Joint Channel and Doppler Offset Estimation in Dynamic Cooperative Relay Networks

Ido Nevat, Gareth W. Peters, Arnaud Doucet, and Jinhong Yuan, *Senior Member, IEEE*

***Abstract*—We develop a new and efficient algorithm to solve the problem of *joint channel and Doppler offset estimation* in time-varying cooperative wireless relay networks. We first formulate the problem as a Bayesian dynamic nonlinear state space model, then develop an algorithm, which is based on *particle adaptive marginal Markov chain Monte Carlo*, method to jointly estimate the time-varying channels and static Doppler offsets. We perform detailed complexity analysis of the proposed algorithm and show that it is very efficient and requires moderate computational complexity. In addition, we develop a new version of the recursive marginal Cramér–Rao lower bound and derive expressions for the achievable mean-square error. Simulation results demonstrate that the proposed algorithm outperforms the state-of-the-art algorithms and performs close to the Cramér–Rao lower bound.**

***Index Terms*—Particle filter, MCMC, channel estimation, Doppler offset estimation, cooperative relay networks.**

I. Introduction

RELAY based systems, first introduced by van der Meulen [1], have recently received considerable attention due to their potential in wireless applications. Relaying techniques provide spatial diversity, improve energy efficiency, and reduce the interference level of wireless channels [2], [3]. To utilize such systems, an accurate channel state information (CSI) is required at the destination. When the communicating terminals and relays are mobile, the wireless channels form a cascade of mobile-to-mobile channels, and change rapidly with time. These wireless channels can be accurately modeled as a dynamic time-varying system (State-Space model) [4]–[6]. In these practical scenarios, the problem is not only to accurately estimate the time varying channels but also the unknown Doppler offsets [5], [7], [8]. This problem is refereed to as joint *channel estimation and Doppler offset estimation and is the focus of this paper*.

A few partial solutions have been proposed in the literature and are based on various simplifying assumptions: a state-space model was proposed in [4] where a sub-optimal algorithm, based on linear estimator (i.e. Kalman filter) was developed. However, the authors assumed that the Doppler offset is known. They reported that the relay speed has a significant impact on the BER performance, and it is important to estimate the Doppler offset in cases where it is unknown *a-priori*. In [5], [8], the problem of joint channel estimation and Doppler offset estimation was considered. However, the authors assumed that the speed of the mobile terminals is known, allowing the formulation of the problem as a non-linear state-space model and solving via Sequential Monte Carlo (SMC) methods. The problem of data detection for OFDM systems with carrier frequency offset over unknown doubly selective channels was addressed in [9] where the authors utilized an Expectation-Maximization (EM) type approach. In [10], an algorithm for data detection utilizing superimposed training symbols was developed utilizing a linear minimum mean square error (LMMSE) criterion.

In contrast to [4], [5], [8], we solve the practical problem of *joint channel and Doppler offset estimation* for the case that the speed of the mobile terminals is unknown. This scenario is of importance since in practice the destination node has no knowledge of the speeds at which the source and the relays are moving. This means that the Doppler offsets are unknown to the destination and need to be jointly estimated along with the channels coefficients. We address this problem by developing a Bayesian state-space model which incorporates the unknown Doppler offsets.

In this paper we develop a new and efficient algorithm to perform joint channel estimation (tracking) and Doppler offset estimation. The algorithm we develop is based on advanced Monte Carlo techniques which combine the strength of non-linear filtering frameworks, such as SMC, with adaptive versions of Markov chain Monte Carlo (MCMC) techniques [11].

In addition, we develop a novel Bayesian model formulation, for which we study the properties and demonstrate efficiently how to make inference for such a state space model structure. In this regard, we perform detailed computational complexity and make a fair comparison between the algorithms we develop. We asses the performance of our algorithm by developing a lower bound on the achievable Mean Square Error (MSE) via the dynamic CRLB. We perform extensive simulations and show that the proposed algorithm performs close to the lower bound.

Notation: random variables are denoted by upper case letters and their realizations by lower case letters; bold face to denote vectors and non-bold for scalars; super script will be used to refer to the index for a particular relay in the network; sub-script will denote discrete time, where $h_{1:T}$ denotes $h_1, \ldots, h_T$; and in the sampling methodology combining MCMC and particle

Manuscript received July 29, 2013; revised February 2, 2014 and June 18, 2014; accepted September 21, 2014. Date of publication October 8, 2014; date of current version December 8, 2014. This paper was presented in part at the 2011 IEEE VTC Conference, San Francisco, CA, USA. The associate editor coordinating the review of this paper and approving it for publication was Z. Wang.

I. Nevat is with the Institute for Infocomm Research (I2R), Singapore 138632.

G. W. Peters is with the Department of Statistical Science, University College London, London WC1E 6BT, U.K., and also with the CSIRO Mathematical and Information Sciences, Clayton South, VIC 3169, Australia.

A. Doucet is with the Department of Statistics, Oxford OX1 3UB, U.K.

J. Yuan is with the School of Electrical Engineering, University of New South Wales, Sydney, NSW 2052, Australia.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2014.2361760

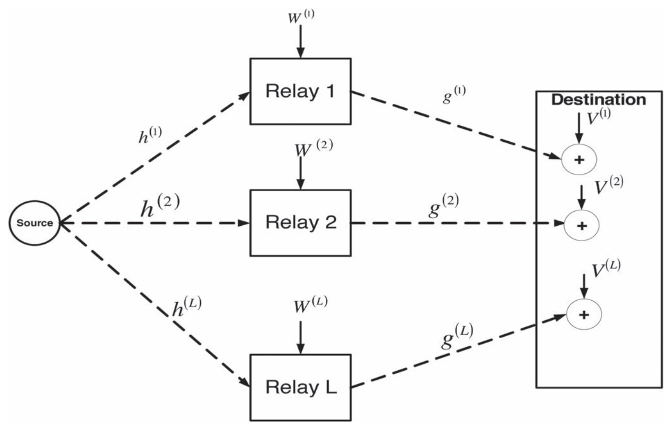


Fig. 1. System model of the relay network with a single source and L relays, communicating to the destination.

filtering we use the following notation $[\cdot](j,i)$ to denote the j-th state of the Markov chain for the i-th particle in the particle filter. In addition, we denote the proposed Markov chain state by $[\cdot]^*(j,i)$, and $\delta(\cdot)$ denotes the delta of Dirac.

II. System Description

We consider the case where one mobile station is transmitting to a Base Station (BS) via L relays, which may be mobile or stationary, see Fig. 1. We consider a training symbol based system [5], thus a data detection is not considered in this paper. We consider frequency-flat fading channels, and the extension to multi-path channels is straight forward, for example via the use of OFDM modulation.

A. Relay Network Model

i. Consider a wireless relay network with one mobile source node, transmitting symbols in frames of length T to a BS via L mobile relays.

ii. We consider a half duplex transmission model in which the data for a given frame are transmitted via a two step procedure. In the first step, the source node broadcasts a frame to all the relay nodes. In the second step, the relay nodes transmit the processed frame, termed the relay signals, to the destination node in **orthogonal fashion, i.e. non-interfering channels via time division or frequency division multiplex**, see for example [12], [13].

iii. The **time-varying channels** follow a Gauss-Markov model [14], [15]. The l-th relay channel is modeled as a two stage latent random process, where $H_n^{(l)}$ is the wireless channel between the source and the l-th relay, and $G_n^{(l)}$ is the wireless channel between the l-th relay and the destination. These channels can be expressed at time n as:

$$H_n^{(l)} = \alpha^{(l)} H_{n-1}^{(l)} + \sqrt{1-\left(\alpha^{(l)}\right)^2}\Upsilon_n^{(l)}$$

$$G_n^{(l)} = \beta^{(l)} G_{n-1}^{(l)} + \sqrt{1-\left(\beta^{(l)}\right)^2}\Omega_n^{(l)},$$

where $\Upsilon_n^{(l)} \sim \mathcal{CN}(0,1)$, $\Omega_n^{(l)} \sim \mathcal{CN}(0,1)$.

iv. The velocities of both the user and the relays are random unknown quantities and constant over a frame and follow a uniform distribution $V_M \sim U[0, v_{\max}]$ and $V_R^{(l)} \sim U[0, v_{\max}]$, respectively, where $v_{\max}$ is a practical upper bound.

v. The channel coefficients $\alpha^{(l)}$ and $\beta^{(l)}$ are modeled according to Jakes' model as [6]

$$\begin{aligned}\alpha^{(l)} &= J_0\left(2\pi\frac{V_M f_c}{c}T_s\right) J_0\left(2\pi\frac{V_R^{(l)} f_c}{c}T_s\right)\\ &= J_0\left(2\pi F_{M\to R}^{(l)} T_s\right) J_0\left(2\pi F_{R\to D}^{(l)} T_s\right),\\ \beta^{(l)} &= J_0\left(2\pi\frac{V_M f_c}{c}T_s\right) = J_0\left(2\pi F_{M\to R}^{(l)} T_s\right),\end{aligned} \tag{1}$$

where J_0 is the zeroth-order Bessel function of the first kind, f_c is the carrier frequency, c is the speed of light and T_s is the sampling period (e.g. symbol duration, frame length etc.) [16], [17].

vi. The received signal at the l-th relay is a random variable given by

$$R_n^{(l)} = s_n H_n^{(l)} + W_n^{(l)},\ l \in \{1,\dots,L\},$$

where at time n, $H_n^{(l)}$ is the channel coefficient between the transmitter and the l-th relay, s_n is the transmitted pilot symbol and $W_n^{(l)}$ is the noise realization associated with the relay receiver.

vii. The received signals at the destination is given by

$$Y_n^{(l)} = f^{(l)}\left(R_n^{(l)}\right) G_n^{(l)} + V_n^{(l)},\ l \in \{1,\dots,L\},$$

where at time n, $G_n^{(l)}$ is the channel coefficient between the l-th relay and the receiver. The model we develop is general enough to allow for many different possible relay functions, in the form of $f^{(l)}(R^{(l)}) : \mathbb{R} \mapsto \mathbb{R}$, for any continuous d-differentiable functions $f^{(l)}(R^{(l)}) \in \mathbb{C}^d[\mathbb{R}]$ on the real line. As an example consider a relay function given by the popular Amplify-and-Forward with constant gain [18], [19], with function

$$f^{(l)}\left(R^{(l)}\right) = \sqrt{\frac{1}{\sigma_h^2 + \sigma_g^2 + \frac{\sigma_v^2}{\mathbb{E}[|s_n|^2]}}} R_n^{(l)}$$

where $\mathbb{E}[|s_n|^2]$ is the average symbols power.

viii. All received signals are corrupted by i.i.d. zero-mean additive white complex Gaussian noise (AWGN). At the l-th relay at the n-th transmitted symbol is denoted by $W_n^{(l)} \sim \mathcal{CN}(0,\sigma_w^2)$. Then at the receiver this is denoted by $V_n^{(l)} \sim \mathcal{CN}(0,\sigma_v^2)$.

