

# Sequential estimation of time-varying multipath channel for MIMO-OFDM systems

Daniele Angelosante

DAEIMI

Università di Cassino

Cassino, Italy

email: d.angelosante@unicas.it

Ezio Biglieri

Departament de Tecnologia

Universitat Pompeu Fabra

Barcelona, Spain

email: e.biglieri@ieee.org

Marco Lops

DAEIMI

Università di Cassino

Cassino, Italy

email: lops@unicas.it

**Abstract**—In this paper, we introduce a pilot-aided multipath channel estimator for Multiple-Input Multiple-Output (MIMO) Orthogonal Frequency Division Multiplexing (OFDM) systems. Typical estimation algorithms assume the number of multipath components and delays to be known and constant, while their amplitudes may vary in time. In this work, we focus on the more realistic assumption that also the number of channel taps is unknown and time-varying. The estimation problem arising from this assumption is solved using Random Set Theory (RST), which is a probability theory of finite sets. Due to the lack of a closed form of the optimal filter, a Rao-Blackwellized Particle Filter (RBPF) implementation of the channel estimator is derived. Simulation results demonstrate the estimator effectiveness.

## I. INTRODUCTION

Wireless systems employing multiple antennas at the transmitter and the receiver have been shown to provide a significant increase in the channel capacity [1]. On the other hand, orthogonal frequency division multiplexing (OFDM) is a technique that has attracted a lot of attention due to its simple implementation and robustness against frequency-selective fading channels [2]. The effectiveness of the combination of OFDM with multiple antennas has been demonstrated in [3]. However, such systems rely upon the knowledge of channel state information (CSI) at the receiver which can be achieved in different ways; one is based on pilot tones that contain symbols known to the receiver, whereas the other is blind, i.e. it does not need the transmission of known symbols exploiting statistical information. Compared with pilot-aided, blind channel estimation generally requires long data record and it is typically adopted to slowly time-varying channels and entails high complexity. In this paper, we consider rapidly time-varying channels, so we restrict our attention to training-based channel estimation.

Traditional channel estimation for MIMO-OFDM is addressed under the assumption that the number of multipath components and their delays is known and constant, while amplitudes are time-varying. In this work, we deal with multipath channel estimation under the more realistic assumption of an unknown and time-varying number of paths. Here, we use Random-Set Theory (RST) which is a probability theory defined on collections of elements where not only each element, but also the number of elements is random. With this theory, Bayesian filtering equations can be derived which allow one

to determine the evolution with time of the channel estimator. Due to the lack of an explicit solution for the Bayesian filtering equations for Random Finite Set (RFS), algorithms that approximate the filter equations are a fundamental issue in RST [5]. Among all techniques, Sequential Monte Carlo (SMC), also known as Particle Filter (PF), seems particularly attractive providing a consistent approximation of the Bayes filtering equation [6]. Hence, under the assumption that the channel dynamic is conditionally linear and Gaussian (CLG), we derived a Rao-Blackwellized particle filter (RBPF), which is an efficient implementation of the Bayes filter [7].

The rest of this paper is organized as follows. In Section II, we briefly overview the basic system model. In Section III, we introduce the Bayes filtering equations and their implementation via SMC filter. A Rao-Blackwellised particle filter is designed in Section IV. Section V contains simulation results.

*Notation:* Normal face letters denote scalar values; Lower (upper) boldface letter are used for column vectors (matrices); upper calligraphic letters are used for RFS;  $(\cdot)^H$  denotes Hermitian operation;  $E[\cdot]$  represents statistical expectation;  $\mathbf{I}_N$  denotes the identity matrix of size  $N$ ;  $\mathbf{0}_{M \times N}$  denotes the  $N \times M$  zeroes matrix; if  $\mathbf{x}_t$  are column vectors of the same dimension  $\forall t = 1, \dots, T$  then  $\mathbf{x}_{1:T} = [\mathbf{x}_1, \dots, \mathbf{x}_T]$ ;  $\text{diag}(\mathbf{x})$  stands for the diagonal matrix with the column vector  $\mathbf{x}$  on its diagonal;  $\det(\mathbf{\Sigma})$  is the determinant of the square matrix  $\mathbf{\Sigma}$ ; finally,  $j = \sqrt{-1}$  and  $\mathcal{N}_c(\mathbf{x}; \boldsymbol{\mu}, \mathbf{\Sigma}) = \frac{1}{\pi \det(\mathbf{\Sigma})} \exp\{-(\mathbf{x} - \boldsymbol{\mu})^H \mathbf{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\}$ .

## II. SIGNAL MODEL AND PROBLEM STATEMENT

In this section, we briefly review the signal model for MIMO-OFDM systems. Successively, we introduce our dynamic channel model based on RFS. Finally, we show that the whole system can be described by a state-space model.

