Wide band Channel Characterization in Coloured Noise using the Reversible Jump Markov Chain Monte Carlo Algorithm

Jean-René Larocque and James P. Reilly*

Department of Electrical and Computer Engineering,

McMaster University, 1280 Main St. W.,

Hamilton, Ontario,

Canada L8S 4K1

Submitted to IEEE Transactions on Vehicular Technology, June, 2001.

Abstract

This paper presents a novel solution to the difficult problem of joint detection/estimation of the parameters characterizing a multipath channel using an array of sensors in coloured noise with unknown arbitrary covariance structure. The proposed solution is defined in the Bayesian framework and uses the Reversible Jump Markov Chain Monte Carlo method to jointly estimate the number of multipath components, their directions of arrival and their times of arrival, after integration of the quickly-varying amplitude parameters and the unknown noise covariance matrix.

The developed method is applied to simulations and, more interestingly, with emphasis to real-life propagation measurements to show the performance and robustness of the approach.

Permission to publish this abstract separately is granted.

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

I. Introduction

In recent years there has been a considerable amount of work done in the estimation of time delays (or times of arrival (TOAs)) and directions of arrival (DOAs) of the components of a multipath channel [1][2][3][4][5][6]. This is a central problem in many fields including radar, sonar, and wireless communications, for both the indoor and outdoor scenarios. In the wireless case, knowledge of the statistics of the TOA/DOA profile of typical multipath channels is necessary to determine the complexity of proposed receivers during the design process and to evaluate the performance of a given wireless configuration. Further, where the receiver is equipped with an array of antennas, the knowledge of the channel is sufficient for the difficult problem of computing the downlink beamformer weights in a frequency–domain duplexed (FDD) system [7], or to locate a user in a cellular cell (911 problem). In radar and sonar applications, knowledge of the channel is necessary for recovering time signatures of the sources of interest.

In this paper, we propose the use of the reversible jump Markov chain Monte Carlo (RJMCMC) technique [8] for joint blind[1] estimation of the number of incident multipath components (model order), and the channel TOA/DOA information in the presence of noise with arbitrary covariance structure, using arrays of sensors. The proposed solution differs from [1] in that the number of multipath components (reflections) is not assumed known in our case. In [2], again, the number of reflections is assumed known, or is *a priori* estimated with the MDL approach. This limits the application of the proposed algorithm to white noise scenarios. Although another algorithm was developed by Wax [3] for the coloured noise case, it appears that it is very sensitive to initialization and local minima [4]. More recently, other methods have been proposed to address this difficult problem. In [5], the number of multipath components is assumed known. Lastly, the application of the SAGE algorithm in [6] is limited to the white noise case. In many other papers, where the joint detection of the number of multipath components and the estimation of the parameters is addressed, including [4], only the directions of arrival are estimated, or only the temporal channel impulse response is estimated, but never jointly. It is shown in [2] that individual estimation of either the TOA or DOA parameters is suboptimal, and

[1] By "blind" we mean no training sequence is required nor is the transmitted source sequence assumed known.

that significantly improved performance is obtained from *joint* TOA/DOA estimation, as is proposed in this paper.

As part of this work, a multipath measurement campaign was conducted on the campus of McMaster University, Hamilton, Ontario, Canada, for the purposes of acquiring measurements of real multipath propagation channels. The proposed method was then applied to these measurements and the results are found to correspond well to those obtained from geographical truthing.

The paper is organized in the following manner. In Section II, we describe the formulation of the problem, state the necessary assumptions and define the notation. In Section III, we develop the statistical model for the received signal in terms of the parameters of interest, where the undesired nuisance parameters are integrated out analytically. In Section IV, we then describe the reversible jump MCMC procedure for joint determination of model order and TOA/DOA estimation. Results from simulated data are presented in Section V. In Section VI, we describe the measurement technique used to gather the real propagation measurements and we apply the proposed method to these measurements. Conclusions are given in Section VII.

Notation: (Column) vectors are represented as bold lower case symbols and matrices as bold upper case symbols. Superscripts ^{T} and ^{H} denote the transpose and Hermitian transpose operation, respectively. The symbol $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the multivariate complex normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ and $\mathcal{U}_{[a,b]^k}$ denotes the $k-$dimensional uniform distribution over $[a, b]$.

II. Problem Formulation

Consider an array composed of M sensors with arbitrary locations and arbitrary spatial response. The array is immersed in a multipath propagation environment where the corresponding channel consists of k_o discrete reflections with time delays (TOAs) $\boldsymbol{\tau} = [\tau_1, \dots, \tau_{k_o}]^T$, incident on the array from directions $\boldsymbol{\phi} = [\phi_1, \dots, \phi_{k_o}]^T$. We assume the sensors and the reflection points are coplanar and that the reflecting points are in the far field of the array, so that ϕ_k is the DOA of the k_{th} reflection. The signal vector $\boldsymbol{y}(n) \in \mathcal{C}^M$

received by the array at time t can be written as

$$\boldsymbol{y}(t) = \sum_{k=1}^{k_o} a_k(t-\tau_k)\boldsymbol{s}(\phi_k) + \boldsymbol{\nu}(t) \tag{1}$$

where

- $a_k(t)$ is the amplitude parameter associated with the k_{th} reflection
- τ_k is the delay (TOA) of the k_{th} reflection
- $\boldsymbol{s}(\phi_k) \in \mathcal{C}^M$ is the array response vector from direction ϕ_k
- $\boldsymbol{\nu}(t) \in \mathcal{C}^M$ is the noise vector at time t. $\boldsymbol{\nu}$ is assumed to be uncorrelated with the signal, zero mean, distributed as $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_\nu)$, where $\mathbf{\Sigma}_\nu \in \mathcal{C}^{M\times M}$ is unknown, positive definite and Hermitian, but otherwise arbitrary.

We choose to express the problem in the context of a single user in a CDMA scenario[2], where the number of chips in one transmitted symbol is denoted by P, and the chip duration is $\triangle T$. After sampling $\boldsymbol{y}(t)$ in (1) in time, we may now express the received signal over the P chips comprising the n_{th} symbol in the form

$$\boldsymbol{Y}(n) = \boldsymbol{S}(\boldsymbol{\phi})\boldsymbol{A}(n)\boldsymbol{T}(\boldsymbol{\tau}) + \boldsymbol{V}(n), \qquad n = 1, \dots, N, \tag{2}$$

where

- $\boldsymbol{Y}(\cdot)$ and $\boldsymbol{V}(\cdot)$ have P columns consisting of their lower- case counterparts
- $\boldsymbol{S}(\boldsymbol{\phi}) \in \mathcal{C}^{M\times k_o} = [\boldsymbol{s}(\phi_1), \dots, \boldsymbol{s}(\phi)_{k_o}]$
- $\boldsymbol{A}(n) \in \mathcal{C}^{k_o\times k_o}$ is a diagonal matrix. The diagonal elements contain the signal amplitudes at the nth symbol
- $\boldsymbol{T}(\boldsymbol{\tau}) \in \mathcal{C}^{k_o\times P}$. Each row of $\boldsymbol{T}(\boldsymbol{\tau})$ consists of zero elements, except for a single one in the pth position. This element indicates that the relative delay of the corresponding scattering component is $p\Delta T$.
- N is the number of observed snapshots.

Given only observations $[\boldsymbol{Y}(1), \dots, \boldsymbol{Y}(N)]$, our objective is to jointly estimate the number of scatterers, k_o, their directions of arrival, $\boldsymbol{\phi} \in [0, 2\pi]^{k_o}$ and their times of arrival within the resolution $\triangle T$ of one chip, represented by the vector of integers $\boldsymbol{\tau} \in [0, P]^{k_o}$.

