

Monte Carlo Methods for Channel, Phase Noise and Frequency Offset Estimation with Unknown Noise Variances in OFDM Systems

François Septier, Yves Delignon, Atika Menhaj-Rivenq, Christelle Garnier

► **To cite this version:**

François Septier, Yves Delignon, Atika Menhaj-Rivenq, Christelle Garnier. Monte Carlo Methods for Channel, Phase Noise and Frequency Offset Estimation with Unknown Noise Variances in OFDM Systems. IEEE Transactions on Signal Processing, Institute of Electrical and Electronics Engineers, 2008, 56 (8), pp.3613-3626. <10.1109/TSP.2008.919629>. <hal-00813272>

HAL Id: hal-00813272

<https://hal-imt.archives-ouvertes.fr/hal-00813272>

Submitted on 16 Apr 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Monte Carlo Methods for Channel, Phase Noise and Frequency Offset Estimation with Unknown Noise Variances in OFDM Systems

François Septier, Yves Delignon, *Member, IEEE*, Atika Menhaj-Rivenq, Christelle Garnier

Abstract

In this paper, we address the problem of OFDM channel estimation in the presence of phase noise (PHN) and carrier frequency offset (CFO). In OFDM systems, PHN and CFO cause two effects: the common phase error (CPE) and the intercarrier interference (ICI) which severely degrade the accuracy of the channel estimate. In literature, several algorithms have been proposed to solve this problem. Nevertheless, in all these existing schemes, both the PHN and the Additive White Gaussian Noise (AWGN) powers are assumed to be known. Because no a priori knowledge of PHN and AWGN powers is available at the receiver, we propose different strategies for the estimation of channel impulse response (CIR), CFO, PHN and also the PHN and the AWGN powers. Based on Monte Carlo methods, the proposed approaches estimate these many unknowns in the time domain from a training OFDM symbol using either off-line or on-line estimators. In the on-line case, we propose Sequential Monte Carlo algorithms and especially an original maximization step of the joint a posteriori probability density function for the unknown parameters. Simulation results are provided to illustrate the efficiency of the proposed algorithms in terms of mean square error (MSE) on channel, phase distortions and also noise power estimation.

Index Terms

Orthogonal frequency-division multiplexing (OFDM), channel estimation, phase noise, carrier frequency offset, sequential Monte-Carlo methods, Rao-blackwellization, optimal importance function, on-

F. Septier and A. Menhaj-Rivenq are with the IEMN-DOAE UMR 8520, university of Valenciennes U.V.H.C. Le Mont Houy, 59313 Valenciennes Cedex 9, France (e-mail: francois.septier@telecom-lille1.eu; atika.menhaj@univ-valenciennes.fr).

Y. Delignon and C. Garnier are with the Department of Communication Systems, GET/INT/TELECOM Lille 1, 59650 Villeneuve d'Ascq, France (e-mail: yves.delignon@telecom-lille1.eu; christelle.garnier@telecom-lille1.eu).

line parameter estimation.

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is known to be an efficient technique for high-rate transmission that can overcome the intersymbol interference (ISI) resulting from the time dispersion of multipath fading channels. It has been adopted as the transmission method of many standards in wireline and wireless communications, such as digital subscriber lines (DSL), digital audio and video broadcasting (DAB/DVB), wireless area networks (IEEE 802.11) and broadband wireless access (IEEE 802.16).

Unfortunately, OFDM systems are very sensitive to phase noise (PHN) and to carrier frequency offset (CFO) respectively caused by oscillator imperfections and Doppler shifts [1]–[3]. Indeed, these phase distortions destroy the orthogonality of the OFDM subcarriers and lead both to rotation of every subcarrier by a random phase, called common phase error (CPE), and to intercarrier interference (ICI).

Many approaches have been proposed to estimate and compensate PHN [4]–[6] or both CFO and PHN [7]. Nevertheless, the channel impulse response (CIR) is assumed known prior to phase distortion suppression. Only two papers [8] and [9] deal with channel estimation in the presence of both PHN and CFO. In [8], the authors mainly focus on channel and CFO estimation and in [9] a maximum *a posteriori* estimator of CIR, PHN and CFO assuming prior knowledge of SNR and PHN statistics has been proposed.