Based on the aforementioned assumptions, the following state-space model expressed as

$$\begin{cases} H_n^{(l)} = \alpha^{(l)} H_{n-1}^{(l)} + \sqrt{1-\left(\alpha^{(l)}\right)^2}\Upsilon_n^{(l)},\\ G_n^{(l)} = \beta^{(l)} G_{n-1}^{(l)} + \sqrt{1-\left(\beta^{(l)}\right)^2}\Omega_n^{(l)},\\ Y_n^{(l)} = f^{(l)}\left(s_n H_n^{(l)} + W_n^{(l)}\right) G_n^{(l)} + V_n^{(l)}.\end{cases}$$

B. Channel and Doppler Offset Estimation Formulation

To complete the problem formulation, we need to derive the distribution of transformed random variables $J_0(2\pi F^{(l)}_{M\to R}T_s)$ and $J_0(2\pi F^{(l)}_{R\to D}T_s)$ in (1). This is presented in the following Lemma, where we set $v_{\max} = \frac{c}{\pi f_c T_s}$, to accommodate for a wide range of velocities, carrier frequencies and symbol rates.

Lemma 1: Under a uniform prior distribution of the mobile velocity, $V_M \sim U[0, v_{\max}]$, its Doppler offset can be approximated to follow a Beta distribution as follows:

$$J_0\left(2\pi \frac{V_M f_c}{c} T_s\right) \sim Beta(1, 1/2).$$

Proof: We expand $J_0\left(2\pi \frac{V_M f_c}{c} T_s\right)$ via Taylor Series expansion around 0:

$$\begin{aligned} J_0\left(\frac{2\pi V_M f_c T_s}{c}\right) &= \sum_{m=0}^{\infty} \frac{(-1)^m}{m!\Gamma(m+1)}\left(\frac{1}{2}\frac{2\pi V_M f_c T_s}{c}\right)^{2m} \\ &\approx \sum_{m=0}^{K} \frac{(-1)^m}{m!\Gamma(m+1)}\left(\frac{\pi V_M f_c T_s}{c}\right)^{2m} \\ &\overset{(K=1)}{\approx} 1 - \left(\frac{\pi V_M f_c T_s}{c}\right)^2. \end{aligned}$$

Next, we note that $\left(\frac{\pi V_M f_c T_s}{c}\right) \sim U[0, 1]$, and therefore, by utilizing the power *function distribution*, we obtain that $\left(\frac{\pi V_M f_c T_s}{c}\right)^2 \sim \mathrm{Beta}(1/2, 1)$. Then, by using the *mirror-image symmetry* of the Beta distribution, we obtain the final result. ■

We can now express the posterior distribution of the wireless channels and the Doppler offsets as, presented in (2), shown at the bottom of the page, where

$$\alpha^{(1:L)}, \beta^{(1:L)}, g^{(1:L)}_{1:T}, h^{(1:L)}_{1:T}, y^{(l:L)}_{1:T} \triangleq \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{g}_{1:T}, \mathbf{h}_{1:T}, \mathbf{y}_{1:T}.$$

Next, based on the marginal posterior in (2), we define the quantities of interest, namely the MAP and MMSE estimators:

$$\begin{aligned} &\left(\widehat{\boldsymbol{\alpha}}^{\mathrm{MAP}}, \widehat{\boldsymbol{\beta}}^{\mathrm{MAP}}, \widehat{\mathbf{h}}^{\mathrm{MAP}}_{1:T}, \widehat{\mathbf{g}}^{\mathrm{MAP}}_{1:T}\right) \\ &\quad = \underset{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{g}_{1:T},\mathbf{h}_{1:T}}{\arg\max}\ p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{g}_{1:T}, \mathbf{h}_{1:T}|\mathbf{y}_{1:T}). \\ &\left(\widehat{\boldsymbol{\alpha}}^{\mathrm{MMSE}}, \widehat{\boldsymbol{\beta}}^{\mathrm{MMSE}}, \widehat{\mathbf{h}}^{\mathrm{MMSE}}_{1:T}, \widehat{\mathbf{g}}^{\mathrm{MMSE}}_{1:T}\right) = \mathbb{E}[\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{g}_{1:T}, \mathbf{h}_{1:T}|\mathbf{y}_{1:T}] \\ &\quad = \int p(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{g}_{1:T},\mathbf{h}_{1:T}|\mathbf{y}_{1:T})\, \boldsymbol{\alpha}\, \boldsymbol{\beta}\, \mathbf{g}_{1:T}\, \mathbf{h}_{1:T}\, d\boldsymbol{\alpha}\, d\boldsymbol{\beta}\, d\mathbf{h}_{1:T}\, d\mathbf{g}_{1:T}. \end{aligned} \tag{3}$$

The posterior model in (2) is very high dimensional with $L(3T+2)$ parameters and highly nonlinear. We now show that the model admits a block factorization structure, which means that the particle filters may be run as two filters separately i.e. independently and therefore can exploit a parallel implementation design.

Lemma 2: The posterior distribution in (2) factorizes according to the following independence structure

$$p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{g}_{1:T}, \mathbf{h}_{1:T}|\mathbf{y}_{1:T}) = \prod_{l=1}^{L} p\left(\alpha^{(l)}, \beta^{(l)}, g^{(l)}_{1:T}, h^{(l)}_{1:T}|y^{(l)}_{1:T}\right),$$

with respect to the number of parallel relay transmission paths. We will take advantage of Lemma 1 in the design of the PMCMC algorithm to reduce the variance of the incremental importance weights, leading to an increased efficiency of the algorithm we develop. Note, the incremental importance weights are the importance sampling weights calculated at each iteration or stage of the Sequential Monte Carlo filter.

C. Deriving the Likelihood Function Via Augmented Bayesian Posterior Model

In order to obtain (3), we need to derive the likelihood function in (2), $p\left(y^{(l)}_n|\alpha^{(l)}, \beta^{(l)}, g^{(l)}_n, h^{(l)}_n\right)$. To achieve that we

$$\begin{aligned} &p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{g}_{1:T}, \mathbf{h}_{1:T}|\mathbf{y}_{1:T}) \\ &\propto p(\mathbf{y}_{1:T}|\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{g}_{1:T}, \mathbf{h}_{1:T}) p(\mathbf{g}_{1:T}, \mathbf{h}_{1:T}|\boldsymbol{\alpha}, \boldsymbol{\beta}) p(\boldsymbol{\alpha}, \boldsymbol{\beta}) \\ &= \prod_{l=1}^{L}\left[\prod_{n=1}^{T}\int p\left(y^{(l)}_n|\alpha^{(l)}, \beta^{(l)}, g^{(l)}_n, h^{(l)}_n, w^{(l)}_n\right) p\left(w^{(l)}_n\right) \mathrm{d}w^{(l)}_n p\left(h^{(l)}_n|h^{(l)}_{n-1}\right) p\left(g^{(l)}_n|g^{(l)}_{n-1}\right)\right] \\ &\times \prod_{l=1}^{L} p\left(g^{(l)}_1\right) p\left(h^{(l)}_1\right) p\left(\alpha^{(l)}\right) p\left(\beta^{(l)}\right) \\ &= \prod_{l=1}^{L}\left[\prod_{n=1}^{T} \frac{1}{\sqrt{2\pi\sigma_v^2}} \int \exp\left(-\frac{\left|y^{(l)}_n - f\left(s_n h^{(l)}_n + w^{(l)}_n\right) g^{(l)}_n\right|^2}{2\sigma_v^2}\right) \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left\{\frac{-\left|w^{(l)}_n\right|^2}{2\sigma_w^2}\right\} \mathrm{d}w^{(l)}_n\right. \\ &\left.\times \frac{1}{\sqrt{2\pi\left(1-\left(\beta^{(l)}\right)^2\right)}} \exp\left\{-\frac{\left|g^{(l)}_n - \beta^{(l)} g^{(l)}_{n-1}\right|^2}{2\left(1-\left(\beta^{(l)}\right)^2\right)}\right\} \frac{1}{\sqrt{2\pi\left(1-\left(\alpha^{(l)}\right)^2\right)}} \exp\left\{-\frac{\left|h^{(l)}_n - \alpha^{(l)} h^{(l)}_{n-1}\right|^2}{2\left(1-\left(\alpha^{(l)}\right)^2\right)}\right\}\right] \\ &\times \frac{1}{\sqrt{2\pi}}\prod_{l=1}^{L}\exp\left\{-\frac{1}{2}\left|g^{(l)}_1\right|^2\right\} \frac{1}{\sqrt{2\pi}}\prod_{l=1}^{L}\exp\left\{-\frac{1}{2}\left|h^{(l)}_1\right|^2\right\} \times \prod_{l=1}^{L}\frac{1}{\mathrm{B}(a,b)}\left(\alpha^{(l)}\right)^{a-1}\left(1-\alpha^{(l)}\right)^{b-1} \prod_{l=1}^{L}\frac{1}{\mathrm{B}(c,d)}\left(\beta^{(l)}\right)^{c-1}\left(1-\beta^{(l)}\right)^{d-1} \end{aligned} \tag{2}$$

write the marginal distribution at the l-th relay as

$$p_{R_n^{(l)}}\left(r|s_n, g_n^{(l)}, h_n^{(l)}\right) = p\left(s_n h_n^{(l)} + w_n^{(l)}|s_n, h_n^{(l)}, g_n^{(l)}\right) = \mathcal{CN}\left(s_n h_n^{(l)}, \sigma_w^2\right).$$

