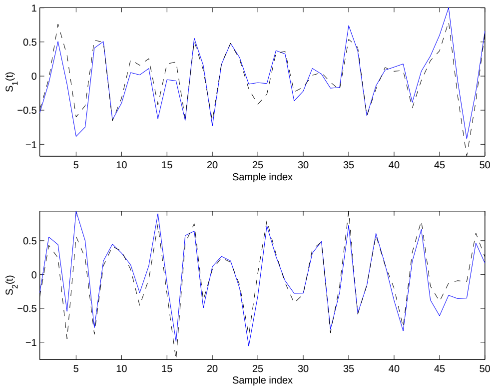

Fig. 3. A comparison between the true and the restored amplitudes using the proposed MCMC method for one realization: solid lines correspond the restored amplitudes using MCMC and the dashed lines correspond the true amplitudes.

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

Table 3.2. Comparison between the true and estimated parameters.

versible jump MCMC procedure, and the source amplitudes can also be recovered. Simulation results support the effectiveness of the method, and demonstrate reliable detection of the number of sources and estimation of TOAs in white noise environment with a single linear array.

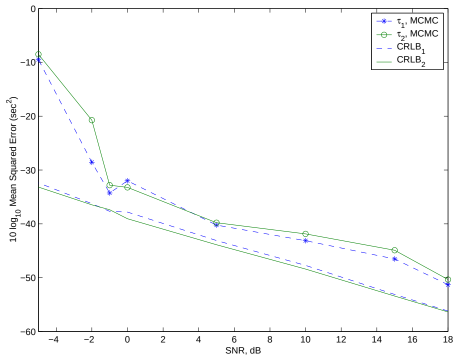

Fig. 4. Mean squared error of $\boldsymbol{\tau}$, and the corresponding CRLB, versus SNR.

5. REFERENCES

[1] D.H. Johnson, "The application of spectral estimation to bearing estimation problem," *Proceedings of the IEEE*, vol. 70, pp. 1018–1028, Sept. 1982.

[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," *IEEE Transactions on Signal Processing*, vol. 47, no. 10, pp. 2667–2676, Oct. 1999.

[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] J.P. Reilly and S. Haykin, "Maximum likelihood receiver for low-angle tracking radar Part 2: the non-symmetric case," *IEE Proceedings*, vol. 129 Pt.F, no. 5, pp. 331–340, Oct. 1982.

[6] 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.

[7] P.J. Green, "Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, 1995.

[8] Jean-René Larocque and James P. Reilly, "Reversible jump MCMC for joint detection and estimation of directions of arrival in coloured noise," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 231–240, Feb. 2002.

[9] W.K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[10] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman and Hall, New York, 1998.

[11] William Ng, James P. Reilly, and Thia Kirubarajan, "The derivation of the theoretical CRLB for array signal processing for wideband signals," May 2002, Available at http://www.ece.mcmaster.ca/∼reilly.