Unlike previous publications, we consider in this paper the channel estimation problem in the presence of PHN and CFO with no *a priori* knowledge of AWGN and PHN powers. Their estimate must be considered with care since both AWGN and PHN powers are required for optimal data detection step of OFDM symbols. To estimate these many unknowns from a single training OFDM symbol, we propose an off-line and an on-line estimator based on Monte Carlo (MC) methods. Firstly, we propose a stochastic version of the Expectation-Maximization (EM) algorithm to perform an off-line estimation of the unknown static parameters. Once these parameters have been estimated, the maximum a posteriori estimator proposed in [9] can be used in order to obtain the other estimates of interest. Secondly, we propose two different schemes based on Sequential Monte Carlo (SMC) methods for the joint and on-line estimation of all the unknown states. Indeed, the sequential nature of SMC methods allows to derive algorithms for on-line parameter estimation in hidden Markov models and thus reducing both the complexity and the memory requirements of off-line parameter estimation techniques such as stochastic version of the EM. More precisely, two different schemes based respectively on point estimation and filtering method are derived for the considered problem. In order to improve the approximation of the

static posterior distribution in the filtering method, we propose an original update of the parameters based on the maximization of the posterior probability density function associated to each particle trajectory. This scheme leads to a concentration of the search space in regions with high posterior probability and consequently to a more accurate estimation.

This paper is organized as follows. In Section II, models of both the OFDM system and the phase distortion are introduced, leading to the dynamic state space (DSS) representation. Section III is devoted to the description of the estimation strategies based on Monte Carlo methods. The different solutions for parameter estimation using MC methods are also described. In Sections IV and V, the different approaches for the estimation problem are detailed. Section VII is devoted to the simulation results. The efficiency of the proposed algorithms is assessed and strategies for parameter estimation are compared. Finally, we focus on the performance of both phase distortion and channel estimation which is evaluated by the mean square error (MSE). Conclusions are given in Section VIII.

II. SYSTEM MODEL

In this paper, N , N_{cp} and T denote respectively the number of subcarriers, the cyclic prefix length and the OFDM symbol duration excluding the cyclic prefix. Let $\mathcal{N}(\mathbf{x}; \mu, \Sigma)$ and $\mathcal{N}_c(\mathbf{x}; \mu, \Sigma)$ represent respectively real and circularly symmetric complex Gaussian random vectors with mean μ and covariance matrix Σ . \mathbf{I}_n and $\mathbf{0}_{n \times m}$ are respectively the $n \times n$ identity matrix and the $n \times m$ matrix of zeros. Finally, lower case bold letters are used for column vectors and capital bold letters for matrices. $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote respectively conjugate, transpose, and Hermitian transpose.

A. Signal Model

We consider an OFDM system with N subcarriers. The transmitted OFDM signal s_k is generated via an inverse FFT operation applied on the subcarrier symbols d_k . To prevent intersymbol interference (ISI), a cyclic prefix of length N_{cp} is placed in front of the OFDM signal useful part of duration T .

We assume a slow fading frequency-selective channel with L paths. The CIR remains constant during the transmission of one burst including several OFDM symbols. Assuming perfect timing synchronization and $L \leq N_{cp}$, the complex baseband received OFDM signal can be written, after removal of the cyclic prefix :

$$r_k = e^{j(\theta_k + 2\pi k\epsilon/N)} \sum_{l=0}^{L-1} h_l s_{k-l} + b_k \quad (1)$$

where k denotes the k -th sample of the OFDM symbol, $\{h_l\}_{l=0}^{L-1}$, $\{s_k\}_{k=0}^{N-1}$, $\{\theta_k\}_{k=0}^{N-1}$ and $\{b_k\}_{k=0}^{N-1}$ are respectively the CIR, the transmitted signal, the PHN and a circular zero mean gaussian white noise with

power σ_b^2 , ϵ is the normalized carrier frequency offset. In the matrix form, equation (1) can be written as :

$$r_k = e^{j(\theta_k + 2\pi k\epsilon/N)} \mathbf{s}_k^T \mathbf{h} + b_k \quad (2)$$

where $\mathbf{s}_k = [s_k \ \cdots \ s_{k-L+1}]^T$ is the transmitted OFDM signal vector and $\mathbf{h} = [h_0 \ \cdots \ h_{L-1}]^T$ is the CIR vector.