### A. Observation model

We have considered a MIMO-OFDM system with  $N$  transmit antennas and  $M$  receive antennas. The number of subcarriers is  $K$ . At each transmit antenna, an OFDM modulator is used. The OFDM symbol that is transmitted from the  $n$ th antenna at discrete time index  $t$  is denoted by the  $K$ -dimensional column vector  $\tilde{\mathbf{d}}_{n;t} \triangleq [d_{1,n;t}, \dots, d_{K,n;t}]^T$ . This vector is processed by an IFFT, and a cyclic prefix of length

$\mu$  is added. We assume that  $\mu \geq L_{max} - 1$ , where  $L_{max}$  is the maximum length of all multipath channels. After removing the cyclic prefix at the  $m$ th receive antenna and performing a FFT, we obtain the  $K$ -dimensional column vector  $\tilde{\mathbf{y}}_{m;t} \triangleq [y_{1,m;t}, \dots, y_{K,m;t}]^T$ , which can be written as [4]

$$\tilde{\mathbf{y}}_{m;t} = \sum_{n=1}^N \tilde{\mathbf{D}}_{n;t} \mathbf{T} \mathbf{h}_{m,n;t} + \tilde{\mathbf{z}}_{m;t} \quad (1)$$

where  $\tilde{\mathbf{D}}_{n;t} = \text{diag}(\tilde{\mathbf{d}}_{n;t})$ ,  $\mathbf{h}_{m,n;t} \triangleq [a_{m,n;t}^{(1)}, \dots, a_{m,n;t}^{(L_{max})}]^T$  is a  $L_{max}$ -dimensional column vector representing the  $L_{max}$  length channel impulse response from the  $n$ th transmit antenna to the  $m$ th receive antenna,  $\tilde{\mathbf{z}}_{m;t} \triangleq [z_{1,m;t}, \dots, z_{K,m;t}]^T$  is the AWGN having power spectral density  $N_0$ , i.e.  $\tilde{\mathbf{z}}_{m;t} \sim \mathcal{N}_c(\tilde{\mathbf{z}}_{m;t}; \mathbf{0}_{K \times 1}, N_0 \mathbf{I}_K)$  and  $\mathbf{T}$  is the  $K \times L_{max}$  Fourier matrix whose  $(k, \ell)$ th entry is defined as,  $\{\mathbf{T}\}_{k,\ell} = e^{-j2\pi(\ell-1)\frac{k-1}{K}}$  for  $k = 1, \dots, K$  and  $\ell = 1, \dots, L_{max}$ .

We assume that  $P$  pilot symbols are uniformly interleaved with data to enable channel estimation and  $P \geq L_{max}N$ . Denote  $k_p$ , for  $p = 1, \dots, P$ , the  $p$ th pilot subcarrier index. Let us collect the signal observed onto the pilot subcarriers of the  $m$ th receive antenna at time  $t$  in the  $P$ -dimensional column vector  $\mathbf{y}_{m;t} \triangleq [y_{k_1;m;t}, \dots, y_{k_P;m;t}]^T$ . Analogously, we collect the observations from the information subcarriers in  $(K - P)$ -dimensional column vector  $\mathbf{y}_{m;t}^I = [y_{k_1^I;m;t}, \dots, y_{k_{K-P}^I;m;t}]^T$  where  $k_i^I$ , for  $i = 1, \dots, K - P$ , is the  $i$ th information subcarrier index. Thus, from (1), it follows that

$$\text{Pilot : } \mathbf{y}_{m;t} = \sum_{n=1}^N \mathbf{D}_{n;t} \mathbf{F} \mathbf{h}_{m,n;t} + \mathbf{z}_{m;t} \quad (2)$$

$$\text{Information : } \mathbf{y}_{m;t}^I = \sum_{n=1}^N \mathbf{D}_{n;t}^I \mathbf{F}^I \mathbf{h}_{m,n;t} + \mathbf{z}_{m;t}^I \quad (3)$$

where  $\mathbf{D}_{n;t} = \text{diag}([d_{k_1,n;t}, \dots, d_{k_P,n;t}]^T)$ ,  $\mathbf{D}_{n;t}^I = \text{diag}([d_{k_1^I,n;t}, \dots, d_{k_{K-P}^I,n;t}]^T)$ ,  $\mathbf{z}_{m;t} = [z_{k_1,m;t}, \dots, z_{k_P,m;t}]^T$ ,  $\mathbf{z}_{m;t}^I = [z_{k_1^I,m;t}, \dots, z_{k_{K-P}^I,m;t}]^T$ ,  $\mathbf{F}$  is a  $P \times L_{max}$  matrix whose  $(p, \ell)$ th entry is defined as,  $\{\mathbf{F}\}_{p,\ell} = e^{-j2\pi(\ell-1)\frac{k_p-1}{K}}$  and  $\mathbf{F}^I$  is a  $(K - P) \times L_{max}$  matrix whose  $(i, \ell)$ th entry is defined as  $\{\mathbf{F}^I\}_{i,\ell} = e^{-j2\pi(\ell-1)\frac{k_i^I-1}{K}}$ . An alternative representation of (2) can be given as

$$\mathbf{y}_{m;t} = \mathbf{A}_t \mathbf{h}_{m;t} + \mathbf{z}_{m;t} \quad (4)$$

where  $\mathbf{A}_t = [\mathbf{D}_{1;t} \mathbf{F} \cdots \mathbf{D}_{N;t} \mathbf{F}]$  is a  $P \times L_{max}N$  matrix and  $\mathbf{h}_{m;t} = [\mathbf{h}_{m,1;t}^T \cdots \mathbf{h}_{m,N;t}^T]^T$  is a  $L_{max}N$ -dimensional column vector.