[2]This is done solely for ease of presentation and is not necessary for the development of the central ideas of this paper. It allows us to consider a single user, isolated by its characteristic CDMA code.

Note that since the number of scatterers k_o is unknown, the dimensions of the parameters $\boldsymbol{\phi}$ and $\boldsymbol{\tau}$ are also unknown. We therefore denote them as $\boldsymbol{\phi}_k$ and $\boldsymbol{\tau}_k$ respectively, where k is the hypothesized number of signals.

We make the further set of assumptions on the model:

- The array steering (response) vector $\boldsymbol{s}(\phi)$ is known for all ϕ within the field of view
- $\boldsymbol{S}(\boldsymbol{\phi})$ is full rank
- the noise $\boldsymbol{\nu}(\cdot)$ is *iid* between symbols
- $\boldsymbol{\phi}$ and $\boldsymbol{\tau}$ are stationary over the entire N observations.
- the symbol amplitudes a_k are *iid* between symbols
- (narrowband assumption) The amplitude of the wavefront is assumed to be approximately stationary during the time taken for the wave to traverse the array.

III. Development of the posterior distribution

The model described by (2) can now be rearranged to a more familiar form using Kronecker algebra, as follows:

$$\boldsymbol{y}(n) = \boldsymbol{Z}(\boldsymbol{\tau}, \boldsymbol{\phi})\boldsymbol{b}(n) + \boldsymbol{\nu}(n), \qquad n = 1, \ldots, N, \tag{3}$$

where:

$$\boldsymbol{y}(n) = \mathrm{vec}(\boldsymbol{Y}(n)), \tag{4}$$

$$\boldsymbol{b}(n) = \mathrm{vec}(\boldsymbol{A}(n)), \tag{5}$$

$$\boldsymbol{Z}(\boldsymbol{\tau}, \boldsymbol{\phi}) = \boldsymbol{T}^T(\boldsymbol{\tau}) \otimes \boldsymbol{S}(\boldsymbol{\phi}). \tag{6}$$

where $\mathrm{vec}(\cdot)$ is the *vec* operator and $\otimes$ is the Kronecker matrix product. Furthermore, noting that the matrix $\boldsymbol{A}(n)$ is diagonal, the vector $\boldsymbol{b}(n)$ as defined only operates on a few columns of $\boldsymbol{Z}$. Regrouping these useful columns into a new matrix $\boldsymbol{H}(\boldsymbol{\tau}, \boldsymbol{\phi})$ we have:

$$\boldsymbol{y}(n) = \boldsymbol{H}(\boldsymbol{\tau}, \boldsymbol{\phi})\boldsymbol{a}(n) + \boldsymbol{\nu}(n) \quad n = 1, \ldots, N, \tag{7}$$

where $\boldsymbol{a}(n)$ holds the diagonal elements of $\boldsymbol{A}(n)$, and the matrix $\boldsymbol{H}(\boldsymbol{\tau}, \boldsymbol{\phi})$ defines the space-time structure of the multipath. This form is more familiar and can now easily be analyzed in the Bayesian framework.

Since the N snapshots are *iid*, the total likelihood function of all the data can be expressed in the following form:

$$
\begin{aligned}
p(\boldsymbol{Y}|\boldsymbol{\phi},\boldsymbol{\tau},\boldsymbol{A},\boldsymbol{\Sigma}_\nu,k) &= \frac{1}{\pi^{NMP}|\boldsymbol{\Sigma}_\nu|^N} \times \\
& e^{-\sum_{n=1}^{N}\left(\boldsymbol{y}(n)-\boldsymbol{H}(\boldsymbol{\phi},\boldsymbol{\tau})\boldsymbol{a}(n)\right)^H\boldsymbol{\Sigma}_\nu^{-1}\left(\boldsymbol{y}(n)-\boldsymbol{H}(\boldsymbol{\phi},\boldsymbol{\tau})\boldsymbol{a}(n)\right)},
\end{aligned}
\tag{8}
$$

where the matrices $\boldsymbol{Y}$ and $\boldsymbol{A}$ represent the sets $\{\boldsymbol{y}(n)\}_{n=1}^N$ and $\{\boldsymbol{a}(n)\}_{n=1}^N$, respectively. This likelihood function can be simplified by integrating out the undesired parameters $\boldsymbol{\Sigma}_\nu$, and $\boldsymbol{A}$. To proceed with this step, we first define an orthonormal matrix $\boldsymbol{U}(\boldsymbol{\phi},\boldsymbol{\tau},k) \in \mathcal{C}^{MP\times MP}$ for a hypothesized model order k, as in [9][10]:

$$
\boldsymbol{U}(\boldsymbol{\phi},\boldsymbol{\tau},k) = [\underbrace{\boldsymbol{U}_s(\boldsymbol{\phi},\boldsymbol{\tau},k)}_{MP\times k} \quad \underbrace{\boldsymbol{U}_\nu(\boldsymbol{\phi},\boldsymbol{\tau},k)}_{MP\times(MP-k)}],
\tag{9}
$$

where $\boldsymbol{U}_s(\boldsymbol{\phi},\boldsymbol{\tau},k) \in \mathcal{H}$, the signal subspace, and $\boldsymbol{U}_\nu(\boldsymbol{\phi},\boldsymbol{\tau},k) \in \mathcal{N}$, the noise subspace. We now transform the received data $\boldsymbol{y}(n)$ into $\mathcal{H}$ and $\mathcal{N}$ to form a signal component $\boldsymbol{z}_s(n)$ and a noise component $\boldsymbol{z}_\nu(n)$ respectively as

$$
\boldsymbol{z}_s(n) = \boldsymbol{U}_s^H(\boldsymbol{\phi},\boldsymbol{\tau},k)\boldsymbol{y}(n),
\tag{10}
$$

and

$$
\boldsymbol{z}_\nu(n) = \boldsymbol{U}_\nu^H(\boldsymbol{\phi},\boldsymbol{\tau},k)\boldsymbol{y}(n).
\tag{11}
$$

The new parameters $\boldsymbol{z}_s(n)$ and $\boldsymbol{z}_\nu(n)$ are both Gaussian. For arbitrary $\boldsymbol{\Sigma}_\nu$, the noise components of $\boldsymbol{z}_s(n)$ and $\boldsymbol{z}_\nu(n)$ may not be uncorrelated; however, for the sake of tractable analysis, we assume them to be so. In the neighbourhood of the true values of the parameters, it can be shown [4] that $\boldsymbol{z}_s(n)$ is distributed as $\mathcal{N}(\tilde{\boldsymbol{a}}(n),\boldsymbol{C})$, and $\boldsymbol{z}_\nu(n)$ is distributed as $\mathcal{N}(\boldsymbol{0},\boldsymbol{W})$, with covariance matrix $\boldsymbol{W} \in \mathcal{C}^{(MP-k)\times(MP-k)} \triangleq \mathrm{E}\{\boldsymbol{U}_\nu^H\boldsymbol{y}_n\boldsymbol{y}_n^H\boldsymbol{U}_\nu\}$. Here, it is not necessary to define $\boldsymbol{C}$ or $\tilde{\boldsymbol{a}}$, since they are integrated out analytically in the final posterior distribution. The joint likelihood function of $\boldsymbol{z}_s$ and $\boldsymbol{z}_\nu$ is then given as:

$$
\begin{aligned}
& p(\boldsymbol{Z}_s,\boldsymbol{Z}_\nu|\tilde{\boldsymbol{A}},\boldsymbol{\phi},\boldsymbol{\tau},k,\boldsymbol{W}^{-1}) \approx \pi^{-Nk}|\boldsymbol{C}^{-1}|^N \\
& \times \exp\left\{-\sum_{n=1}^{N}(\boldsymbol{z}_s(n)-\tilde{\boldsymbol{a}}(n))^H\boldsymbol{C}^{-1}(\boldsymbol{z}_s(n)-\tilde{\boldsymbol{a}}(n))\right\} \\
& \times \pi^{-N(MP-k)}|\boldsymbol{W}^{-1}|^N \exp\left\{-\sum_{n=1}^{N}\boldsymbol{z}_\nu^H(n)\boldsymbol{W}^{-1}\boldsymbol{z}_\nu(n)\right\}.
\end{aligned}
\tag{12}
$$