Then, finding the distribution of the random variable after the relay function is applied i.e. the distribution of $\widetilde{f}_n^{(l)} \triangleq f\left(r_n^{(l)}\right) G_n^{(l)}$ given $\mathbf{s}_n, h_n^{(l)}, g_n^{(l)}$, involves the following change of variable formula

$$p\left(\widetilde{f}_n^{(l)}|s_n, h_n^{(l)}, g_n^{(l)}\right) = p_{R_n^{(l)}}\left(\left(\widetilde{f}_n^{(l)}\right)^{-1}|s_n, h_n^{(l)}, g_n^{(l)}\right)\left|\frac{\partial \widetilde{f}^{(l)}}{\partial r_n^{(l)}}\right|^{-1},$$

which can not always be written down analytically for an arbitrary relay function. The second complication is that even if $\widetilde{f}_n^{(l)} = f\left(r_n^{(l)}\right) G_n^{(l)}$ is known, one must then solve the following convolution to obtain the likelihood:

$$p\left(y_n^{(l)}|s_n, g_n^{(l)}, h_n^{(l)}\right) = p\left(\widetilde{f}_n^{(l)}|s_n, g_n^{(l)}, h_n^{(l)}\right) \otimes p_{V_n^{(l)}} = \int_{-\infty}^{\infty} p\left(\widetilde{f}\left(z|s_n, g_n^{(l)}, h_n^{(l)}\right)\right) p_{V_n^{(l)}}\left(y_n^{(l)} - z\right) dz,$$

where $\otimes$ denotes the convolution operation. Typically this will be intractable to evaluate pointwise.

We solve this problem via augmented Bayesian posterior $p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{g}_{1:T}, \mathbf{h}_{1:T}, \mathbf{w}_{1:T}|\mathbf{y}_{1:T})$, containing auxiliary variables $\mathbf{W}_{1:T}$ which we marginalize out numerically in our sampling algorithm to obtain the posterior corresponding to (2). Under the augmented model, the likelihood function can now be expressed as

$$y_n^{(l)}|s_n, g_n^{(l)}, h_n^{(l)}, w_n^{(l)} \sim \mathcal{CN}\left(f^{(l)}\left(s_n h_n^{(l)} + w_n^{(l)}\right) g_n^{(l)}, \sigma_v^2\right).$$

III. Joint Channel and Doppler Estimation via Adaptive Particle MCMC

In this section we develop a novel algorithm which is based on adaptive Particle MCMC (PMCMC) methodology [20]. This will provide a very efficient algorithm to find the MAP and MMSE estimates in (3). In particular, we develop the *Marginal Metropolis-Hastings within Rao–Blackwellized particle filter* algorithm. Our algorithm provides very efficient sampling mechanism from the full conditional posterior in (2) by operating on the factorisation of the joint space $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{g}_{1:T}, \mathbf{h}_{1:T})$. The key advantage of our algorithm is that it allows us to jointly update the entire set of posterior parameters at each iteration and only requires calculation of the marginal acceptance probability in the Metropolis-Hastings (MH) algorithm. We achieve this by embedding a particle filter estimate of the optimal proposal distribution for the latent process into the MCMC algorithm. This allows the Markov chain to mix efficiently in the high dimensional posterior parameter space due to the particle filter approximation of the optimal proposal distribution in the MCMC algorithm, thereby allowing high-dimensional parameter block updates even in the presence of strong posterior parameter dependence. To develop the algorithm we begin by writing the MH acceptance probability for the augmented model $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T})$:

$$\begin{aligned} &A\left([\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}]^*; [\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}](j)\right) \\ &= \min\left(1, \frac{p\left([\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta}]^*|\mathbf{y}_{1:T}\right)}{p\left([\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta}](j)|\mathbf{y}_{1:T}\right)} \right. \\ &\left. \times \frac{q\left([\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta}](j); [\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta}]^*\right)}{q\left([\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta}]^*; [\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right)}\right) \end{aligned} \quad (4)$$

Next we specify the Markov transition kernel, given by $q([\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta}](j); [\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta}]^*)$. We propose a particular choice of proposal that will provide a significant dimension reduction in evaluation of the acceptance probability. Our choice to move from a state at iteration j to a new state at iteration $(j+1)$ is split into two components:

$$\begin{aligned} &q\left([\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta}]^*; [\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right) \\ &= p\left([\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}]^*|[\boldsymbol{\alpha}, \boldsymbol{\beta}]^*, \mathbf{y}_{1:T}\right) q\left([\boldsymbol{\alpha}, \boldsymbol{\beta}]^*|[\boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right). \end{aligned} \quad (5)$$

The first component involves the sampling of a trajectory for $\mathbf{g}_{1:T}, \mathbf{h}_{1:T}, \mathbf{w}_{1:T}$, while the second component involves a Markov transition kernel to sample the static parameters $\boldsymbol{\alpha}, \boldsymbol{\beta}$. Plugging (5) into the acceptance probability in (4) results in the dimension reduction in the acceptance probability over the latent path space. This solution involves marginalization over the path space $\mathbf{g}_{1:T}, \mathbf{h}_{1:T}, \mathbf{w}_{1:T}$ to obtain the marginal likelihood required to evaluate the dimension reduced marginal MH acceptance probability. In other words, the acceptance probability is evaluated only on the parameter space and not on the full parameter space and the latent space.

To utilize this MH algorithm, we need to solve the following two problems:

Problem I: The marginal likelihood $p(\mathbf{y}_{1:T}|\boldsymbol{\alpha}, \boldsymbol{\beta})$ which is used in the evaluation of (4) can not be obtained analytically. This is due to the fact that marginalization of the joint likelihood over the path space involves the following integration

$$\begin{aligned} p(\mathbf{y}_{1:T}|\boldsymbol{\alpha}, \boldsymbol{\beta}) &= \prod_{n=1}^{T} p(\mathbf{y}_n|\mathbf{y}_{1:n-1}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \\ &= \int \left[\prod_{n=1}^{T} p(\mathbf{y}_n|\mathbf{h}_n, \mathbf{g}_n, \mathbf{w}_n, \boldsymbol{\alpha}, \boldsymbol{\beta}) \right. \\ &\left. \times p(\mathbf{h}_n, \mathbf{g}_n, \mathbf{w}_n|\mathbf{y}_{1:n-1}, \boldsymbol{\alpha}, \boldsymbol{\beta})\right] d\mathbf{h}_{1:T} d\mathbf{g}_{1:T} d\mathbf{w}_{1:T} \end{aligned}$$

which can not be performed analytically.

Problem II: We need to evaluate and sample from the distribution of the latent path space $p(\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}|\mathbf{y}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta})$. This can be achieved by constructing the sequence of distributions recursively, over the path space given by $\{p(\mathbf{h}_1, \mathbf{g}_1, \mathbf{w}_1|\mathbf{y}_1, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)), p(\mathbf{h}_{1:2}, \mathbf{g}_{1:2}, \mathbf{w}_{1:2}|\mathbf{y}_{1:2}, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)), \ldots, p(\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}|\mathbf{y}_{1:T}, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j))\}$ via a two step filter recursion involving the Chapman-Kolmogorov equation:

Stage I (Prediction):

$$
\begin{aligned}
&p\left(\mathbf{h}_n, \mathbf{g}_n, \mathbf{w}_n | \mathbf{y}_{1:n-1}, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right) \\
&= \int p\left(\mathbf{h}_n, \mathbf{g}_n, \mathbf{w}_n | \mathbf{h}_{n-1}, \mathbf{g}_{n-1}, \mathbf{w}_{n-1}, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right) \\
&\times p(\mathbf{h}_{n-1}, \mathbf{g}_{n-1}, \mathbf{w}_{n-1} | \mathbf{y}_{1:n-1}, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)) d\mathbf{h}_{n-1} d\mathbf{g}_{n-1} d\mathbf{w}_{n-1}.
\end{aligned}
$$

Stage II (Update):

$$
\begin{aligned}
&p\left(\mathbf{h}_n, \mathbf{g}_n, \mathbf{w}_n | \mathbf{y}_{1:n}, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right) \\
&= \frac{p(\mathbf{y}_n | \mathbf{h}_n, \mathbf{g}_n, \mathbf{w}_n, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j))\, p\left(\mathbf{h}_n, \mathbf{g}_n, \mathbf{w}_n | \mathbf{y}_{1:n-1}, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right)}{p\left(\mathbf{y}_n | \mathbf{y}_{1:n-1}, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right)}
\end{aligned}
$$

which can not be performed analytically.