### B. Dynamic channel model

Denote by  $\mathcal{H}_{m,n;t}$  the random set representing the multipath channel experienced by the transmission from the  $n$ th antenna to the  $m$ th antenna at discrete time  $t$ . For the sake of simplicity we have assumed a Tapped Delay Line (TDL) channel model where the tap delays are equispaced. The same theory could accommodate the case of unknown and time-varying multipath delays. We assume that the maximum number of admissible path is  $L_{max}$ . Clearly,  $\mathcal{H}_{m,n;t}$  can be represented as

$$\mathcal{H}_{m,n;t} = \bigcup_{\ell=1}^{L_{max}} \mathcal{H}_{m,n;t}^{(\ell)}$$

where  $\mathcal{H}_{m,n;t}^{(\ell)}$  denotes a singleton-or-empty set. Indeed,  $\mathcal{H}_{m,n;t}^{(\ell)} = [m, n, \ell, a_{m,n;t}^{(\ell)}]^T$ , if the  $\ell$ th path from the  $n$ th transmitting antenna to the  $m$ th receive antenna is active at time  $t$ , otherwise  $\mathcal{H}_{m,n;t}^{(\ell)} = \emptyset$ . Notice that  $\mathcal{H}_{m,n;t}$  is a random set in the hybrid space  $m \times n \times \{1, \dots, L_{max}\} \times \mathbb{C}$ .

Denoting by  $\mathcal{S}_{m,n;t}$  the surviving path from epoch  $t - 1$  to epoch  $t$ , from the  $n$ th transmission antenna to the  $m$ th receive antenna, and by  $\mathcal{B}_{m,n;t}$  the newly born paths, we have  $\mathcal{H}_{m,n;t} = \mathcal{S}_{m,n;t} \cup \mathcal{B}_{m,n;t}$ . Moreover, we define the random sets  $\pi(\mathcal{H}_{m,n;t})$  and  $\pi'(\mathcal{H}_{m,n;t})$ , denoting the projection of  $\mathcal{H}_{m,n;t}$  onto  $m \times n \times \{1, \dots, L_{max}\}$  and onto  $\mathbb{C}$ , respectively:

$$\pi(\mathcal{H}_{m,n;t}) = \bigcup_{\ell: \mathcal{H}_{m,n;t}^{(\ell)} \neq \emptyset} \{m, n, \ell\} \quad (5)$$

$$\pi'(\mathcal{H}_{m,n;t}) = \bigcup_{\ell \in \pi(\mathcal{H}_{m,n;t})} \{a_{m,n;t}^{(\ell)}\} \quad (6)$$

The constraints

$$\begin{aligned} \pi(\mathcal{H}_{m,n;t-1}) \cap \pi(\mathcal{B}_{m,n;t}) &= \emptyset \\ \pi(\mathcal{S}_{m,n;t}) &\subseteq \pi(\mathcal{H}_{m,n;t-1}) \end{aligned}$$

reflect the facts that no component being active at time  $t-1$  can migrate to the set of new paths, and that the paths surviving at epoch  $t$  are a subset of those active at epoch  $t-1$ .

The conditional density of the newly born path, assuming that new paths arise independently, can be expressed as

$$\begin{aligned} f_{\mathcal{B}_{m,n;t} | \mathcal{H}_{m,n;t-1}}(\mathcal{B}_{m,n;t} | \mathcal{H}_{m,n;t-1}) &= P_{birth}^{|\mathcal{B}_{m,n;t}|} \\ &\times (1 - P_{birth})^{L_{max} - |\mathcal{H}_{m,n;t-1}| - |\mathcal{B}_{m,n;t}|} \\ &\times \prod_{\ell \in \pi(\mathcal{B}_{m,n;t})} f_{a_{m,n;t}^{(\ell)}}(a_{m,n;t}^{(\ell)}) \end{aligned} \quad (7)$$

where  $P_{birth}$  the probability that a new path arises and  $f_{a_{m,n;t}^{(\ell)}}(a_{m,n;t}^{(\ell)})$  is the probability density function of the  $\ell$ th path gain from the  $n$ th transmission antenna to the  $m$ th receiver antenna at epoch  $t$ .

Similarly, we have

$$\mathcal{S}_{m,n;t} = \bigcup_{\ell \in \pi(\mathcal{H}_{m,n;t-1})} \mathcal{S}_{m,n;t}^{(\ell)} \quad (8)$$

with

$$\mathcal{S}_{m,n;t}^{(\ell)} = \begin{cases} \emptyset & \text{with probability } P_{death} \\ \{\mathbf{h}_{m,n;t}^{(\ell)}\} & \text{with probability } 1 - P_{death} \end{cases}$$

where  $P_{death}$  is the probability that an active path disappears. It follows that the conditional density of the random set  $\mathcal{S}_{m,n;t}$  given  $\mathcal{H}_{m,n;t-1}$  can be given as