To complete the model, prior distributions are chosen to be non-informative where possible. When convenient, we also choose the structural form of these distributions for their desirable conjugate properties. The priors distributions are described as follows:

• $\tilde{\boldsymbol{A}}$ is assigned a conjugate non-informative prior distribution described as a Gaussian function with a large covariance matrix $\boldsymbol{D}$ (compared to $\boldsymbol{C}$), and zero mean. Thus,

$$p(\tilde{\boldsymbol{A}}|\boldsymbol{\phi}_k, \boldsymbol{\tau}, k, \boldsymbol{W}^{-1}) = \prod_{n=1}^{N} \mathrm{N}(\boldsymbol{0}, \boldsymbol{D}), \tag{13}$$

where $\boldsymbol{D} \triangleq d^2 \boldsymbol{I}_k$, which assumes the projected signals are independent with the same large variance. The choice of the hyper-parameter d is addressed later.

• The prior distributions for both $\boldsymbol{\phi}$ and $\boldsymbol{\tau}$ are chosen to be uniform:

$$p(\boldsymbol{\phi}|k) = \mathcal{U}_{[0,2\pi]^k} \qquad p(\boldsymbol{\tau}|k) = \frac{1}{P^k}. \tag{14}$$

• The prior on k is chosen to be Poisson with expectation Λ:

$$p(k) = \Lambda^k e^{-\Lambda}/k!. \tag{15}$$

Although not strictly non-informative, this prior is used to speed up convergence. In previous simulations, a uniform prior over a range $[0, k_{max}]$ was successfully used. This mildly informative prior does not significantly influence the estimate of $\hat{k}$, and it has been shown through simulations that the algorithm is robust to the value of the hyper-parameter Λ.

• $\boldsymbol{W}^{-1}$: We use a non-informative multi-dimensional Jeffreys' prior for the unknown transformed noise covariance matrix [9]

$$p(\boldsymbol{W}^{-1} \mid \boldsymbol{\phi}, \boldsymbol{\tau}, k) \propto \mid \boldsymbol{W}^{-1} \mid^{-(MP-k)}. \tag{16}$$

The posterior distribution, after carrying out the integration of the nuisance parameters and ignoring the constant terms, is then (see [4] for more detail):

$$\begin{aligned} p(k, \boldsymbol{\phi}, \boldsymbol{\tau}|\boldsymbol{Z}_\nu) \propto \; & \frac{\pi^{\frac{1}{2}(MP-k)(MP-k-1)} \prod_{i=1}^{MP-k} \Gamma(N-i+1)}{(2\pi P/\Lambda)^k k! (d^2)^{kN}} \\ & \times \left| N\hat{\boldsymbol{W}}(\boldsymbol{\phi}, \boldsymbol{\tau}, k) \right|^{-N}, \end{aligned} \tag{17}$$

with $N\hat{\boldsymbol{W}}(\boldsymbol{\phi},\boldsymbol{\tau},k) \triangleq \sum_{n=1}^{N} \boldsymbol{z}_\nu(n)\boldsymbol{z}_\nu(n)^H$. Take note that this function depends only on the slowly varying parameters of interest. The objective is to estimate the parameters of this highly non-linear function, as the Maximum A Posteriori (MAP) estimates:

$$\{\hat{k},\hat{\boldsymbol{\phi}},\hat{\boldsymbol{\tau}}\} = \arg \max_{k,\boldsymbol{\phi},\boldsymbol{\tau}\in\Theta} p(k,\boldsymbol{\phi},\boldsymbol{\tau}|\mathbf{Z}_{\mathbf{v}}). \tag{18}$$

A. White noise hypothesis

In the event that the noise is known to be spatially white, the integration of the nuisance parameters is straightforward. The detailed development is presented in [11]. The resulting posterior distribution is given as:

$$\begin{aligned} p_{white}(\boldsymbol{\phi},\boldsymbol{\tau},k|\{\boldsymbol{y}(n)\}) \propto & \left(\sum_{n=1}^{N} \boldsymbol{y}^H(n)\boldsymbol{P}_H^{\perp}(\boldsymbol{\phi},\boldsymbol{\tau},k)\boldsymbol{y}(n)\right)^{NMP} \\ & \times \frac{\Lambda^k}{(2\pi)^k P^k k!(1+\delta^2)^{kN}}, \end{aligned} \tag{19}$$

where

$$\boldsymbol{P}_H^{\perp}(\boldsymbol{\phi},\boldsymbol{\tau},k) = \boldsymbol{I} - \boldsymbol{H}\boldsymbol{M}\boldsymbol{H}^H, \tag{20}$$

$$\boldsymbol{M}^{-1} = \boldsymbol{H}^H\boldsymbol{H}(1+\delta^{-2}). \tag{21}$$

and the hyperparameter δ^2 is the estimated SNR.

However, this form remains quite sensitive to the degree of colour in the noise. Our experiences show that the performance of the detection of model order degrades very quickly in the presence of even mildly coloured noise.

IV. The Reversible Jump MCMC Algorithm

We now propose the Metropolis-Hastings (MH) algorithm[12] as a suitable MCMC method to perform the Bayesian computation in extracting the parameters of interest from the posterior distribution (17). MCMC computational techniques are now emerging as important methods in the signal processing arena. The method proposed here is similar to that presented in [11] which describes the white noise case and that in [4] which describes the coloured noise case, but for DOA estimation alone. The following presentation is therefore brief.

The idea of the MH algorithm is to establish an ergodic Markov chain, which under weak conditions, produces samples distributed according to an invariant distribution equal to the desired posterior distribution.

With the MH method, assume at the ith iteration we are in state $(\boldsymbol{\phi}^{(i)}, \boldsymbol{\tau}^{(i)}, k^{(i)})$. A candidate $(\boldsymbol{\phi}^{\star}, \boldsymbol{\tau}^{\star}, k^{\star})$ for the next state of the chain is drawn at random from a proposal distribution $q(\cdot|\cdot)$, which is chosen to be easy to sample from and which may be conditional on $(\boldsymbol{\phi}^{(i)}, \boldsymbol{\tau}^{(i)}, k^{(i)})$. An acceptance ratio r is then generated according to:

$$r = \frac{p(\boldsymbol{\phi}^{\star}, \boldsymbol{\tau}^{\star}, k^{\star}|\boldsymbol{Z}_{\nu})q(\boldsymbol{\phi}^{(i)}, \boldsymbol{\tau}^{(i)}, k^{(i)}|\boldsymbol{\phi}^{\star}, \boldsymbol{\tau}^{\star}, k^{\star})}{p(\boldsymbol{\phi}^{(i)}, \boldsymbol{\tau}^{(i)}, k^{(i)}|\boldsymbol{Z}_{\nu})q(\boldsymbol{\phi}^{\star}, \boldsymbol{\tau}^{\star}, k^{\star}|\boldsymbol{\phi}^{(i)}, \boldsymbol{\tau}^{(i)}, k^{(i)})}. \tag{22}$$

An acceptance parameter α is then defined as

$$\alpha = \min\{r, 1\}. \tag{23}$$

The proposed candidate $(\boldsymbol{\phi}^{\star}, \boldsymbol{\tau}^{\star}, k^{\star})$ is accepted as the current state at iteration $(i+1)$ with probability α.