B. Phase Distortion Model

In a baseband complex equivalent form, the carrier delivered by the noisy oscillator can be modeled as :

$$p(t) = \exp(j\phi(t)) \quad (3)$$

where the phase distortion $\phi(t)$ represents both the phase noise (PHN) and the carrier frequency offset (CFO) and can be written as follows :

$$\phi(t) = \theta(t) + 2\pi\Delta f t \quad (4)$$

where $\theta(t)$ and Δf correspond respectively to the PHN and the carrier frequency offset. The PHN is modeled as a Brownian process [1], [2], i.e. for $k = 0, \dots, N - 1$:

$$\theta_k = \theta_{k-1} + v_k \quad (5)$$

where k denotes the k -th time sample of the OFDM symbol and v_k is an independent and identically distributed (i.i.d.) zero mean Gaussian variable with variance $\sigma_v^2 = 2\pi\beta T/N$ with βT the PHN rate.

As a consequence, using (4) and (5) and assuming the initial condition $\phi_{-1} = 0$ due to perfect phase synchronization at the beginning of the OFDM symbol as in [2], [6], a discrete recursive relation for phase distortions including both PHN and CFO can be summarized as :

$$\phi_k = \begin{cases} v_0, & \text{for } k = 0 \\ \phi_{k-1} + \frac{2\pi\epsilon}{N} + v_k, & \text{for } k > 0 \end{cases} \quad (6)$$

where $\epsilon = \Delta f T$, denoted below by CFO, is the normalized carrier frequency offset with respect to the subcarrier spacing.

C. Dynamic State-Space Model

This paper focuses on the accurate estimation of CIR, CFO and PHN from a single OFDM symbol. At the receiver, the transmitted OFDM signal $\{s_k\}_{k=0}^{N-1}$ is perfectly known unlike both the instantaneous PHN power σ_v^2 and the AWGN power σ_b^2 . The mathematical foundation of our solution is the Bayesian theory which requires a dynamic state-space (DSS) modeling both the observation and the hidden process.

By combining (2) and (6), we obtain the following dynamic state-space model :

$$\begin{cases} \phi_k = \begin{cases} v_0, & \text{for } k = 0 \\ \phi_{k-1} + \frac{2\pi\epsilon}{N} + v_k, & \text{for } k > 0 \end{cases} \\ r_k = e^{j\phi_k} \mathbf{s}_k^T \mathbf{h} + b_k \end{cases} \quad (7)$$

The PHN process ϕ is the only dynamic state and the others \mathbf{h} , ϵ and $\sigma^2 = \{\sigma_v^2, \sigma_b^2\}$ are static parameters. In a Bayesian perspective, all necessary information is contained in the joint posterior probability distribution function (pdf), $p(\mathbf{h}, \phi_{0:N-1}, \epsilon, \sigma^2 | r_{0:N-1})$. Unfortunately, this distribution is analytically intractable, so we propose its numerical approximation via Monte Carlo methods.

Let us remarked that when CIR, PHN and CFO are jointly estimated, it can be shown that the phase distortions can be accurately estimated, but with a constant gap Φ from the true values :

$$\widehat{\phi}_{0:k} \rightarrow \phi_{0:k} + \Phi \quad (8)$$

$$\widehat{\mathbf{h}} \rightarrow e^{-j\Phi} \mathbf{h} \quad (9)$$

Indeed, from the observation equation of the DSS model (7), introducing two opposite phase rotations to PHN and CIR results in the same overall distortion $e^{j\phi_k} \mathbf{h}$, and hence the likelihood function is unchanged. The exact analysis of this residual common phase rotation is difficult, especially when PHN and AWGN powers are assumed unknown. A fairly good understanding of its origin is proposed in [9] where PHN and AWGN powers are perfectly known by the receiver. Nevertheless, this rotation can be estimated during the data detection step using pilot symbols [9] and does not affect the final symbol detection.