$$\begin{aligned} f_{\mathcal{S}_{m,n;t} | \mathcal{H}_{m,n;t-1}}(\mathcal{S}_{m,n;t} | \mathcal{H}_{m,n;t-1}) &= P_{death}^{|\mathcal{H}_{m,n;t-1}| - |\mathcal{S}_{m,n;t}|} \\ &\times (1 - P_{death})^{|\mathcal{S}_{m,n;t}|} \prod_{\ell \in \pi(\mathcal{S}_{m,n;t})} f_{a_{m,n;t}^{(\ell)} | a_{m,n;t-1}^{(\ell)}}(a_{m,n;t}^{(\ell)} | a_{m,n;t-1}^{(\ell)}) \end{aligned} \quad (9)$$

with  $\mathcal{S}_{m,n;t} \subseteq \mathcal{H}_{m,n;t-1}$  and  $f_{a_{m,n;t}^{(\ell)} | a_{m,n;t-1}^{(\ell)}}(a_{m,n;t}^{(\ell)} | a_{m,n;t-1}^{(\ell)})$  the transition density describing the evolution of the gains of

the surviving paths.

The transition density  $f_{\mathcal{H}_{m,n;t}|\mathcal{H}_{m,n;t-1}}(\mathcal{H}_{m,n;t} | \mathcal{H}_{m,n;t-1})$  can in turn be determined through the generalized convolution formula, yielding:

$$\begin{aligned} f_{\mathcal{H}_{m,n;t}|\mathcal{H}_{m,n;t-1}}(\mathcal{H}_{m,n;t} | \mathcal{H}_{m,n;t-1}) \\ = f_{\mathcal{S}_{m,n;t}}(\mathcal{H}_{m,n;t} \cap \mathcal{H}_{m,n;t-1} | \mathcal{H}_{m,n;t-1}) \\ \times f_{\mathcal{B}_{m,n;t}}(\mathcal{H}_{m,n;t} \setminus (\mathcal{H}_{m,n;t} \cap \mathcal{H}_{m,n;t-1}) | \mathcal{H}_{m,n;t-1}) \end{aligned}$$

Let  $\mathcal{H}_{m;t}$  be

$$\mathcal{H}_{m;t} = \bigcup_{n=1}^N \mathcal{H}_{m,n;t} \quad (10)$$

We may map the random set  $\mathcal{H}_{m,n;t}$  into a column vector of dimension  $L_{max}$  whose  $\ell$ th entry is defined as

$$\{\mathbf{h}_{m,n;t}\}_{\ell} = \begin{cases} a_{m,n;t}^{(\ell)} & \text{if } \ell \in \pi(\mathcal{H}_{m,n;t}) \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

The mapping in eq. (11) makes the RFS channel model compatible with the model in (1). Analogously, the mapping between  $\mathcal{H}_{m;t}$  and  $\mathbf{h}_{m;t}$  follows straightforwardly.

### C. Problem formulation

The problem under consideration can now be formulated as follows: Given the observation model in (4) and the dynamic channel model in Section II-B, the goal of this paper is to derive a causal estimator for the multipath channel for MIMO-OFDM systems. Successively, given this channel estimate, standard data detection algorithms can be easily implemented. While channel estimation is addressed in Secs. III-IV, next we describe data detection once the channel state information (CSI) has been obtained.

Let  $\hat{\mathbf{h}}_{m,n;t} = [\hat{a}_{m,n;t}^{(1)}, \dots, \hat{a}_{m,n;t}^{(L_{max})}]^T$  for  $m = 1, \dots, M$  and  $n = 1, \dots, N$  the channel estimates obtained from the pilot subcarriers. It is possible to prove that the Maximum Likelihood (ML) detection of the information symbols can be disjoint into  $K - P$  problems, i.e. one for each information subcarriers, as

$$\hat{\mathbf{s}}_{i;t} = \arg \min_{\mathbf{s}_{i;t}} \|\mathbf{r}_{i;t} - \mathbf{C}_{i;t} \mathbf{s}_{i;t}\|^2, \quad \text{for } i = 1, \dots, K - P$$

where

$$\begin{aligned} \mathbf{r}_{i;t} &= [y_{k_i^I,1;t}, \dots, y_{k_i^I,M;t}]^T \\ \mathbf{s}_{i;t} &= [d_{k_i^I,1;t}, \dots, d_{k_i^I,N;t}]^T \\ \hat{H}_{k_i^I,m,n;t} &= \sum_{\ell=1}^{L_{max}} \hat{a}_{m,n;t}^{(\ell)} e^{-j2\pi(k_i^I-1)\frac{\ell-1}{K}} \\ \mathbf{C}_{i;t} &= \begin{bmatrix} \hat{H}_{k_i^I,1,1;t} & \dots & \hat{H}_{k_i^I,1,N;t} \\ \vdots & \ddots & \vdots \\ \hat{H}_{k_i^I,M,1;t} & \dots & \hat{H}_{k_i^I,M,N;t} \end{bmatrix} \end{aligned}$$

## III. BAYES RECURSIVE FILTER AND ITS IMPLEMENTATION

In the next section, we will consider the problem of sequential estimation of MIMO-OFDM multipath channel within a Bayesian framework. Next, due to the lack of a closed form for the optimal estimator, we present a SMC implementation of the Bayes Recursive filter.

### A. Bayes Recursive filter

Let us consider the observation model in eq. (4) relative to the  $m$ th receive antenna and the dynamic channel model in Sec. II-B. Consider the variable  $\mathbf{h}_{m;t}$ , depending by  $\mathcal{H}_{m;t}$ , to be the state that fully characterize the wireless channel experienced by the  $m$ th receive antenna. Clearly, what we will develop further will be repeated for all the  $M$  receive antennas.