The set of accepted candidates represents a set of samples drawn form the posterior distribution of interest. These samples can then be used to construct a histogram from which the desired statistical inferences can be made. The difficulty with MCMC methods and the traditional MH algorithm is that the parameter space must be of fixed dimension.

In our application, since the dimension of the parameter space $\boldsymbol{\Phi} = \{\boldsymbol{\phi}_k, \boldsymbol{\tau}_k, k\}$ varies with k, we use the *reversible jump* MCMC method [8]. This procedure samples directly from different model orders from the joint distribution on $\{\boldsymbol{\phi}_k, \boldsymbol{\tau}_k, k\}$. In effect, the process jumps between subspaces of different dimensions, thus visiting all model orders for $k \in [0, \ldots, k_o]$.

In the reversible jump case, candidate samples are chosen from a *set* of proposal distributions, which are randomly accepted according to an acceptance ratio that ensures reversibility, and therefore the invariance of the Markov chain with respect to the desired posterior distribution. Here, we choose our set of proposal distributions to correspond to the following set of moves:

1. the *birth* move, valid for $k < M$. Here, a new scattering point is proposed at random on $(0, 2\pi]$ and $[0, P]$.

2. the *death* move, valid for $k > 0$. Here, a randomly chosen scatterer is removed.
3. the *update* move. Here, the parameters describing the propagation channel are updated for a fixed value of k.

   The probabilities for choosing each move are denoted u_k, b_k and d_k, respectively, such that $u_k + b_k + d_k = 1$ for all k. In accordance with [8], we choose:

$$b_k = c\min\{\frac{p(k+1)}{p(k)}, 1\}, \quad d_{k+1} = c\min\{\frac{p(k)}{p(k+1)}, 1\}, \tag{24}$$

   where $p(k)$ is the prior probability of the kth model according to (15), and c is the tuning parameter for the ratio of update moves to jump moves. We choose $c = 0.5$ so that the probability of a jump is between 0.5 and 1 at each iteration [8]. The overall description of the reversible jump MCMC algorithm is determined by the choice of move at each iteration. This description is summarized as follows:

Reversible Jump MCMC

1. Initialization: set $\mathbf{\Phi}^{(0)} = (\boldsymbol{\phi}^{(0)}, \boldsymbol{\tau}^{(0)}, k^{(0)})$ according to the prior distributions.
2. Iteration i,
   - Sample $u \sim \mathcal{U}_{[0,1]}$
   - if $(u < b_{k^{(i)}})$ then execute a "birth move" (see IV-B)
   - else if $(u < b_{k^{(i)}} + d_{k^{(i)}})$ then execute a "death move" (see IV-B)
   - else, execute an update move (see IV-A)
3. $i \leftarrow i + 1,$ goto step 2

The algorithm repeats for $L \to \infty$ iterations according to procedure described earlier. Then, estimates of the incident angles and delays and the model order can be obtained using, e.g., (18), where the histogram of the collected samples is used an approximation to the true posterior density.

These proposal distributions for each of the three move types are chosen heuristically. These choices satisfy the reversibility conditions for the Markov chain to converge to the desired invariant distribution. Other possibilities certainly exist. The ones chosen have been shown to work well in simulations.

A. Update move

Here, we assume that the current state of the algorithm is $(\boldsymbol{\phi}_k, \boldsymbol{\tau}_k, k)$. When the update move is selected, the algorithm samples all parameters for k fixed. The proposal distributions in the general case of global exploration of the space are given as

$$q(\boldsymbol{\phi}^{(i)}, \boldsymbol{\tau}^{(i)}, k) = q(\boldsymbol{\phi}^{\star}, \boldsymbol{\tau}^{\star}, k) = \frac{1}{(2\pi P)^k}. \tag{25}$$

The acceptance ratio $r = r_{update}$ for the update move, in the case of coloured noise is obtained by substituting (25) and (17) into (22) to give:

$$r_{update}(\boldsymbol{\phi}_k^{\star}, \boldsymbol{\tau}_k^{\star}, k, \boldsymbol{\phi}_k, \boldsymbol{\tau}_k, k) = \frac{\left| N\hat{\boldsymbol{W}}(\boldsymbol{\phi}^{\star}, \boldsymbol{\tau}^{\star}, k) \right|^{-N}}{\left| N\hat{\boldsymbol{W}}(\boldsymbol{\phi}, \boldsymbol{\tau}, k) \right|^{-N}}, \tag{26}$$

$$\alpha_{update} = \min[r_{update}, 1]. \tag{27}$$

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

For the specific case at hand, the performance of the proposed method is enhanced by selecting randomly between two types of proposal distributions for the update case rather that the one given by (25): one type involves a global exploration of the parameter space, while the other involves a local exploration. The method is summarized as follows. (Further detail is presented in [11]).

Update Move (Metropolis-one-at-the-time)

At iteration i, DO

• **Sampling the directions of arrival**:

– Propose new directions of arrival, maintaining fixed values for $\boldsymbol{\tau}^{(i)}$. The choice of proposal distribution is described as follows:

– Sample $u \sim \mathcal{U}_{[0,1]}$

∗ if $u < 0.5$, then propose a global exploration

$$\boldsymbol{\phi}_k^{\star} \sim \mathcal{U}_{[0,2\pi]^k} \tag{28}$$

∗ else, propose a local exploration [3]

$$\boldsymbol{\phi}_k^\star \sim \mathcal{N}(\boldsymbol{\phi}_k^{(i)}, \boldsymbol{\Sigma}_\phi) \tag{29}$$

where $\boldsymbol{\Sigma}_\phi$ is a covariance matrix such that samples from (29) are closely clustered around $\boldsymbol{\phi}_k^{(i)}$.

– Evaluate α_{update} with (27).

– Sample $u \sim \mathcal{U}_{[0,1]}$.

– if $(u \leq \alpha_{update})$ then the state of the Markov Chain at iteration $i+1$ becomes $(\boldsymbol{\phi}_k^\star, k)$, else it remains at $(\boldsymbol{\phi}_k^{(i)}, k)$.

• **Sampling the times of arrival**. (This process is analogous to sampling the DOAs, as above).

– Propose new times of arrival, maintaining the directions of arrival fixed at their current values.

– Sample $u \sim \mathcal{U}_{[0,1]}$

∗ if $u < 0.5$, then propose a global exploration

$$\boldsymbol{\tau}_k^\star \sim \mathcal{U}_{[0,P]^k} \tag{30}$$

∗ else, propose a local exploration

$$\boldsymbol{\tau}_k^\star \sim \mathcal{N}(\boldsymbol{\tau}_k^{(i)}, \boldsymbol{\Sigma}_\tau) \tag{31}$$

– Evaluate α_{update} with (27).

– Sample $u \sim \mathcal{U}_{[0,1]}$.

– if $(u \leq \alpha_{update})$ then the state of the Markov Chain becomes $(\boldsymbol{\tau}_k^\star, k)$, else it remains at $(\boldsymbol{\tau}_k^{(i)}, k)$.