We now develop the solutions to **Problem I** and **Problem II**.

A. *Solution for* ***Problem I*** *and* ***Problem II***

We solve **Problem I** and **Problem II** by decomposing the Markov transition kernel according to (4). We approximate the first component in (4) via a Rao–Blackwellized particle filter and the second component in (4) is constructed as an adaptive MH Markov transition kernel. These solutions detailed below provide the particle filter based estimates which are utilized in the PMCMC algorithm, to approximate the acceptance probability of the ideal choice, given in (4).

$$
\begin{aligned}
&\widehat{p}\left(\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T} | \mathbf{y}_{1:T}, [\boldsymbol{\alpha}, \boldsymbol{\beta}]^*(j+1)\right) \\
&\quad = \sum_{i=1}^{N} [\Xi_{1:T}](j, i) \delta_{[\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}](j,i)}(\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T})
\end{aligned}
$$

$$
\widehat{p}\left(\mathbf{y}_{1:T} | [\boldsymbol{\alpha}, \boldsymbol{\beta}]^*(j+1)\right) = \prod_{n=1}^{T} \left(\frac{1}{N} \sum_{i=1}^{N} [\xi_n](j, i) \right), \tag{6}
$$

where $[\Xi_{1:T}](j, i)$ and $[\xi_n](j, i)$ are the normalized importance weight on the path space and the incremental importance weight at time n, respectively, for the i-th particle at the j-th iteration.

We now present how to construct the Markov transition kernel in (5) for the PMCMC algorithm. We first present the proposal for the latent states (Rao–Blackwellized particle filter) followed by the proposal for the static parameters (adaptive MCMC).

Obtaining the approximation given in (6) involves a Rao–Blackwellized particle filter detailed below.

1) Particle Filter for $p(\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T} | \mathbf{y}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta})$*:* We develop a Rao–Blackwellized version of SMC to recursively approximate $p(\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T} | \mathbf{y}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta})$. The proposal kernel can be decomposed as

$$
\begin{aligned}
&p\left(\mathbf{g}_{1:T}, \mathbf{h}_{1:T}, \mathbf{w}_{1:T} | \mathbf{y}_{1:T}, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right) \\
&\quad = \underbrace{p\left(\mathbf{g}_{1:T} | \mathbf{h}_{1:T}, \mathbf{w}_{1:T}, \mathbf{y}_{1:T}, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right)}_{\text{Kalman filter}} \\
&\quad \times \underbrace{p\left(\mathbf{h}_{1:T}, \mathbf{w}_{1:T} | \mathbf{y}_{1:T}, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right)}_{\text{Particle filter}},
\end{aligned} \tag{7}
$$

which involves a particle filter with a conditional Rao–Blackwellization achieved via a Kalman filter recursion conditional on each particles state realization. The Rao–Blackwellized particle filter estimate of (7) is given by the importance weighted particle approximation

$$
\begin{aligned}
&\widehat{p}\left(\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T} | \mathbf{y}_{1:T}, [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right) \\
&\quad = \sum_{i=1}^{N} [\tilde{\Xi}](j, i) p\left(\mathbf{g}_{1:T} | [\mathbf{h}_{1:T}, \mathbf{w}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta}](j, i), \mathbf{y}_{1:T}\right).
\end{aligned} \tag{8}
$$

We can now perform a Kalman filter recursion to obtain $p(\mathbf{g}_{1:T} | [\mathbf{h}_{1:T}, \mathbf{w}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\beta}](j, i), \mathbf{y}_{1:T})$ for each particle, which is optimal for this conditionally linear Gaussian state structure. The recursive construction of the SIR particle filter under Rao–Blackwellized scheme estimate in (8) therefore proceeds according to the following recursive steps involving unnormalized importance weights, given by

$$
[\widetilde{\Xi}_n](j, i) \propto [\Xi_{n-1}](j, i) p(\mathbf{y}_n | \mathbf{y}_{1:n-1}, [\mathbf{h}_{1:n}, \mathbf{w}_{1:n}, \boldsymbol{\alpha}, \boldsymbol{\beta}](j, i)),
$$

where the incremental importance weight is given by the marginal evidence at time n, $p(\mathbf{y}_n | \mathbf{y}_{1:n-1}, [\mathbf{h}_{1:n}, \mathbf{w}_{1:n}, \boldsymbol{\alpha}, \boldsymbol{\beta}](j, i))$ and it is directly obtained as an output in each stage of the Kalman filter. The conditional Kalman filter recursion at time n, for the i-th particle and the j-th marginal Metropolis proposed static parameters involves obtaining recursively the sufficient statistics for the conditional mean (MMSE) and covariance matrix of $p(\mathbf{g}_t | [\mathbf{h}_{1:t}, \mathbf{w}_{1:t}, \boldsymbol{\alpha}, \boldsymbol{\beta}](j, i), \mathbf{y}_{1:T})$ according to the following Kalman filter recursion:

$$
\begin{aligned}
&\widetilde{\mathbf{y}} = \mathbf{y}_n - f\left(s_n[\mathbf{h}_n](j, i)\right) [\boldsymbol{\mu}_{n|n-1}](j, i) \\
&\mathbf{S} = f(s_n[\mathbf{h}_n](j, i)) [\boldsymbol{\mu}_{n|n-1}](j, i) [\Sigma_{n|n-1}](j, i) f(s_n[\mathbf{h}_n](j, i)) \\
&\quad \times [\boldsymbol{\mu}_{n|n-1}](j, i) + \sigma_w^2 \mathbf{I} \\
&\mathbf{K} = [\Sigma_{n|n-1}](j, i) f\left(s_n[\mathbf{h}_n](j, i)\right) [\boldsymbol{\mu}_{n|n-1}](j, i)^\top \mathbf{S}^{-1} \\
&[\boldsymbol{\mu}_{n|n}](j, i) = [\boldsymbol{\mu}_{n-1|n-1}](j, i) + \mathbf{K} \widetilde{\mathbf{y}} \\
&[\Sigma_{n|n}](j, i) = \left(\mathbf{I} - \mathbf{K} f(s_n[\mathbf{h}_n](j, i)) [\boldsymbol{\mu}_{n|n-1}](j, i)\right) [\Sigma_{n|n-1}](j, i).
\end{aligned}
$$

2) Adaptive MCMC for Static Parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$*:* We now present the proposal kernel, $q([\boldsymbol{\alpha}, \boldsymbol{\beta}]^* | [\boldsymbol{\alpha}, \boldsymbol{\beta}](j))$, for the static parameters in (5). The static parameters are updated via an adaptive MH proposal comprised of a mixture of Gaussians. Adaptive MCMC attempts to improve the mixing rate by automatically learning better parameter values of the MCMC algorithm while it is running. In particular, one of the mixture components has a covariance structure which is adaptively learnt on-line. The mixture proposal distribution for parameters $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$ at iteration j of the Markov chain is given by,

$$
\begin{aligned}
&q\left([\boldsymbol{\alpha}, \boldsymbol{\beta}](j); [\boldsymbol{\alpha}, \boldsymbol{\beta}](j+1)\right) \\
&\quad = w_1 N \left([\boldsymbol{\alpha}, \boldsymbol{\beta}](j+1); [\boldsymbol{\alpha}, \boldsymbol{\beta}](j), \frac{(2.38)^2}{d} \Sigma_j \right) \\
&\qquad + (1 - w_1) N \left([\boldsymbol{\alpha}, \boldsymbol{\beta}](j+1); [\boldsymbol{\alpha}, \boldsymbol{\beta}](j), \frac{0.1^2}{d} \mathbf{I}_{2L,2L} \right),
\end{aligned} \tag{9}
$$

where $\mathbf{I}_{2L,2L}$ is the identity matrix of size $2L$. Here, Σ_j is the current empirical estimate of the covariance between the parameters of $\boldsymbol{\alpha}, \boldsymbol{\beta}$ estimated using samples from the PMCMC chain up to time j, and w_1 is a mixture proposals weight which we set according to the recommendation of [21] and are based on optimality conditions presented in [22]. The PMCMC joint channel and Doppler offset estimation algorithm is presented in Algorithm 1.

Algorithm 1 Joint Channel and Doppler Estimation via Particle Adaptive MCMC

Initialize $[\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}](1)$ by sampling each value from the corresponding priors.

for $j = 2, \ldots, J$ **do**

  Sample $[\boldsymbol{\alpha}, \boldsymbol{\beta}]^*(j+1) \sim q([\boldsymbol{\alpha}, \boldsymbol{\beta}](j); [\boldsymbol{\alpha}, \boldsymbol{\beta}](j+1))$ according to (9):

  Sample a realization u_1 from $U_1 \sim U[0,1]$

  if $u_1 \geq w_1$ then sample $[\boldsymbol{\alpha}, \boldsymbol{\beta}](j+1)$ **then**

    Sample $[\boldsymbol{\alpha}, \boldsymbol{\beta}](j+1)$ from the adaptive component as follows:

    Estimate Σ_j, the empirical covariance of α, β, using samples $\{[\boldsymbol{\alpha}, \boldsymbol{\beta}](i)\}_{i=1:j}$.