The best description of  $\mathcal{H}_{m;t}$  based on the observation  $\mathbf{y}_{m;1:t}$  relies upon the evaluation of the so called filtering distribution, i.e.  $f(\mathcal{H}_{m;t} | \mathbf{y}_{m;1:t})$ . The basic step to achieve it is the evaluation of the one-step ahead predicted distribution via Chapman-Kolmogorov equation given the filtering distribution at time  $t - 1$ , i.e.

$$f(\mathcal{H}_{m;t} | \mathbf{y}_{m;1:t-1}) = \int f(\mathcal{H}_{m;t} | \mathcal{H}_{m;t-1}) f(\mathcal{H}_{m;t-1} | \mathbf{y}_{m;1:t-1}) \delta \mathcal{H}_{m;t-1} \quad (12)$$

where  $f(\mathcal{H}_{m;t} | \mathcal{H}_{m;t-1})$  is the state transition distribution. The symbol  $\delta \mathcal{H}_{m;t-1}$  emphasize that the integral in (12) is a set integral. After the arrival of the observation  $\mathbf{y}_{m;t}$  at time  $t$ , the filtering distribution can be updated via Bayes rule as

$$f(\mathcal{H}_{m;t} | \mathbf{y}_{m;1:t}) = \frac{f(\mathbf{y}_{m;t} | \mathcal{H}_{m;t}) f(\mathcal{H}_{m;t} | \mathbf{y}_{m;1:t-1})}{f(\mathbf{y}_{m;t} | \mathbf{y}_{m;1:t-1})} \quad (13)$$

where  $f(\mathbf{y}_{m;t} | \mathcal{H}_{m;t})$  is the likelihood distribution. Equations (12) and (13) are called Bayes Recursion (BR) and define the so-called Bayes Recursive filter. The analytical solution to the above equations is intractable and approximated techniques are necessary. Next, we address the approximation of the BR via SMC.

### B. Sequential Monte Carlo approximation of the Bayes Recursive filter

Among all techniques, SMC seems particularly attractive for a RFS implementation of the BR providing a consistent approximation of the Bayes Recursive filter [5]. The SMC approach relies on a sample-based construction of distributions involved in (12) and (13). Multiple particles of the state variable are generated, each one associated with a weight which characterizes the belief that the channel is in that state. The filtering distribution is approximated by a set of particles as

$$f(\mathcal{H}_{m;t} | \mathbf{y}_{1:t}) \approx \sum_{r=1}^R w_t^{(r)} m_{\mathcal{H}_{m;t}}(\mathcal{H}_{m;t}^{(r)}) \quad (14)$$

where  $m_{\mathcal{X}}(\mathcal{Y})$  is the "0-1" measure, defined as follows

$$\int_{\mathcal{C}} m_{\mathcal{X}}(\mathcal{Y}) \delta \mathcal{X} = \begin{cases} 1, & \text{if } \mathcal{Y} \subseteq \mathcal{C} \\ 0, & \text{otherwise} \end{cases}$$

In (14),  $\mathcal{H}_{m;t}^{(r)}$  is the  $r$ th set "particle",  $w_t^{(r)}$  is its "weight", and  $R$  is the total number of particles. Asymptotic convergence properties of the SMC RFS filter has been proved in [8], showing that, for sufficient large  $R$ , the mean-square approximation error of the SMC RFS filter is inversely proportional to  $R^\alpha$ , for some constant  $0 < \alpha \leq 1$ , while the implementation complexity is approximately linear with  $R$ .

Suppose to have a particle representation of the filtering distribution at time  $t - 1$ , i.e.

$$f(\mathcal{H}_{m;t-1} | \mathbf{y}_{1:t-1}) \approx \sum_{r=1}^R w_{t-1}^{(r)} m_{\mathcal{H}_{m;t-1}}(\mathcal{H}_{m;t-1}^{(r)})$$

The Sequential Monte Carlo algorithm proceeds to approximate the posterior distribution at time  $t$  by a new set of weighted particles  $\{w_t^{(r)}, \mathcal{H}_{m;t}^{(r)}\}_{r=1}^R$ .

#### IV. RAO-BLACKWELLIZED PARTICLE FILTER

A major drawback of SMC is that its efficiency decrease with the dimension of the state space requiring a large number of particles. Hence, in some application its implementation can become prohibitive. However, the complexity can be reduced by a procedure called Rao-Blackwellization which is a technique for improving particle filtering by analytically marginalizing some of the variables (linear and Gaussian) from the joint filtering distribution. The linear part of the system model is then estimated by a Kalman filter (KF) while the nonlinear part is estimated by a SMC. RBPF can results in a decrease of the variance of the SMC estimates and of the complexity. On the other hand, Rao-Blackwellization restrict the form of the system model, requiring a conditionally linear Gaussian (CLG) model. Indeed, consider the dynamic channel model in Sec. II-B under the assumption that the distributions  $f_{a_{m,n;t}}^{(\ell)}(a_{m,n;t}^{(\ell)})$  and  $f_{a_{m,n;t}|a_{m,n;t-1}}^{(\ell)}(a_{m,n;t}^{(\ell)}|a_{m,n;t-1}^{(\ell)})$  in (7) and (9), respectively, are Gaussian.