■

B. Birth and Death moves

In the death move case, we assume the current state is $(\boldsymbol{\phi}_{k+1}, \boldsymbol{\tau}_{k+1}, k+1)$ and we wish to determine whether the state is $(\boldsymbol{\phi}_k, \boldsymbol{\tau}_k, k)$ at the next iteration. This involves the removal

[3] Strictly speaking, the acceptance ratio given by (26) is not the correct form of (22) in this case, since the proposal distributions no longer cancel with the prior distributions as they do in the global case. However, in practice, the error introduced by using (26) in negligible.

of an incident signal, which is chosen randomly amongst the $(k+1)$ existing incident signals. The proposal distribution $q(\boldsymbol{\phi}_k^\star, \boldsymbol{\tau}_k^\star, k|\boldsymbol{\phi}_{k+1}, \boldsymbol{\tau}_{k+1}, k+1)$ for the death move is therefore chosen as

$$q(\boldsymbol{\phi}_k^\star, \boldsymbol{\tau}_k^\star, k|\boldsymbol{\phi}_{k+1}, \boldsymbol{\tau}_{k+1}, k+1) = p(k) \div \binom{k+1}{1} \propto \frac{\Lambda^k}{k!} \frac{1}{(k+1)}, \tag{32}$$

where $p(k)$ is the prior distribution defined in (15). The remaining values of $\boldsymbol{\phi}_k^\star$ and $\boldsymbol{\tau}_k^\star$ are set to the corresponding values of $\boldsymbol{\phi}_k^{(i)}$ and $\boldsymbol{\tau}_k^{(i)}$ respectively.

Similarly, in the birth move case, we assume the current state is $(\boldsymbol{\phi}_k, \boldsymbol{\tau}_k, k)$ and we wish to determine whether the next state is $(\boldsymbol{\phi}_{k+1}, \boldsymbol{\tau}_{k+1}, k+1)$. This involves the addition of a new incident signal, which is proposed uniformly over $(0, 2\pi]$ and over $(0, P]$. The proposal distribution $q(\boldsymbol{\phi}_{k+1}^\star, \boldsymbol{\tau}_{k+1}^\star, (k+1)|\boldsymbol{\phi}_k, \boldsymbol{\tau}_k, k)$ for the birth move is therefore

$$q(\boldsymbol{\phi}_{k+1}^\star, \boldsymbol{\tau}_{k+1}^\star, (k+1)|\boldsymbol{\phi}_k, \boldsymbol{\tau}_k, k) = p(k+1)\frac{1}{2\pi}\frac{1}{P} \propto \frac{\Lambda^{k+1}}{(k+1)!}\frac{1}{2\pi}\frac{1}{P} \tag{33}$$

A sufficient condition for the invariant distribution of the reversible jump MCMC algorithm to assume the desired distribution with respect to model order is that the Markov chain be *reversible* with respect to moves across subspaces of different model orders; i.e., the probability of moving from model order k to $k+1$ must be equal to that of moving from $k+1$ to k. It is shown [8] that reversibility is satisfied if

$$r_{death} = \frac{1}{r_{birth}}. \tag{34}$$

In the birth move case, to satisfy (34) the proposal distribution $q(\boldsymbol{\phi}^\star, \boldsymbol{\tau}^\star, (k+1)|\boldsymbol{\phi}^{(i)}, \boldsymbol{\tau}^{(i)}, k)$ in (22) is chosen to be (33), whereas the proposal distribution $q(\boldsymbol{\phi}^{(i)}, \boldsymbol{\tau}^{(i)}, k|\boldsymbol{\phi}^\star, \boldsymbol{\tau}^\star, (k+1))$ is taken as the right-hand side of (32).[4] In the death move case, the reverse assignments are made.

Using this arrangement for the proposal distributions, and using (17), the acceptance ratio $r = r_{death}$ from (22) for the death move is given as

$$\begin{aligned} r_{death}(\boldsymbol{\phi}_k^\star, \boldsymbol{\tau}_k^\star, k, \boldsymbol{\phi}_{k+1}, \boldsymbol{\tau}_{k+1}, k+1) &= \frac{\left|N\hat{\boldsymbol{W}}(\boldsymbol{\phi}^\star, \boldsymbol{\tau}^\star, k)\right|^{-N}}{\left|N\hat{\boldsymbol{W}}(\boldsymbol{\phi}, \boldsymbol{\tau}, k+1)\right|^{-N}} \\ &\times \pi^{MP-k-1}\Gamma(N-MP+k+1)(k+1)d^{2^N}. \end{aligned} \tag{35}$$

[4]To interpret this assignment on an intuitive level, (33) may be considered the proposal from order $k+1$ to k, whereas (32) is the proposal from k to $k+1$.

The quantity α_{death} is then defined according to

$$\alpha_{death} = \min[r_{death}, 1]. \tag{36}$$

The acceptance ratio $r = r_{birth}$ for the birth move is therefore

$$\alpha_{birth} = \min[1, \frac{1}{r_{death}}]. \tag{37}$$

The following block describes the algorithm for the birth move.

Birth Move

- Propose a new direction of arrival and time of arrival,

$$\boldsymbol{\phi}^{\star}_{k+1} = [\boldsymbol{\phi}^{(i)}_{k} \vdots U_{[0,2\pi]}] \qquad \boldsymbol{\tau}^{\star}_{k+1} = [\boldsymbol{\tau}^{(i)}_{k} \vdots U_{[0,P]}] \tag{38}$$

- Evaluate α_{birth} with (37).
- Sample $u \sim \mathcal{U}_{[0,1]}$.
- if $(u \leq \alpha_{birth})$ then the state of the Markov Chain becomes $(\boldsymbol{\phi}^{\star}_{k+1}, \boldsymbol{\tau}^{\star}_{k+1}, k+1)$, else it remains at $(\boldsymbol{\phi}^{(i)}_{k}, \boldsymbol{\tau}^{(i)}_{k}, k)$.

■

The description for the death move is similar, with appropriate modifications.

Remark 1: At this point, it is enlightening to take the logarithm of the previously obtained posterior distribution (17):

$$\begin{aligned}\log p(k, \boldsymbol{\phi}_k, \boldsymbol{\tau}_k | \boldsymbol{Z}_\nu) = C - N \log(\left| N\hat{\boldsymbol{W}} \right|) + kN \log(d^2) + k \log(\Lambda/2\pi) + \log(k!) \\ + \frac{1}{2}(MP-k)(MP-k-1)\log(\pi) + \log(\prod_{i=1}^{MP-k} \Gamma(N-i+1))\end{aligned} \tag{39}$$

where C is a constant independent of k. In this form, the similarities with previous model selection criteria such as AIC, MDL, D-MAP of Djuric[13] or W-MDL of Wax [3] are made apparent. The first term represents the likelihood term, while the remaining ones jointly constitute a "penalty term", which is dependent on the prior distributions of the parameters. □

C. Selection of the hyper-parameter d^2

In this section, we summarize the conditions which must be applied on the hyper-parameter d^2 in (13) for consistent determination of model order.

Let the eigenvalues of $N\hat{\boldsymbol{W}}$ in (17) at $\boldsymbol{\phi} = \boldsymbol{\phi}_o$ be given as $\lambda_1, \lambda_2, \ldots, \lambda_{MP-k}$, arranged in *ascending*, rather than the usual *descending* order. We also note from the definition of $N\hat{\boldsymbol{W}}$ that the eigenvalues λ_i are directly proportional to N, and thus they can be written as

$$\lambda_i = N\tilde{\lambda}_i, \qquad i = 1, \ldots, MP, \tag{40}$$

where $\tilde{\lambda}_i$ is the normalized version of λ_i. It is shown in [4] that the detection of model order is consistent if the hyper-parameter d^2 is chosen so that

$$\tilde{\lambda}_{MP-k_o} > \frac{d^2}{e} > \tilde{\lambda}_{MP-k_o+1}. \tag{41}$$

Strictly speaking, this procedure for determining d^2 cannot be used for consistent detection because (41) depends on the unknown k_o. However, it should be possible in the practical scenario to propose an *ad hoc* scheme to approximate (41). For example, it is usually possible to form an estimate of $\boldsymbol{\Sigma}_\nu$ during periods where it is known with reasonable certainty there are no signals present. The largest eigenvalue of this matrix could then be used as an upper bound on $\tilde{\lambda}_{k_o+1}$. If some *a priori* knowledge on the DOAs were available, then an estimate of $\boldsymbol{U}_\nu$ can be evaluated and a better estimate of $\tilde{\lambda}_{k_o+1}$ could be determined. In either case, d^2 could be given as an empirically determined constant times this eigenvalue estimate.