    Sample

$$[\boldsymbol{\alpha}, \boldsymbol{\beta}](j+1) \sim N\left(\boldsymbol{\alpha}, \boldsymbol{\beta}; [\boldsymbol{\alpha}, \boldsymbol{\beta}](j), \frac{2.38^2}{d}\Sigma_j\right);$$

  else

    Sample $[\boldsymbol{\alpha}, \boldsymbol{\beta}](j+1)$ from the non-adaptive component as follows:

    Sample

$$[\boldsymbol{\alpha}, \boldsymbol{\beta}](j+1) \sim N\left(\boldsymbol{\alpha}, \boldsymbol{\beta}; [\boldsymbol{\alpha}, \boldsymbol{\beta}](j), \frac{0.1^2}{d} I_{d,d}\right)$$

  end if

  Run the Rao–Blackwellized particle filter with N particles to obtain:

$$\widehat{p}\,(\mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T} | \mathbf{y}_{1:T}, [\boldsymbol{\alpha}, \boldsymbol{\beta}]^*(j+1))$$

$$\widehat{p}\,(\mathbf{y}_{1:T} | [\boldsymbol{\alpha}, \boldsymbol{\beta}]^*(j+1))$$

  Compute the MH acceptance probability in (4).

  Sample a realization u_1 from $U_1 \sim U[0,1]$.

  if

$$A\,([\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}](j) \quad [\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}]^*(j+1)) > u_1$$

  then

$$[\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}](j+1) = [\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}]^*(j+1)$$

  else

$$[\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}](j+1) = [\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{h}_{1:T}, \mathbf{g}_{1:T}, \mathbf{w}_{1:T}](j)$$

  end if

end for

IV. Cramér–Rao Lower Bound for Channel Coefficients and Doppler Offsets

In this section we derive the Bayesian Cramér–Rao Lower Bound (BCRLB) for the channel coefficients and the Doppler offset parameters.

A. Bayesian CRLB of Channel Coefficients $\mathbf{g}_{1:T}, \mathbf{h}_{1:T}$

The BCRLB provides a lower bound on the MSE matrix for estimation of the path space parameters which correspond in our model to the estimation of the latent process states $\mathbf{x}_{1:T} \triangleq \mathbf{g}_{1:T}, \mathbf{h}_{1:T}, \mathbf{w}_{1:T}$. We denote the Fisher Information Matrix (FIM), used in the CRLB, on the path space by $[F_{1:T}(\mathbf{x}_{1:T})](j)$ and marginally by $[F_n(\mathbf{x}_n)](j)$ for time n in the path space, conditional on the proposed static parameters at iteration j of the algorithm. Assuming regularity holds for the probability density functions, the BCRLB inequality states that the MSE of any estimator is bounded as:

$$\left[F_n^{-1}(\mathbf{X}_n)\right](j) \leq \mathbb{E}_{p(\mathbf{x}_n, \mathbf{y}_{1:n}|\boldsymbol{\alpha}, \boldsymbol{\beta})}\left[(\mathbf{X}_n - \widehat{\mathbf{X}}_n)(\mathbf{X}_n - \widehat{\mathbf{X}}_n)^H\right].$$

This formulation assumes prior knowledge of the static parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. In the following Lemma we develop a modified version of the BCRLB by marginalizing numerically over the static parameters.

Lemma 3: The Bayesian CRLB for the channel coefficients, $\mathbf{h}$ and $\mathbf{g}$, is expressed as

$$\mathbb{E}_{p(\mathbf{x}_n, \mathbf{y}_{1:n})}\left[(\mathbf{X}_n - \widehat{\mathbf{X}}_n)(\mathbf{X}_n - \widehat{\mathbf{X}}_n)^H\right]$$

$$\simeq \frac{1}{J}\sum_{j=1}^{J} \mathbb{E}_{p(\mathbf{x}_n, \mathbf{y}_{1:n}|\boldsymbol{\alpha}, \boldsymbol{\beta})}\left\{[\mathbf{X}_n - \widehat{\mathbf{X}}_n][\mathbf{X}_n - \widehat{\mathbf{X}}_n]^H | [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right\},$$

where $\mathbf{x}_n \triangleq [\mathbf{g}_n, \mathbf{h}_n, \mathbf{w}_n]$.

Proof:

$$\begin{aligned}
&\mathbb{E}_{p(\mathbf{x}_n, \mathbf{y}_{1:n})}\left[(\mathbf{X}_n - \widehat{\mathbf{X}}_n)(\mathbf{X}_n - \widehat{\mathbf{X}}_n)^H\right] \\
&= \int \left\{[\mathbf{X}_n - \widehat{\mathbf{X}}_n][\mathbf{X}_n - \widehat{\mathbf{X}}_n]^H\right\} \\
&\quad \times p(\mathbf{x}_n, \mathbf{y}_{1:n}, \boldsymbol{\alpha}, \boldsymbol{\beta}) d\mathbf{x}_n d\mathbf{y}_{1:n} \mathrm{d}\boldsymbol{\alpha} \mathrm{d}\boldsymbol{\beta} \\
&= \int \mathbb{E}_{p(\mathbf{x}_n, \mathbf{y}_{1:n}|\boldsymbol{\alpha}, \boldsymbol{\beta})}\left\{[\mathbf{X}_n - \widehat{\mathbf{X}}_n][\mathbf{X}_n - \widehat{\mathbf{X}}_n]^H | \boldsymbol{\alpha}, \boldsymbol{\beta}\right\} p(\boldsymbol{\alpha}, \boldsymbol{\beta}) \mathrm{d}\boldsymbol{\alpha} \mathrm{d}\boldsymbol{\beta} \\
&\approx \frac{1}{J}\sum_{j=1}^{J} \mathbb{E}_{p(\mathbf{x}_n, \mathbf{y}_{1:n}|\boldsymbol{\alpha}, \boldsymbol{\beta})}\left\{[\mathbf{X}_n - \widehat{\mathbf{X}}_n][\mathbf{X}_n - \widehat{\mathbf{X}}_n]^H | [\boldsymbol{\alpha}, \boldsymbol{\beta}](j)\right\}.
\end{aligned}$$

■

Next, we develop a recursive algorithm to perform the evaluation of the BCRLB for the expectation $\mathbb{E}_{p(\mathbf{x}_n, \mathbf{y}_{1:n}|\boldsymbol{\alpha}, \boldsymbol{\beta})}\{[\mathbf{X}_n - \widehat{\mathbf{X}}_n][\mathbf{X}_n - \widehat{\mathbf{X}}_n]^H | [\alpha, \beta](j)\}$. This is based on the representation in Lemma 3 which can be combined with the BCRLB recursive expression for the channel coefficients given in Theorem 1.

Theorem 1: The Bayesian CRLB for the channel coefficients, $\mathbf{h}$ and $\mathbf{g}$ is given in (10), shown at the bottom of the next page.

Proof: See Appendix A. ■

B. Bayesian CRLB of Doppler Offset Parameters α, β

We now develop a numerical solution to obtain the CRLB for the Doppler offset parameters. The expression for the FIM for α, β is:

$$\nabla\nabla_\theta \int p(\mathbf{y}_{1:T}|\mathbf{x}_{1:T}, \theta) p(\mathbf{x}_{1:T}|\theta) d\mathbf{x}_{1:T}$$

where $\theta = [\alpha, \beta]$. To begin, we obtain an unbiased particle estimate of $\int p(\mathbf{y}_{1:T}|\mathbf{x}_{1:T}, \theta) p(\mathbf{x}_{1:T}|\theta) d\mathbf{x}_{1:T}$ in each stage of the PMCMC algorithm. This results in a set of J evaluations of this marginal likelihood from the PMCMC algorithm giving $\{\widehat{p}(\mathbf{y}_{1:T}|[\theta](j))\}_{j=1:J}$ according to the unbiased estimator described previously. We then obtain the following expression:

$$\begin{aligned} p(\mathbf{y}_{1:T}|\theta) &= \int p(\mathbf{y}_{1:T}|\mathbf{x}_{1:T}, \theta) p(\mathbf{x}_{1:T}|\theta) d\mathbf{x}_{1:T} \\ &\approx \widehat{p}(\mathbf{y}_{1:T}|\theta) \\ &= \prod_{n=1}^{T} \left(\frac{1}{N} \sum_{i=1}^{N} [\xi_n](j, i) \right), \end{aligned} \tag{12}$$

with incremental particle weight $[xi_n](j, i)$ given at iteration j of the PMCMC for the i-th particle a function of parameter θ. The estimate in (12) is unbiased as studied in [23] and furthermore, according to [23, Proposition 9.4.1, page 301] the following central limit theorem holds,

$$\sqrt{N} \left(\frac{\widehat{p}(\mathbf{y}_{1:T}|\theta)}{p(\mathbf{y}_{1:T}|\theta)} - 1 \right) \xrightarrow{d} \mathcal{N}\left(0, \gamma^2(\theta)\right)$$

for a $\gamma^2(\theta)$ which is problem specific and finite. Therefore, we can confidently utilize these estimates $\{\widehat{p}(\mathbf{y}_{1:T}|[\theta](j))\}_{j=1:J}$ from each iteration of the PMCMC algorithm to construct a continuous and differentiable kernel density estimator given generically for all θ by

$$\widehat{p}_K(\mathbf{y}_{1:T}|\theta) = \frac{1}{Jh} \sum_{j=1}^{J} K\left(\frac{\theta - [\theta](j)}{h} \right),$$

where $K(\cdot)$ is the kernel which is a symmetric but not necessarily positive function that integrates to one and $h > 0$ is a smoothing parameter called the bandwidth. Given the smooth kernel density estimator which can be differentiated to obtain the Hessian around the mode and obtain the CRLB for the Doppler offset parameters:

$$\begin{aligned} \nabla\nabla_\theta \int p(\mathbf{y}_{1:T}|\mathbf{x}_{1:T}, \theta) p(\mathbf{x}_{1:T}|\theta) d\mathbf{x}_{1:T} \\ \approx \nabla\nabla_\theta\, \widehat{p}_K(\mathbf{y}_{1:T}|\theta)\big|_{\theta=\text{mode}(\theta)}. \end{aligned} \tag{13}$$