Instead of directly evaluating the filtering distribution, i.e.  $f(\mathcal{H}_{m;t}|\mathbf{y}_{m;1:t})$ , we address the evaluation of the so-called joint posterior distribution, i.e.  $f(\mathcal{H}_{m;1:t}|\mathbf{y}_{m;1:t})$ , from which the filtering distribution can be easily obtained. The joint posterior distribution satisfies the following recursion

$$\begin{aligned} f(\mathcal{H}_{m;t}|\mathbf{y}_{m;1:t}) &= \frac{f(\mathbf{y}_{m;t}|\mathcal{H}_{m;t})f(\mathcal{H}_{m;t}|\mathcal{H}_{m;t-1})}{f(\mathbf{y}_{m;t}|\mathbf{y}_{m;1:t-1})} \\ &\times f(\mathcal{H}_{m;1:t-1}|\mathbf{y}_{m;1:t-1}) \end{aligned}$$

Let us divide the random set  $\mathcal{H}_{m;t}$  in two parts, according to the definition in eqs. (5) and (6), i.e.

$$\mathcal{H}_{m;t} = \{\pi(\mathcal{H}_{m;t}), \pi'(\mathcal{H}_{m;t})\}$$

Clearly,

$$\begin{aligned} f(\mathcal{H}_{m;t}|\mathcal{H}_{m;t-1}) &= f(\pi(\mathcal{H}_{m;t})|\pi(\mathcal{H}_{m;t-1})) \\ &\times f(\pi'(\mathcal{H}_{m;t})|\pi(\mathcal{H}_{m;t-1}), \pi'(\mathcal{H}_{m;t-1})) \end{aligned}$$

Notice that  $\pi(\mathcal{H}_{m;t})$  is a random finite set defined on a countable space, while  $\pi'(\mathcal{H}_{m;t})$ , conditioned to  $\pi(\mathcal{H}_{m;t})$ , is a random vector whose dimension is dictated by the cardinality of  $\pi(\mathcal{H}_{m;t})$ . Moreover, due to the CGL hypothesis, the conditional posterior distribution  $f(\pi'(\mathcal{H}_{m;t})|\pi(\mathcal{H}_{m;t}), \mathbf{y}_{m;1:t})$  is analytically tractable through Kalman filtering. Hence, we can easily marginalize out  $\pi'(\mathcal{H}_{m;t})$  from the posterior distribution, and seek for an estimator of  $f(\pi(\mathcal{H}_{m;t})|\mathbf{y}_{m;1:t})$ , which lies in a space of reduced dimension. Suppose to have at time  $t-1$  an approximation of the posterior distribution  $f(\pi(\mathcal{H}_{m;1:t-1})|\mathbf{y}_{m;1:t-1})$  in term of particles, i.e.

$$\begin{aligned} f(\pi(\mathcal{H}_{m;1:t-1})|\mathbf{y}_{m;1:t-1}) &= \sum_{r=1}^R w_{t-1}^{(r)} m_{\pi(\mathcal{H}_{m;1:t-1})}(\pi(\mathcal{H}_{m;1:t-1})^{(r)}) \end{aligned}$$

When the measurement  $\mathbf{y}_{m;t}$  is available, the RBPF performs the operations described in Algorithm 1 in order to determine an approximation in term of particles of  $f(\pi(\mathcal{H}_{m;1:t})|\mathbf{y}_{m;1:t})$ .

#### Algorithm 1 Generic RBPF

- 
- 1: Given  $\{w_{t-1}^{(r)}, \pi(\mathcal{H}_{m;1:t-1})^{(r)}\}_{r=1}^R$
  - for**  $r = 1$  to  $R$
  - 2:   Sample  $\pi(\mathcal{H}_{m;t})^{(r)} \sim q(\cdot|\pi(\mathcal{H}_{m;1:t-1})^{(r)}, \mathbf{y}_{m;1:t})$
  - 3:   Set  $\pi(\mathcal{H}_{m;1:t})^{(r)} \triangleq \{\pi(\mathcal{H}_{m;1:t-1})^{(r)}, \pi(\mathcal{H}_{m;t})^{(r)}\}$
  - 4:   Update the weight  $\tilde{w}_t^{(r)} = \frac{w_{t-1}^{(r)} f(\mathbf{y}_{m;t}|\mathbf{y}_{m;1:t-1}, \pi(\mathcal{H}_{m;1:t})^{(r)}) f(\pi(\mathcal{H}_{m;t})^{(r)}|\pi(\mathcal{H}_{m;1:t-1})^{(r)})}{q(\pi(\mathcal{H}_{m;t})^{(r)}|\pi(\mathcal{H}_{m;1:t-1})^{(r)}, \mathbf{y}_{m;1:t})}$
  - end**
  - Normalization
  - for**  $r = 1$  to  $R$
  - 5:    $w_t^{(r)} = \frac{\tilde{w}_t^{(r)}}{\sum_{r'=1}^R \tilde{w}_t^{(r')}} \pi(\mathcal{H}_{m;t})^{(r)}$
  - end**
  - 6: Resampling with replacement
- 

Resampling is necessary because the variance of the particle weights can only increase in time which will ends up in the degeneracy problem. In practice, it has been observed that, after a few iterations, one of the normalized importance weights tends to 1, while the remaining weights tend to zero. The basic idea here is to eliminate particles that have small weights, and to concentrate on particles with large weights whenever the degeneracy phenomenon becomes relevant (see [9] for a review of resampling techniques).