V. Simulation Results

The proposed algorithm for coloured noise is now applied to simulation data, generated for $k_o = 2$ scatterers with the parameters described in Table 5.1. The receiver array is composed of 5 elements. The amplitudes are *iid* Rayleigh distributed over $N = 125$ received symbols (or snapshots) for $P = 25$ chips, with an SNR of 5dB. The noise is coloured with an AR filter, the poles of which are $0.95e^{-j1.07\pi}$ and $0.95e^{-j0.88\pi}$. The corresponding spatial spectrum is shown in Figure 1. The hyper-parameter d^2 was set to

Scatterers	DOA (deg)	TOA (bins)	Amplitude (dB)
S1	65°	8	10
S2	20°	2	10

TABLE 5.1

Parameters of the multipath environment for simulated data

the value 10, according to (41). The hyper-parameter Λ was set to the true number of multipath components.

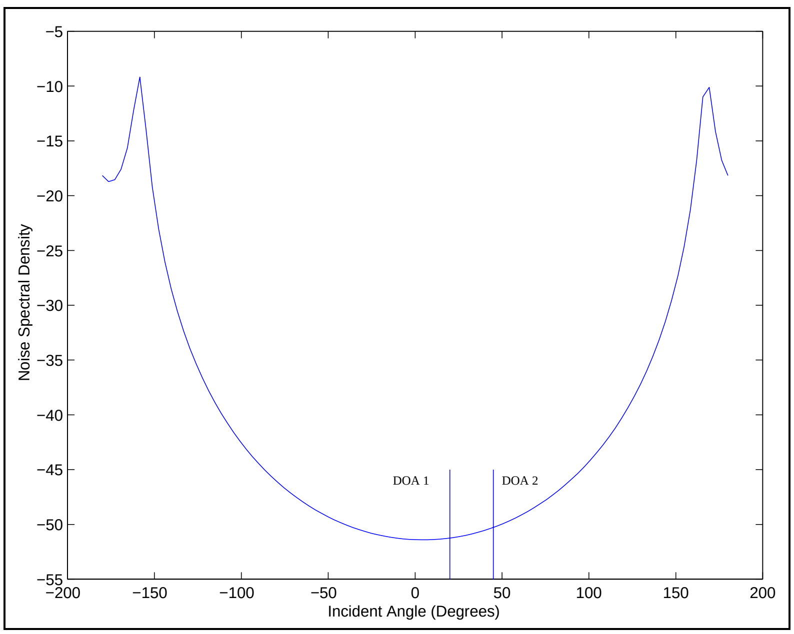

Fig. 1. Noise Spectrum

The Reversible Jump MCMC scheme goes through 1500 iterations after a burn-in period of 300 iterations. The results, as found by the algorithm, are summarized in Figure 2. It is clear from the histograms that the DOA and TOA estimates concentrate around their true values. The posterior probability of the number of scatterers $\hat{k}$ being equal to the true value of two was evaluated at 75% (due to a high noise level), as summarized in Table 5.2. The bottom line, "measurements", describes the detection performance using real measured data, as presented in section VI. Clearly, the algorithm correctly identified the parameters of the simulated multipath scenario. Further simulation results for the DOA case only are presented in [14].

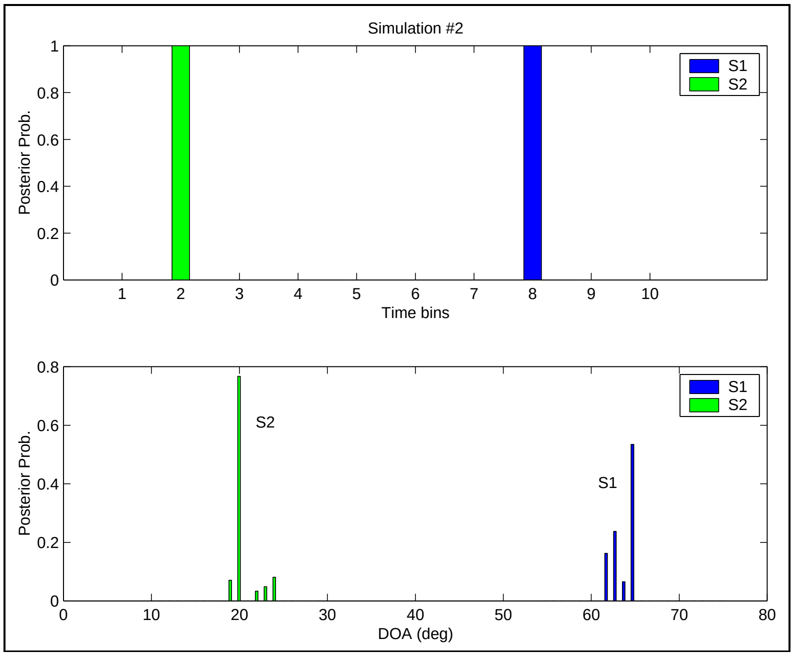

Fig. 2. Simulations: Histogram of the TOA (top); Histogram of DOA (bottom).

$p(\hat{k}=i)$ %	1	2	3
Simulations	0	75	25
Measurements	1	98	1

TABLE 5.2

Posterior estimate of the number of paths using MCMC, for both simulated data and real data.

A. Performance of the method

The TOA parameter is discrete; therefore the estimate may be considered fixed for reasonable values of SNR. Thus the joint TOA/DOA estimation performance is approximately determined by examining the DOA estimation performance alone.

Thus, in Figure 3, we show the variances of the DOA estimates vs. SNR for the proposed algorithm. The variance of the DOA estimates was determined on the basis of 50 independent trials at each value of SNR. The other parameter values are given as before in Table 5.1. Specifically, the value of d^2 was held at the value 10, which was verified to satisfy (41) over the entire range of SNR values considered. The results are compared with the Cramér-Rao lower bound, evaluated from the results in [15][16].

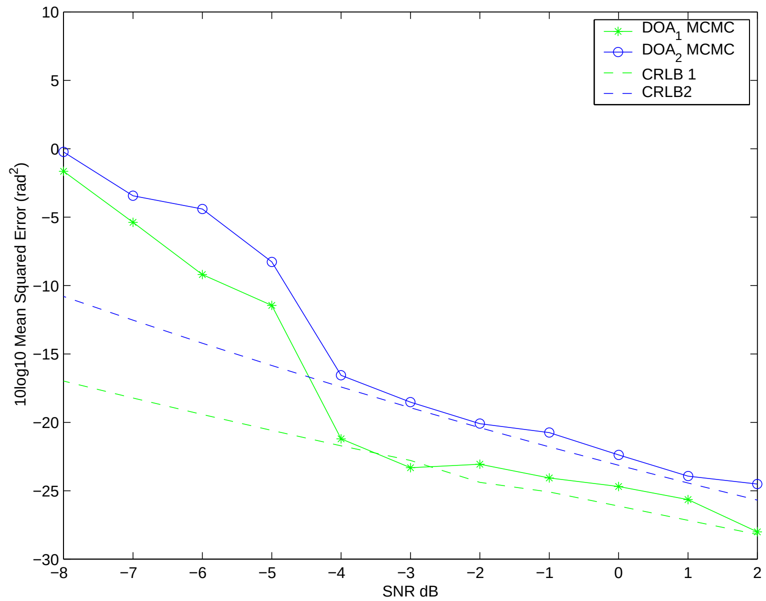

Fig. 3. Coloured noise with unknown covariance matrix: Mean Squared Error of the DOA estimates compared with the CRLB.