III. ESTIMATION APPROACHES BASED ON MONTE CARLO METHODS

A. Introduction

In this paper, we propose an off-line and an on-line estimator of the unknown states. Whereas off-line estimation is done using a block of data which corresponds in our context to the entire OFDM symbol, on-line estimation is recursive, i.e. the parameter estimates are updated at each time instant.

In [9], the authors propose an off-line maximum a posteriori estimator of $\{\phi_{0:N-1}, \mathbf{h}, \epsilon\}$, the Joint CFO/PHN/CIR Estimator (JCPCE) assuming that noise variances σ^2 are perfectly known at the receiver side. Moreover, since an analytical form of the optimal solution of the CFO cannot be found, an optimization scheme such as conjugate gradient algorithm has to be carried out in a first step of the JCPCE. Its complexity is scaled with the inverse of the resolution required for $\hat{\epsilon}$. In this paper, we deal with the difficult task where the noise variances are unknown. As a consequence, we firstly propose an estimator of $\lambda = \{\sigma^2, \epsilon\}$ in a batch way, i.e. given the entire OFDM symbol $r_{0:N-1}$ (off-line context), based on the Expectation-Maximization algorithm. Due to the complex nature of the system, the expectation step of the EM is analytically intractable, so we propose a stochastic version of EM, namely the Stochastic Approximation of Expectation Maximization (SAEM). Once the parameter λ has been estimated after I iterations of the SAEM algorithm, the JCPCE can be performed with the noise variance estimates. Moreover, the estimate of ϵ can be used as the initial value of the search method proposed in the JCPCE thus reducing its complexity.

On the other hand, we propose two different Sequential Monte Carlo (SMC) methods. In these schemes, the estimation of unknown state and parameters is performed jointly and sequentially (on-line context) thus reducing both the computational cost and the memory requirement induced by the use of the SAEM in the first solution. Indeed in the on-line mode, the N observations contained in the OFDM training symbol is only treated one time in sequential manner unlike in the off-line case. Indeed in the batch mode, the algorithm requires several iterations over the block of N observations in order to converge.

In the following, we focus on the crucial point of our estimation problem which is the parameter estimation task in hidden Markov model using Monte Carlo methods.

B. Parameter Estimation using Monte Carlo methods

In this section, MC-based algorithms for parameter estimation are reviewed in off-line and in on-line framework. The off-line estimation can be done by stochastic versions of EM-type algorithms. For on-line estimation, one can resort to SMC methods. In this case, we propose a new strategy to improve the accuracy of the parameter estimates.

1) *MC-based algorithms*: Since the paper of Dempster et al. [10], the Expectation-Maximization (EM) algorithm has become a highly appreciated tool for the estimation of parameters in models with missing data. Each iteration of an EM algorithm formally consists of an E-step and a separate M-step. The E-step calculates a conditional expectation while the M-step maximizes this expectation. Given an estimate λ_{n-1}

of the parameter, the estimate at the next iteration (n) is given by :

$$\lambda_n = \arg \max_{\lambda} Q(\lambda, \lambda_{n-1}) \quad (10)$$

where

$$Q(\lambda, \lambda_{n-1}) = \int \log p(z_{0:K}, r_{0:K} | \lambda) p(z_{0:K} | r_{0:K}, \lambda_{n-1}) dz_{0:k} \quad (11)$$

with K the length of observations available for the batch estimation.