Notice that from the particle approximation of the posterior distribution, i.e.

$$f(\pi(\mathcal{H}_{m;1:t})|\mathbf{y}_{m;1:t}) \approx \sum_{r=1}^R w_t^{(r)} m_{\pi(\mathcal{H}_{m;1:t})}(\pi(\mathcal{H}_{m;1:t})^{(r)})$$

it is straightforward to obtain the full joint posterior distribution, i.e.

$$\begin{aligned} f(\pi(\mathcal{H}_{m;1:t}), \pi'(\mathcal{H}_{m;1:t})|\mathbf{y}_{m;1:t}) &= \sum_{r=1}^R w_t^{(r)} m_{\pi(\mathcal{H}_{m;1:t}), \pi'(\mathcal{H}_{m;1:t})}(\pi(\mathcal{H}_{m;1:t})^{(r)}, \pi'(\mathcal{H}_{m;1:t})^{(r)}) \end{aligned}$$

On-line channel estimation can be performed by maximizing the discrete parameter filtering distribution as

$$\widehat{\pi(\mathcal{H}_{m;t})} = \arg \max_{\pi(\mathcal{H}_{m;t})} f(\pi(\mathcal{H}_{m;t})|\mathbf{y}_{m;1:t}) \quad (15)$$

Clearly,  $\widehat{\pi(\mathcal{H}_{m;t})}$  will carry the information about the paths that are active at time  $t$ . Given  $\widehat{\pi(\mathcal{H}_{m;t})}$  the estimation of the continuous part will be performed as

$$\widehat{\pi'(\mathcal{H}_{m;t})} = E[\pi'(\mathcal{H}_{m;t})|\mathbf{y}_{m;1:t}, \pi(\mathcal{H}_{m;1:t-1}), \widehat{\pi(\mathcal{H}_{m;t})}]$$

which can be evaluated by  $R$  KFs.

An important implementation issue is the design of the importance distribution  $q(\pi(\mathcal{H}_{m;t})|\pi(\mathcal{H}_{m;1:t-1})^{(r)}, \mathbf{y}_{m;1:t})$  in Algorithm 1 which highly affect the SMC algorithm performance. The optimal importance distribution, in sense of minimizing the variance of the particle weight, is given as



$$\begin{aligned}
& q(\pi(\mathcal{H}_{m;t})|\pi(\mathcal{H}_{m;1:t-1})^{(r)}, \mathbf{y}_{m;1:t}) \\
&= f(\pi(\mathcal{H}_{m;t})|\pi(\mathcal{H}_{m;1:t-1})^{(r)}, \mathbf{y}_{m;1:t}) \\
&= \frac{f(\mathbf{y}_{m;t}|\mathbf{y}_{m;1:t-1}, \pi(\mathcal{H}_{m;1:t-1})^{(r)}, \pi(\mathcal{H}_{m;t}))}{f(\mathbf{y}_{m;t}|\mathbf{y}_{m;1:t-1}, \pi(\mathcal{H}_{m;1:t-1})^{(r)})} \\
&\quad \times f(\pi(\mathcal{H}_{m;t})|\pi(\mathcal{H}_{m;1:t-1})^{(r)})
\end{aligned}$$

The weights are updated as

$$\tilde{w}_t^{(r)} = w_{t-1}^{(r)} f(\mathbf{y}_{m;t}|\mathbf{y}_{m;1:t-1}, \pi(\mathcal{H}_{m;1:t-1})^{(r)})$$

Clearly, the calculation of above distribution requires an additional integration stage as

$$\begin{aligned}
& f(\mathbf{y}_{m;t}|\mathbf{y}_{m;1:t-1}, \pi(\mathcal{H}_{m;1:t-1})^{(r)}) \\
&= \int f(\mathbf{y}_{m;t}|\mathbf{y}_{m;1:t-1}, \pi(\mathcal{H}_{m;1:t-1})^{(r)}, \pi(\mathcal{H}_{m;t})) \\
&\quad \times f(\pi(\mathcal{H}_{m;t})|\pi(\mathcal{H}_{m;1:t-1})^{(r)}) \delta\pi(\mathcal{H}_{m;t})
\end{aligned}$$

The first term in the integral can be evaluated by a Kalman filter for each value of  $\pi(\mathcal{H}_{m;t})$ . Moreover, since  $\pi(\mathcal{H}_{m;t})$  is a countable set, the integral becomes a summation.