With reference to Figure 3, it may be observed that the performance of the MCMC method almost achieves the CRLB. The slight degradation may be caused by neglecting the off-diagonal terms of the covariance matrix in the development of the posterior distribution (12).

We note from Figure 3 that the performance of the signal associated with DOA_2 is degraded over that for DOA_1. This is because, with reference to Figure 1, the signal with DOA_2 receives a higher noise level than that for DOA_1. For this reason, Figure 3 shows variances for both DOA values.

Further, in Figure 4, we show the probability of detection of the correct number of sources vs. N, for the same simulation scenario as described previously, with the SNR held constant at 5dB. The value of d^2 was again set at the value 10. It is clear from this figure that the probability of correct detection approaches unity with increasing N, thus verifying the consistency of the proposed method under condition (41).

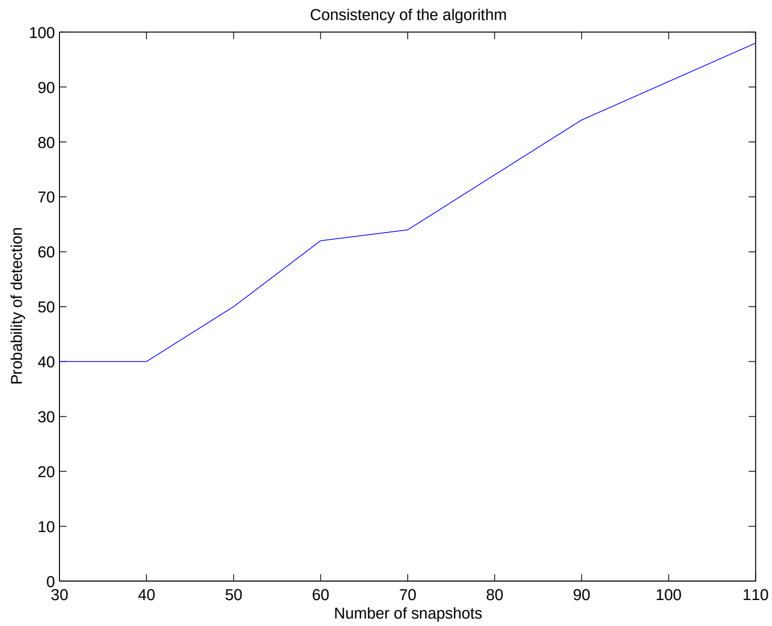

Fig. 4. Probability of detection vs. number of observations.

VI. Application to real-life problem

In this section, we apply our proposed scheme for the coloured noise case to real-life outdoor propagation measurements with a typical data set collected on McMaster University campus.

A. The measurement scenario

The complex channel impulse response, in time and space, is measured directly in the time domain by transmitting a wide-band spread-spectrum signal and correlating the received signal with the known transmitted sequence, at each element of the receiving array.

The receiving base station is a circular antenna array made of 8 mono-pole antennae. The transmitted signal is a 255 chip pseudonoise (PN) sequence at 5 MHz. The received signal of each element is I-Q demodulated, converted to baseband, sampled at 10 MHz, and then stored for further processing. One time bin ΔT therefore corresponds to 100ns. The measurements were conducted on the McMaster University campus, with the receiving base station at different locations and different heights in a pico-cell scenario that offered rich multipath characteristics with severe fading. An initial calibration is based on measurements with the antenna array inside an anechoic chamber and illuminated from 64 different angles. This calibration data is necessary to extract phase information from the data.

B. Processing

For the purpose of demonstrating the proposed algorithm, $N = 20$ received data symbols (or snapshots) are used, measured from a position where the multipath characteristics have been geographically observed to be 2 rays incident approximatively from angles of arrival of 30° and 135°, as shown in Figure 5.

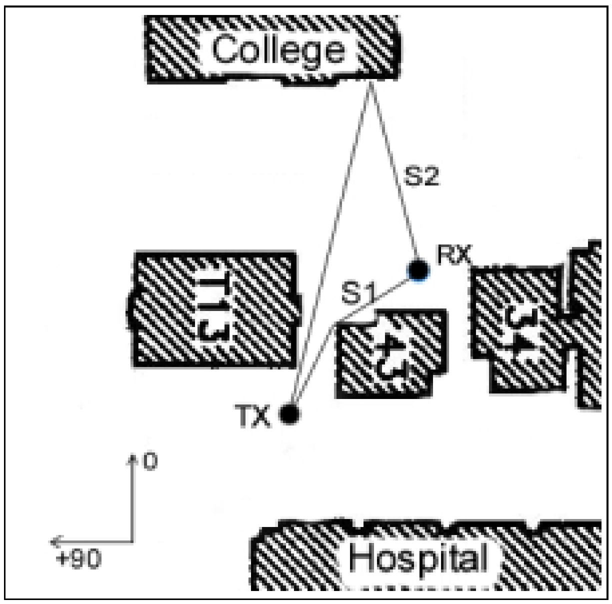

Fig. 5. Map describing the geometry of the setup.

Using traditional beamforming techniques, one can easily obtain an angular power spectrum. The temporal impulse response is directly obtained from the output of the correlator. Both results are presented in Figure 6, where the estimates found by the proposed algorithm are indicated. As the figure shows, it is not easy to first determine the number of multipath components; it can easily be over-estimated. Furthermore, there is ambiguity in which DOA estimate corresponds with which TOA estimate. With the proposed algorithm, since the optimization is done jointly for the two sets of parameters, the ambiguity is resolved.

The results from application of the proposed algorithm are now presented. Figure 7

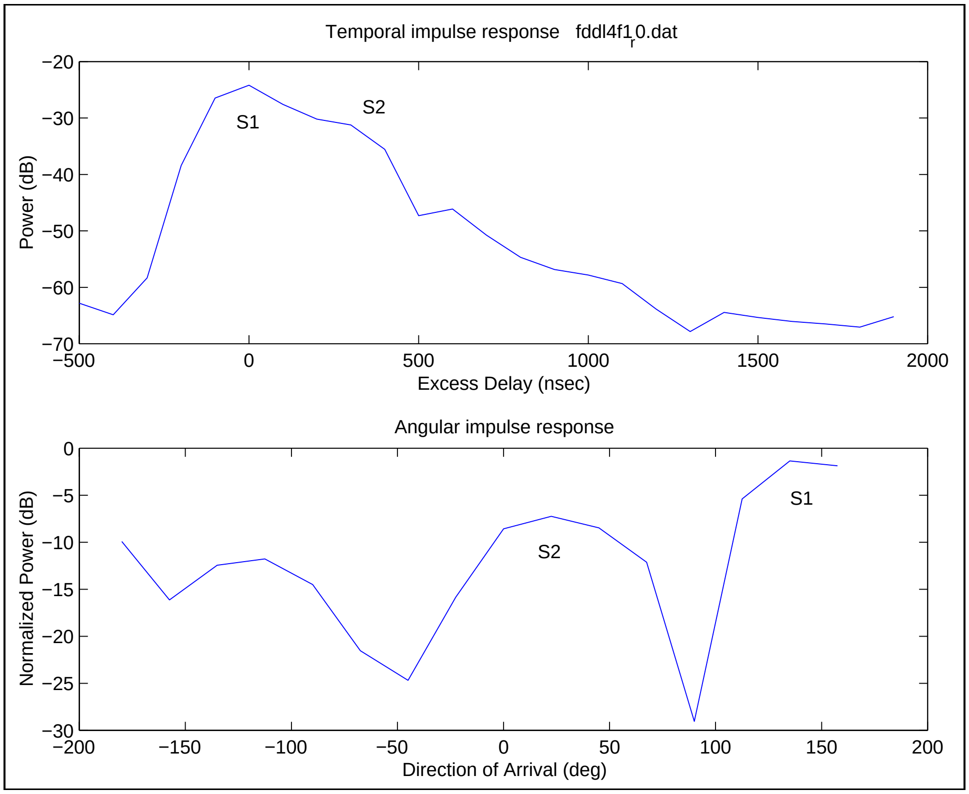

Fig. 6. Channel Impulse Response: Temporal (top); Angular (bottom).