V. Algorithmic Computational Complexity Analysis

We now derive the computational complexity comparison between each of the algorithms. The computational cost of each of these algorithms is split into three parts: the first cost involves constructing and sampling from the proposal; the second significant computational cost comes from the evaluation of the acceptance probability for the proposed new Markov chain state; and the third is related to the mixing rate of the overall MCMC algorithm as affected by the length of the Markov chain required to obtain estimators of a desired accuracy. We define the following building blocks for a single MCMC iteration and their associated complexity, measured by $\mathbb{O}(\cdot)$ as order of magnitude and C_m and C_a are the complex multiplications and complex additions, respectively:

1) Exact sampling of a random variable $\approx \mathbb{O}(1)$.
2) Likelihood evaluation of

$$\prod_{n=1}^{T} \prod_{l=1}^{L} p\left(y_n^{(l)} | \alpha^{(l)}, \beta^{(l)}, g_n^{(l)}, h_n^{(l)} \right) \approx TL(C_m + C_a) + \mathbb{O}(1).$$

$$\begin{aligned} &\left[J_n(\widehat{\mathbf{X}}_n) \right](j) \\ &= \begin{bmatrix} \frac{1}{1-[\beta](j)^2} & 0 & 0 \\ 0 & \frac{1}{1-[\alpha](j)^2} & 0 \\ 0 & 0 & \frac{1}{\sigma_w^2} \end{bmatrix} \\ &\quad - \frac{1}{\sigma_v^2} \begin{bmatrix} f^2(r_n) & g_n\left(f(r_n) \frac{\partial f(r_n)}{\partial h_n} + \frac{\partial f^2(r_n)}{\partial h_n} \right) & g_n\left(f(r_n) \frac{\partial f(r_n)}{\partial w_n} + \frac{\partial f^2(r_n)}{\partial w_n} \right) \\ g_n\left(f(r_n) \frac{\partial f(r_n)}{\partial h_n} + \frac{\partial f^2(r_n)}{\partial h_n} \right) & -g_n^2 \left(\frac{\partial f(r_n)}{\partial h_n} \right)^2 \frac{\partial^2 f(r_n)}{\partial h_n^2} & g_n^2 \frac{\partial f(r_n)}{\partial h_n} \frac{\partial f(r_n)}{\partial w_n} \\ g_n f(r_n) \frac{\partial f(r_n)}{\partial w_n} & g_n^2 \frac{\partial f(r_n)}{\partial w_n} \frac{\partial f(r_n)}{\partial h_n} & g_n^2 \left(\frac{\partial f(r_n)}{\partial w_n} \right)^2 \end{bmatrix} \\ &\quad - \begin{bmatrix} \frac{[\beta](j)}{1-[\beta](j)^2} & 0 & 0 \\ 0 & \frac{[\alpha](j)}{1-[\alpha](j)^2} & 0 \\ 0 & 0 & 0 \end{bmatrix} \left(\left[J_{n-1}(\widehat{\mathbf{X}}_{n-1}) \right](j) + \begin{bmatrix} \frac{[\beta](j)^2}{1-[\beta](j)^2} & 0 & 0 \\ 0 & \frac{[\alpha]^2(j)}{1-[\alpha](j)^2} & 0 \\ 0 & 0 & 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} \frac{[\beta](j)}{1-[\beta](j)^2} & 0 & 0 \\ 0 & \frac{[\alpha](j)}{1-[\alpha](j)^2} & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned} \tag{10}$$

3) Prior evaluations of

$$\prod_{n=1}^{T}\prod_{l=1}^{L} p\left(h_n^{(l)}|h_{n-1}^{(l)}\right) p\left(g_n^{(l)}|g_{n-1}^{(l)}\right) p\left(g_1^{(l)}\right) p\left(h_1^{(l)}\right)$$
$$\times p\left(\alpha^{(l)}\right) p\left(\beta^{(l)}\right) \approx 6TL(C_m + C_a) + \mathbb{O}(1).$$

Based on these building blocks we estimate the overall complexity of the proposed algorithms as follows.

A. Computational Complexity of Proposed PMCMC Algorithm

1) Adaptive MCMC Component in (9): Complexity $(2L \times 2L + 2L)\mathbb{O}(1)$.

2) Rao–Blackwellized SIR Filter Component (7):

- Kalman filter component: $TL\mathbb{O}(1)$.
- SIR filter component: $2NLT\mathbb{O}(1)$.
- Evaluation of marginal likelihood: $NT\mathbb{O}(1)$.
- Sampling SIR filter path space proposal: $N\mathbb{O}(1)$.

3) Evaluation of Acceptance Probability in (4): Complexity $(NT + 4L)\mathbb{O}(1)$.

Therefore, the total cost of a single PMCMC iteration can be approximated as $(2L^2 + TL + NT(2L + 2) + N)\mathbb{O}(1)$.

Now that we have obtained the computational complexity of each MCMC iteration for both methods, we are able to perform a fair comparison with respect to algorithmic complexity. This will be presented in the next section.

VI. Simulation Results

In this section, we present the performance of the proposed algorithm via Monte Carlo simulations. Simulation Set-Up: the channels are generated according to Rayleigh flat-fading channel model with Jakes' Doppler spectrum [24]. We consider a carrier frequency of 6 GHz and a bandwidth of 10 kHz, which is suitable for IEEE 802.16e [25]. The velocities of the mobile terminal and the relay were set to 80 km/h and 100 km/h, respectively, which correspond to $\alpha = 0.98$ and $\beta = 0.95$. The network topology used in the simulations involved a single relay network, $L = 1$, and the relay processing function $f^{(l)}(R^{(l)})$ is a constant gain Amplify and Forward. In all simulations we ran a Markov chain of length 20,000 iterations, discarding the first 5,000 samples as burnin, and used $N = 5000$ particles.

A. Analysis of Estimated MSE and BER Versus SNR

In this section we compare the performance of the proposed PMCMC algorithm with: (i) MCMC-within-Gibbs sampler of [26], where the number of iterations of the Gibbs sampler is set such its overall computational complexity is the same as our PMCMC algorithm. (ii) Extended Kalman filter based approach coupled with Maximum Likelihood (see [27] for details) to estimate the Doppler offsets α, β. These results are evaluated for the channel estimation $\mathbf{h}, \mathbf{g}$ and Doppler offsets α, β MSE, followed by BER curves for different number of relays. In addition we present the BCRLB results which serve as the lower bound on the achievable MSE. The MSE results are presented in Fig. 2. Clearly, there is a decrease in the MSE as the SNR increases. In addition, since the estimate of the channels $\mathbf{g}_{1:T}$ is performed using the Rao–Blackwellizing conditionally optimal Kalman Filter, this is reflected in the level of the total MSE. The estimates for $\mathbf{g}_{1:T}$ are more accurate than those for the particle filter sampled estimates $\mathbf{h}_{1:T}$. We note that the PMCMC algorithm achieves lower MSE than the Gibbs sampler at every SNR. This indicates the superiority of the PMCMC algorithm over the Gibbs sampler. We also note that the Extended Kalman filter is unable to track correctly the non-linear dynamic observation models, resulting in poor estimates of the channels as well as the Doppler offsets. In addition we present the BCRLB according to Theorem 1. The results demonstrate that the proposed PMCMC algorithm performs close to optimally.

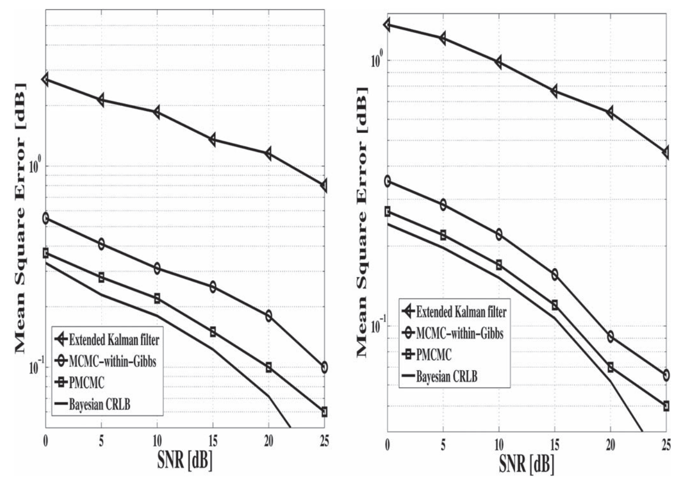


Fig. 2. Channel estimation performance for $\mathbf{h}$ (left panel) and $\mathbf{g}$ (right panel) of the PMCMC algorithm compared with Gibbs-within-MCMC [26] and the Extended Kalman filter [27] algorithms and the BCRLB. In both figures the performance of the PMCMC algorithm is very close to the CRLB.

In Fig. 3 we present the MSE for the Doppler offsets $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as a function of the SNR. We compare these with the BCRLB developed in (13). Clearly as the SNR increases, the estimates converge to the true parameter values. The results show that the estimation of the Doppler offset $\boldsymbol{\beta}$ is more accurate than of $\boldsymbol{\alpha}$, which is due to the use of the Rao–Blackwell theorem.

We now present the effect that the channel estimation quality has on BER as a function of number of relays. To perform the detection, we implement the Maximum Likelihood detection algorithm. In Fig. 4 we present the BER curves for different number of relays $L = \{1, 4, 8\}$, where we used 2 pilot symbols at the beginning of each frame followed by 98 data symbols (i.e. frame length of 100 symbols). The results demonstrate that spatial diversity does exist and that the BER provided by our PMCMC algorithm is much lower than both the MCMC-within-Gibbs approach as well as the Extended Kalman filter.