## V. SIMULATION RESULTS

We consider a MIMO-OFDM system employing  $K = 64$  subcarriers, with  $N = 2$ ,  $M = 3$ ,  $P = 8$ , and a frequency spacing  $D_f = 8$  between two pilot subcarriers. The average energy per symbol,  $\sigma_s^2 = E[|d_{k,n;t}|^2]$ , is uniform, and a 4-QAM modulation scheme is assumed. As to the channel, we assume a uniform multipath delay profile, a multipath spread smaller than the guard time, and uncorrelated path gains. The overall channel energy is normalized to one. Consider first the transmission of  $T = 50$  OFDM symbols through a channel with  $L_{max} = 4$ ,  $P_{birth} = 0.05$ , and  $P_{death} = 0.05$ . The path gains are assumed to follow a Gauss-Markov model as

$$\begin{aligned}
f(a_{m,n;1}^{(\ell)}) &= \mathcal{N}_c(a_{m,n;1}^{(\ell)}; 0, \sigma_h^2) \\
f(a_{m,n;t}^{(\ell)}|a_{m,n;t-1}^{(\ell)}) &= \mathcal{N}_c(a_{m,n;t}^{(\ell)}; \lambda a_{m,n;t-1}^{(\ell)}, (1 - \lambda^2)\sigma_h^2)
\end{aligned}$$

with  $\sigma_h^2$  the average energy of one path, and  $\lambda = 0.999$ .

We have simulated a RBPF with optimal importance function and  $R = 50$ , and we have compared it to the Least-Square channel estimator in [4]. In fig. 1, we have plotted the frequency-domain mean-square error (FMSE) defined as

$$FMSE = \frac{\sum_{i=1}^{K-P} \sum_{m=1}^M \sum_{n=1}^N \sum_{t=1}^T |H_{k_i^I, m, n, t} - \hat{H}_{k_i^I, m, n, t}|^2}{MNT(K-P)}$$

versus the signal to noise ratio (SNR), defined as  $SNR = \frac{\sigma_s^2}{N_0}$ . The pilot symbols are assumed to be phase shift orthogonal (see [4]). Fig. 1 shows that the RBPF achieves a gain of about 6-dB at a FMSE of  $10^{-2}$  with respect to the standard LS estimator. Fig. 2 shows the Bit Error Rate (BER) using the ML detector in (12) based on the channel estimate of the LS and the RBPF estimators, and a ML detector assuming Complete Channel State Information (CCSI). Using RBPF, we can achieve a 3-dB gain at a BER of  $10^{-4}$  with respect to LS, while approaching the CCSI performance.

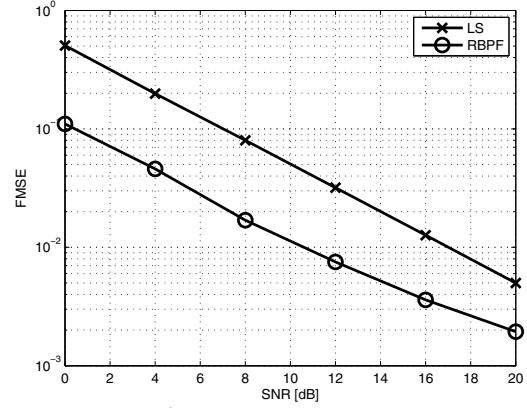


Fig. 1. FMSE versus SNR.

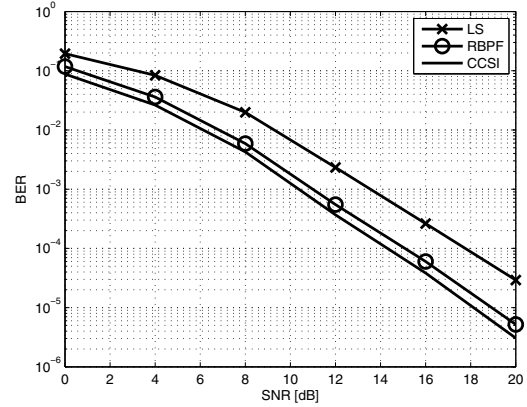


Fig. 2. BER versus SNR.

## ACKNOWLEDGMENT

The work of Ezio Biglieri was supported by the Spanish Ministry of Education and Science under Project TEC2006-01428/TCM, and by the STREP project No. IST-026905 (MASCOT) within the 6th framework program of the European Commission.

## REFERENCES

- [1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311-335, 1998.
- [2] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. on Comm.*, vol. 19, pp. 628-634, Oct. 1971.
- [3] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, pp. 225-234, Feb. 2002.
- [4] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for mimo ofdm systems in mobile wireless channels," *IEEE Trans. Signal Processing*, vol. 5, pp. 1615-1624, June 2003.
- [5] R. Mahler, *Statistical Multisource-Multitarget Information Fusion*, Artech House, 2007.
- [6] M.S. Arulampalam, S. Maskell, N. Gordon, T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. on Signal Processing* Vol. 50, pp. 174-188, Feb. 2002.
- [7] M. Vihola, "Rao-Blackwellised Particle Filtering in Random Set Multitarget Tracking," *IEEE Transactions on Aerospace and Electronic Systems*, Volume 43, Issue 2, April 2007, pp. 689-705.
- [8] B.-N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo Methods for Bayesian multi-target filtering with random finite sets," *IEEE Trans. Aerosp. Electron. Syst.*, Vol. 41, No. 4, pp. 1224-1245, October 2005.
- [9] M. Bolic, P. M. Djuric, and S. Hong, "Resampling algorithms and architectures for distributed particle filters," *IEEE Trans. Signal Proc.*, vol. 53, no. 7, pp. 2442-2450, July 2005.