shows typical results for 1000 iterations of the Reversible Jump MCMC Sampler. The initialization of the algorithm was totally random, with no prior information being used. It is clear that the algorithm identifies the two major multipath components, in time and in direction. These estimates are presented in Table 6.3. From the geometry of the system, the first multipath ray goes around the corner of building 43 (or over the top), while the second ray impinges from another direction, after bouncing off the Teachers' College building across the street. The receiving array is located approximately mid-way between the College and the transmitter. The extra propagation time is thus simply the path from the transmitter to the College, which was measured to be approximately 100m which corresponds to a time delay of 330ns. Within the resolution of the 100ns time bins, the excess delay measurement of S2 obtained from geographical truthing is therefore 300ns.

It is seen the parameters extracted by the algorithm are in accordance with the physical characteristics of the scenario, as described on the map. The performance of the detection

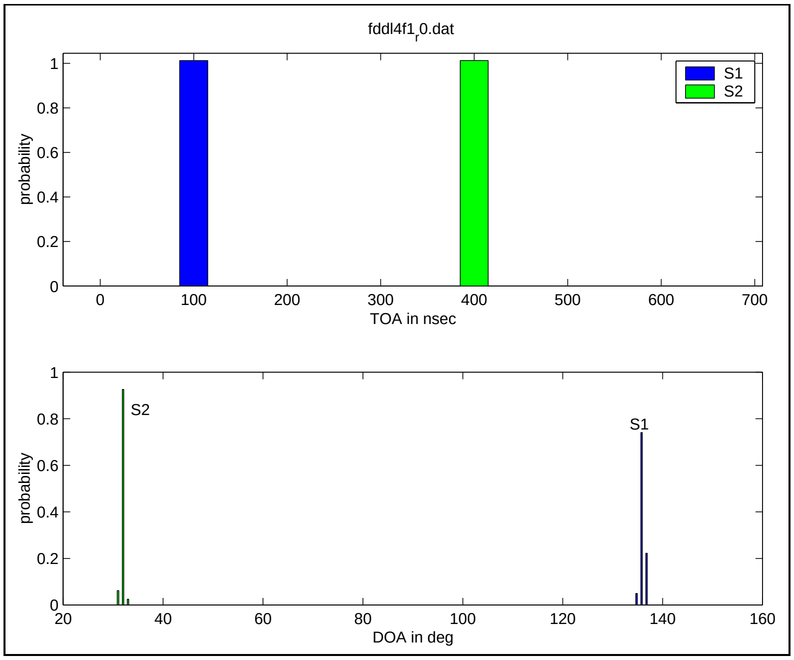

Fig. 7. Measurements: Histogram of the TOA (top); Histogram of the DOA (bottom).

Sources	DOA	TOA (Excess delay)
S1	136°	0 ns
S2	32°	300 ns

TABLE 6.3

Estimate of the multipath using MCMC

procedure for the real measurement case is shown in the bottom line of Table 5.2.

VII. Conclusion

In this paper, a new and innovative approach to channel characterization in unknown coloured noise with arbitrary covariance is presented in the Bayesian framework. We used a Markov Chain Monte Carlo method to perform the joint estimation of the model order, times of arrival and direction of arrival characterizing the multipath channel. The nuisance parameters (unknown noise variance and amplitudes) are integrated out analytically. This approach permits computation of the downlink beamforming weights used in FDD communication systems.

Simulation results support the effectiveness of the method, and demonstrate reliable detection of the number of sources and estimation of their directions and times of arrival in coloured noise with a single array. The performance is reinforced with the successful application to real data.

References

[1] M. C. Vanderveen, C. B. Papadias, and A. Paulraj, "Joint angle and delay estimation (JADE) for multipath signals arriving at an antenna array," *IEEE Communications Letters*, vol. 1, pp. 12–14, Jan. 1997.

[2] M. Wax and A. Leshem, "Joint estimation of time delays and directions of arrival of multiple reflections of a known signal," *IEEE Transactions on Signal Processing*, vol. 45, pp. 2477–2484, Oct. 1997.

[3] M. Wax, "Detection and localization of multiple sources in noise with unknown covariance," *IEEE Transactions on Signal Processing*, vol. 40, pp. 245–249, Jan. 1992.

[4] J.-R. Larocque and J. P. Reilly, "Reversible jump MCMC for joint detection and estimation of directions of arrival in coloured noise," *IEEE Transactions on Signal Processing*, 2000. Submitted for publication, also available at http://www.ece.mcmaster.ca/~reilly.

[5] J. J. Blanz, A. Papathanassiou, M. Haardt, I. Furio, and P. W. Baier, "Smart antennas for combined DOA and joint channel estimation in time-slotted CDMA mobile radio systems with joint detection," *IEEE Transactions on Vehicular Technology*, vol. 49, pp. 293–306, Mar. 2000.

[6] A. Logothetis and C. Carlemalm, "SAGE algorithms for multipath detection and parameter estimation in asynchronous CDMA systems," *IEEE Transactions on Signal Processing*, vol. 48, pp. 3162–3174, Nov. 2000.

[7] K. Pahlavan and A. H. Levesque, *Wireless Information Networks*. New York: John Wiley and Sons, 1995.

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

[9] K.-M. Wong, J. P. Reilly, Q. Wu, and S. Qiao, "Estimation of the directions of arrival of signals in unknown correlated noise, Part 1: The MAP approach and its implementation," *IEEE Transactions on Signal Processing*, vol. 40, pp. 2007–2017, August 1992.

[10] C.-M. Cho and P. M. Djuric, "Bayesian detection and estimation of cisoids in colored noise," *IEEE Transactions on Signal Processing*, vol. 43, pp. 2943–2951, Dec. 1995.

[11] 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, pp. 2667–2676, Oct. 1999.

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

[13] P. M. Djuric, "A model selection rule for sinusoids in white Gaussian noise," *IEEE Transactions on Signal Processing*, vol. 44, pp. 1744–1751, July 1996.

[14] J.-R. Larocque and J. P. Reilly, "Reversible jump MCMC for joint detection and estimation of directions of arrival in coloured noise," in *Proceedings of the IMA International Conference on Mathematics in Signal Processing*, (University of Warwick, England), Dec. 2000.

[15] A. B. Gershman, P. Stoica, M. Pesavento, and E. Larsson, "Stochastic Cramér-Rao bound for direction estimation in unknown noise fields," *IEEE Transactions on Signal Processing*, 2001. To appear.

[16] H. Ye and R. DeGroat, "Maximum likelihood DOA estimation and asymptotic Cramér-Rao bounds for additive unknown colored noise," *IEEE Transactions on Signal Processing*, vol. 43, pp. 938–949, Apr. 1995.