VII. Conclusion

We solved the problem of *joint channel and Doppler offset estimation* in dynamic cooperative wireless mobile relay networks. We developed two algorithms to jointly estimate the

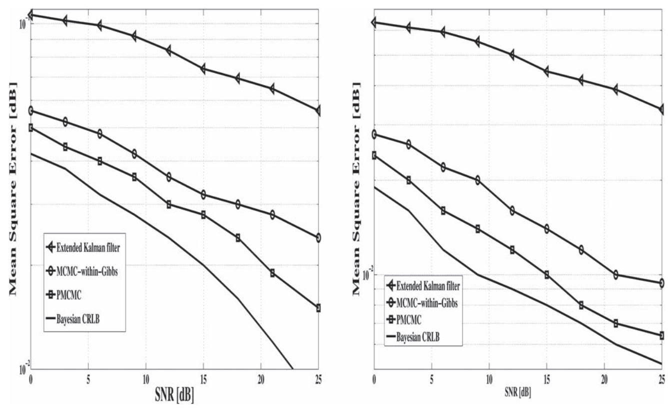


Fig. 3. Estimation of Doppler offsets α (left panel) and β (right panel) of the PMCMC compared with Gibbs-within-MCMC sampler of [26] and the Extended Kalman filter approach of [27]. We also plot the BCRLB.

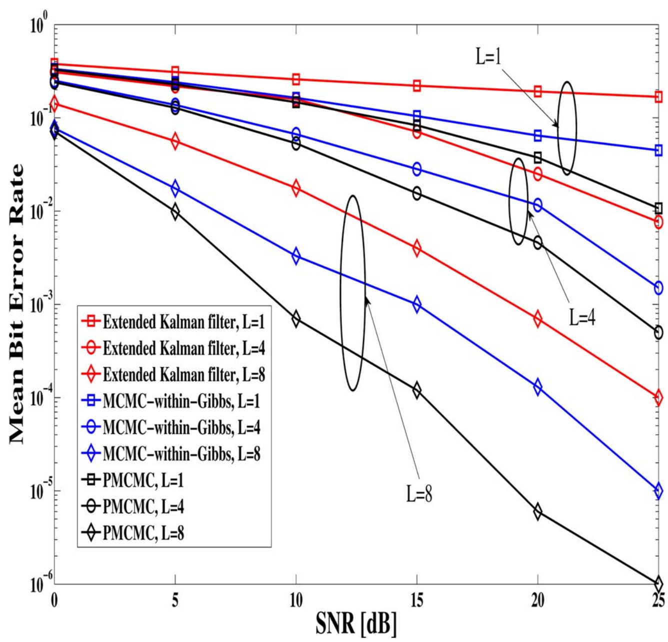


Fig. 4. Bit error rate (BER) curves for $L = \{1, 4, 8\}$ relays for the PMCMC algorithm and the MCMC-within-Gibbs sampler of [26] and the Extended Kalman filter approach of [27].

time-varying channels and static Doppler offsets. We presented complexity analysis of the proposed algorithms and showed that they are very efficient. We developed a novel version of the recursive marginal Cramér–Rao lower bound and derived expressions for the achievable MSE. Simulation results demonstrated that the proposed algorithms perform close to the Cramér–Rao Lower Bound.

APPENDIX A
BAYESIAN CRLB ELEMENTS OF THEOREM 1

We utilize the recursive expression for the FIM [28]:

$$\left[J_n(\widehat{\mathbf{X}}_n)\right](j) = \left[D_{n-1}^{22}(\widehat{\mathbf{x}}_n)\right](j) - \left[D_{n-1}^{21}(\widehat{\mathbf{x}}_n)\right](j) \times \left(\left[J_{n-1}(\widehat{\mathbf{x}}_n)\right](j) + \left[D_{n-1}^{11}(\widehat{\mathbf{x}}_n)\right](j)\right)^{-1} \left[D_{n-1}^{12}(\widehat{\mathbf{x}}_n)\right](j).$$

Utilsing Lemma 3, and conditional on the previous Markov chain state $[\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{X}_{1:T}](j-1)$ and the new sampled Markov chain proposal for the static parameters at iteration j, $[\boldsymbol{\alpha}, \boldsymbol{\beta}](j)$, we obtain the BCRLB.

We derive the terms $[J_1(\widehat{\mathbf{x}}_1)](j)$, $[D_{n-1}^{12}(\widehat{\mathbf{x}}_n)](j)$ and $[D_{n-1}^{11}](j)$ for the Bayesian CRLB:

$$\begin{aligned}\left[J_1(\widehat{\mathbf{x}}_1)\right](j) &= -\mathbb{E}\left[\boldsymbol{\nabla}_{\mathbf{x}_1}\left\{\boldsymbol{\nabla}_{\mathbf{x}_1}\log p(\mathbf{x}_1)\right\}^T\right] \\ &= -\mathbb{E}\left[\boldsymbol{\nabla}_{\mathbf{x}_1}\left\{\boldsymbol{\nabla}_{\mathbf{x}_1}\left[\frac{g_1^2}{2\sigma_g^2}, \frac{h_1^2}{2\sigma_h^2}, \frac{w_1^2}{2\sigma_w^2}\right]\right\}^T\right] \\ &= \begin{bmatrix}\frac{1}{\sigma_{\mathbf{g}}^2} & 0 & 0 \\ 0 & \frac{1}{\sigma_{\mathbf{h}}^2} & 0 \\ 0 & 0 & \frac{1}{\sigma_{\mathbf{w}}^2}\end{bmatrix}.\end{aligned}$$

$$\left[D_{n-1}^{11}\right](j) = -\mathbb{E}\left[\boldsymbol{\nabla}_{\mathbf{x}_{n-1}}\left\{\boldsymbol{\nabla}_{\mathbf{x}_{n-1}}\log p(\mathbf{x}_n|\mathbf{x}_{n-1})\right\}^T\right]$$

$$= -\mathbb{E}\left[\boldsymbol{\nabla}_{\mathbf{x}_{n-1}}\left\{\boldsymbol{\nabla}_{\mathbf{x}_{n-1}}\left[\frac{(g_n - \beta g_{n-1})^2}{2(1-(\beta)^2)}, \frac{(h_n - \alpha h_{n-1})^2}{2(1-(\alpha)^2)}, \frac{w_n^2}{2\sigma_w^2}\right]\right\}^T\right]$$

$$= \begin{bmatrix}\frac{[\beta](j)^2}{1-[\beta](j)^2} & 0 & 0 \\ 0 & \frac{[\alpha](j)^2}{1-[\alpha](j)^2} & 0 \\ 0 & 0 & 0\end{bmatrix}.$$

$$\left[D_{n-1}^{12}(\widehat{\mathbf{x}}_n)\right](j) = \left[D_{n-1}^{21}(\widehat{\mathbf{x}}_n)\right](j)$$

$$= -\mathbb{E}\left[\boldsymbol{\nabla}_{\mathbf{x}_n}\left\{\boldsymbol{\nabla}_{\mathbf{x}_{n-1}}\log p\left(\mathbf{x}_n|\mathbf{x}_{n-1}\right)\right\}^T\right]$$

$$= -\mathbb{E}\left[\boldsymbol{\nabla}_{\mathbf{x}_n}\left\{\boldsymbol{\nabla}_{\mathbf{x}_{n-1}}\left[\frac{(g_n - \beta g_{n-1})^2}{2(1-(\beta)^2)}, \frac{(h_n - \alpha h_{n-1})^2}{2(1-(\alpha)^2)}, \frac{w_n^2}{2\sigma_w^2}\right]\right\}^T\right]$$

$$= \begin{bmatrix}\frac{[\beta](j)}{1-[\beta](j)^2} & 0 & 0 \\ 0 & \frac{[\alpha](j)}{1-[\alpha](j)^2} & 0 \\ 0 & 0 & 0\end{bmatrix}.$$

$$\left[D_{n-1}^{22}(\widehat{\mathbf{x}}_n)\right](j) = -\mathbb{E}\left[\boldsymbol{\nabla}_{\mathbf{x}_n}\left\{\boldsymbol{\nabla}_{\mathbf{x}_n}\log p(\mathbf{x}_n|\mathbf{x}_{n-1})\right\}^T\right] - \mathbb{E}\left[\boldsymbol{\nabla}_{\mathbf{x}_n}\left\{\boldsymbol{\nabla}_{\mathbf{x}_n}\log p\left(\mathbf{y}_n|\mathbf{x}_n\right)\right\}^T\right]$$

$$= -\mathbb{E}\left[\boldsymbol{\nabla}_{\mathbf{x}_n}\left\{\boldsymbol{\nabla}_{\mathbf{x}_n}\left[\frac{(g_n - \beta g_{n-1})^2}{2(1-(\beta)^2)}, \frac{(h_n - \alpha h_{n-1})^2}{2(1-(\alpha)^2)}, \frac{w_n^2}{2\sigma_w^2}\right]\right\}^T\right] - \mathbb{E}\left[\boldsymbol{\nabla}_{\mathbf{x}_n}\left\{\boldsymbol{\nabla}_{\mathbf{x}_n}\left[\frac{\left(y_n - f\left(sh_n + w_n\right)g_n\right)^2}{2\sigma_v^2}\right]\right\}^T\right],$$

which is given in (14), shown at the bottom of the page.

$$\left[D_{n-1}^{22}(\widehat{\mathbf{x}}_n)\right](j) = \begin{bmatrix}\frac{1}{1-[\beta](j)^2} & 0 & 0 \\ 0 & \frac{1}{1-[\alpha](j)^2} & 0 \\ 0 & 0 & \frac{1}{\sigma_w^2}\end{bmatrix} - \begin{bmatrix}\frac{f^2(r_n)}{\sigma_v^2} & \frac{g_n f(r_n)\frac{\partial f(r_n)}{\partial h_n} + \frac{\partial f^2(r_n)}{\partial h_n} g_n}{\sigma_v^2} & \frac{g_n f(r_n)\frac{\partial f(r_n)}{\partial w_n} + \frac{\partial f^2(r_n)}{\partial w_n} g_n}{\sigma_v^2} \\ \frac{g_n f(r_n)\frac{\partial f(r_n)}{\partial h_n} + \frac{\partial f^2(r_n)}{\partial h_n} g_n}{\sigma_v^2} & \frac{-g_n^2\left(\frac{\partial f(r_n)}{\partial h_n}\right)^2 \frac{\partial^2 f(r_n)}{\partial h_n^2}}{\sigma_v^2} & \frac{g_n^2 \frac{\partial f(r_n)}{\partial h_n}\frac{\partial f(r_n)}{\partial w_n}}{\sigma_v^2} \\ \frac{g_n f(r_n)\frac{\partial f(r_n)}{\partial w_n}}{\sigma_v^2} & \frac{g_n^2 \frac{\partial f(r_n)}{\partial w_n}\frac{\partial f(r_n)}{\partial h_n}}{\sigma_v^2} & \frac{g_n^2\left(\frac{\partial f(r_n)}{\partial w_n}\right)^2}{\sigma_v^2}\end{bmatrix} \quad (14)$$

REFERENCES

[1] E. Van Der Meulen, "Three-terminal communication channels," *Advances in Applied Probability*, vol. 3, no. 1, pp. 120–154, 1971.

[2] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 74–80, Oct. 2004.

[3] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[4] X. Zhou, T. Lamahewa, and P. Sadeghi, "Kalman filter-based channel estimation for amplify and forward relay communications," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, 2009, pp. 1498–1502.

[5] K. Kim, R. Iltis, and H. Poor, "Frequency offset and channel estimation in cooperative relay networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3142–3155, Sep. 2011.

[6] C. Patel, G. Stuber, and T. Pratt, "Simulation of Rayleigh-faded mobile-to-mobile communication channels," *IEEE Trans. Commun.*, vol. 53, no. 10, pp. 1876–1884, Oct. 2005.

[7] E. Simon, L. Ros, H. Hijazi, and M. Ghogho, "Joint carrier frequency offset and channel estimation for OFDM systems via the em algorithm in the presence of very high mobility," *IEEE Trans. Signal Process.*, vol. 603, no. 2, pp. 754–765, Feb. 2012.

[8] K. Kim, M. Pun, and R. Iltis, "Joint carrier frequency offset and channel estimation for uplink MIMO-OFDMA systems using parallel Schmidt Rao-Blackwellized particle filters," *IEEE Trans. Commun.*, vol. 58, no. 9, pp. 2697–2708, Sep. 2010.

[9] L. He, S. Ma, Y.-C. Wu, and T.-S. Ng, "Semiblind iterative data detection for OFDM systems with CFO and doubly selective channels," *IEEE Trans. Commun.*, vol. 58, no. 12, pp. 3491–3499, Dec. 2010.

[10] L. He, Y.-C. Wu, S. Ma, T.-S. Ng, and H. V. Poor, "Superimposed training-based channel estimation and data detection for OFDM amplify-and-forward cooperative systems under high mobility," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 274–284, Jan. 2012.

[11] C. Andrieu, N. de Freitas, A. Doucet, and M. Jordan, "An introduction to MCMC for machine learning," *Mach. Learn.*, vol. 50, no. 1/2, pp. 5–43, Jan. 2003.

[12] J. Laneman and G. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[13] J. Laneman and G. Wornell, "Exploiting distributed spatial diversity in wireless networks," in *Proc. Allerton Conf. Commun., Control, Comput.*, 2000, pp. 1–10.

[14] H. Wang and P. Chang, "On verifying the first-order Markovian assumption for a Rayleigh fading channel model," *IEEE Trans. Veh. Technol.*, vol. 45, no. 2, pp. 353–357, May 1996.

[15] S. Ghandour-Haidar, L. Ros, and J.-M. Brossier, "On the use of first-order autoregressive modeling for Rayleigh flat fading channel estimation with Kalman filter," *Signal Process.*, vol. 92, no. 2, pp. 601–606, 2012.

[16] A. Nasir, S. Durrani, and R. Kennedy, "Mixture Kalman filtering for joint carrier recovery and channel estimation in time-selective Rayleigh fading channels," in *Proc. IEEE ICASSP*, 2011, pp. 3496–3499.

[17] T. Ghirmai, "Sequential Monte Carlo method for fixed symbol timing estimation and data detection," in *Proc. Conf. Inf. Sci. Syst.*, 2006, pp. 1291–1295.

[18] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1907–1916, May 2008.

[19] A. Behbahani and A. Eltawil, "On channel estimation and capacity for amplify and forward relay networks," in *Proc. IEEE Global Telecommun. Conf.*, 2008, pp. 1–5.

[20] C. Andrieu, A. Doucet, and R. Holenstein, "Particle Markov chain Monte Carlo methods," *J. R. Stat. Soc. Series B*, vol. 72, no. 3, pp. 269–342, Jun. 2010.

[21] G. Roberts and J. Rosenthal, "Examples of adaptive MCMC," *J. Comput. Graphical Statist.*, vol. 18, no. 2, pp. 349–367, 2009.

[22] G. Roberts and J. Rosenthal, "Optimal scaling for various Metropolis-Hastings algorithms," *Statist. Sci.*, vol. 16, no. 4, pp. 351–367, Nov. 2001.

[23] P. Del Moral, *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems With Applications*. New York, NY, USA: Springer-Verlag, 2004.

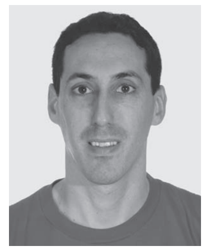
[24] W. Jakes and D. Cox, *Microwave Mobile Communications*. Hoboken, NJ, USA: Wiley-IEEE Press, 1994.

[25] I. W. Group, IEEE Standard, vol. 802 IEEE Standard for Local and Metropolitan Area Networks-Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems-Amendment 3: Management Plane Procedures and Services 2005, IEEE Standard, vol. 802.

[26] F. Septier and Y. Delignon, "MCMC sampling for joint estimation of phase distortions and transmitted symbols in OFDM systems," *Digit. Signal Process.*, vol. 21, no. 2, pp. 341–353, Mar. 2011.

[27] X. Sun, L. Jin, and M. Xiong, "Extended Kalman filter for estimation of parameters in nonlinear state-space models of biochemical networks," *PloS One*, vol. 3, no. 11, p. e3758, 2008.

[28] P. Tichavský, C. Muravchik, and A. Nehorai, "Posterior Cramer-Rao bounds for discrete-time nonlinear filtering," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, May 1998.



Ido Nevat received the B.Sc. degree in electrical engineering from the Technion-Israel Institute of Technology, Haifa, Israel, in 1998 and the Ph.D. degree in electrical engineering from the University of New South Wales, Sydney, Australia, in 2010. From 2010 to 2013, he was a Research Fellow with the Wireless and Networking Technologies Laboratory, Commonwealth Scientific and Industrial Research Organisation, Australia. Currently, he is a Scientist with the Institute for Infocomm Research (I2R), Singapore. His main areas of interests include statistical signal processing and Bayesian statistical modeling.

Gareth W. Peters is an Assistant Professor with the Department of Statistical Science, Principle Investigator in Computational Statistics and Machine Learning, and Academic Member of the U.K. Ph.D. Centre of Financial Computing, University College London, London, U.K. He is also an Adjunct Scientist with the Commonwealth Scientific and Industrial Research Organisation, Australia; an Associate Member with the Oxford-Man Institute, Oxford University, Oxford, U.K.; and an Associate Member with the Systemic Risk Centre, London School of Economics, London.

Arnaud Doucet received the Ph.D. degree from the University Paris XI, Orsay Cedex, France, in 1997. He has held previous faculty positions with Cambridge University, Cambridge, U.K.; Melbourne University, Richmond, Australia; University of British Columbia, Vancouver, BC, Canada; and the Institute of Statistical Mathematics, Tokyo, Japan. Since 2011, he has been a Professor with the Department of Statistics, Oxford University, Oxford, U.K. His main research interests are Monte Carlo methods and Bayesian statistics.

Jinhong Yuan (M'02–SM'11) received the B.E. and Ph.D. degrees in electronics engineering from Beijing Institute of Technology, Beijing, China, in 1991 and 1997, respectively. From 1997 to 1999, he was a Research Fellow with the School of Electrical Engineering, University of Sydney, Sydney, Australia. In 2000, he joined the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, Australia, where he is currently a Telecommunications Professor. He has published two books, three book chapters, over 200 papers in telecommunications journals and conference proceedings, and 40 industrial reports. He is a coinventor of one patent on MIMO systems and two patents on low-density-parity-check codes. He has coauthored three Best Paper Awards and one Best Poster Award, including the Best Paper Award from the IEEE Wireless Communications and Networking Conference, Cancun, Mexico, in 2011, and the Best Paper Award from the IEEE International Symposium on Wireless Communications Systems, Trondheim, Norway, in 2007. He is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS. He serves as the IEEE NSW Chair of Joint Communications/Signal Processions/Ocean Engineering Chapter. His current research interests include error control coding and information theory, communication theory, and wireless communications.