However in many complex problems such as nonlinear and/or non-Gaussian models, at least one of these steps is often analytically intractable. Many authors have suggested that E-step may be overcome by approximating the expectation with the MC methods [11], leading to the principle of the Monte Carlo EM (MCEM). Using MC methods and importance sampling for the estimation of $p(z_{0:K} | r_{0:K}, \lambda_{n-1})$, Eq. (11) is approximated by :

$$\widehat{Q}(\lambda, \lambda_{n-1}) = \sum_{j=1}^{M(n)} \tilde{w}^{(j),n} \log p(z_{0:K}^{(j),n}, r_{0:K} | \lambda) \quad (12)$$

where $\{z_{0:K}^{(j),n}\}_{j=1}^{M(n)}$ and $\{\tilde{w}^{(j),n} = w^{(j),n} / \sum_{i=1}^{M(n)} w^{(i),n}\}_{j=1}^{M(n)}$ are respectively the $M(n)$ samples of missing data and the normalized importance weights obtained at the n -th iteration. The importance weights are obtained as follows :

$$\tilde{w}^{(j),n} = \frac{p(z_{0:K}^{(j),n} | r_{0:K}, \lambda_{n-1})}{\pi(z_{0:K}^{(j),n} | r_{0:K}, \lambda_{n-1})} \quad (13)$$

where $\pi(z_{0:K}^{(j),n} | r_{0:K}, \lambda_{n-1})$ is the importance function used for the generation of $\{z_{0:K}^{(j),n}\}_{j=1}^{M(n)}$. If $\{z_{0:k}^{(j),n}\}_{j=1}^{M(n)}$ is either sampled directly from $p(z_{0:K} | r_{0:K}, \lambda_{n-1})$ or obtained using MCMC methods [12] then $\{\tilde{w}^{(j),n}\}_{j=1}^{M(n)} = 1/M(n)$.

This function can be easily maximized when $p(z_{0:K}^{(j),n}, r_{0:K} | \lambda)$ belongs to the exponential family. In this case, $p(z_{0:K}, r_{0:K} | \lambda)$ depends on $(z_{0:K}, r_{0:K})$ only through a set of sufficient statistics.

However, when approximating Q by \widehat{Q} , the Monte Carlo sample size $M(n)$ has to be increased as the algorithm moves along. In fact, Booth et al. argue that MCEM will never converge if $M(n)$ is held fixed across iterations because of a persevering Monte Carlo error [13].

A variant of the basic approach is to use a stochastic approximation version of EM, the SAEM [14]. Delyon et al. have proved that, in contrast to MCEM, SAEM converges with a fixed Monte Carlo sample size thus reducing the computational demand of the parameter estimation algorithm. Similarly to the MCEM, the first step of the SAEM consists in generating realizations of the missing data vector under

the posterior distribution using MC methodology. Then, the Monte Carlo integration (11) is replaced by the following stochastic averaging procedure :

$$\widehat{Q}_n(\lambda) = (1 - \gamma_n)\widehat{Q}_{n-1}(\lambda) + \gamma_n \sum_{j=1}^M \widetilde{w}^{(j)} \log p(z_{0:K}^{(j),n}, r_{0:K} | \lambda) \quad (14)$$

where $\{\gamma_n\}_{k \geq 1}$ is a sequence of positive step-size which is typically $\gamma_n = n^{-\alpha}$ with $\alpha \in (1/2; 1]$. The new parameter estimate is obtained by maximizing the quantity $\widehat{Q}_n(\lambda)$.

In comparison with the MCEM algorithm, the SAEM makes more efficient the use of the imputed missing values. At each new iteration of the MCEM algorithm, a whole set of missing values needs to be simulated and all the missing values simulated during the previous iteration are dropped. In the SAEM algorithm, all the simulated missing values contribute to the evaluation of the auxiliary quantity $\widehat{Q}_n(\lambda)$ with a forgetting factor inversely proportional to the step-size.

2) *SMC-based algorithms*: Particle filtering algorithms are powerful tools for tracking time-varying signals in random dynamic systems [15]. These methods, also called particle filters, have become very popular over the last few years for non-Gaussian and nonlinear problems [16]. More precisely by exploiting the principle of importance sampling, SMC methods approximate the sequence of probability distribution $\{p(z_{0:k} | r_{0:k}), k = 0, 1, \dots, K\}$ without modifying the past simulated states $\left\{z_{0:k-1}^{(j)}\right\}_{j=1}^M$:

$$\widehat{p}(z_{0:k} | r_{0:k}) = \sum_{j=1}^M \widetilde{w}_k^{(j)} \delta(z_{0:k} - z_{0:k}^{(j)}) \quad (15)$$

The Sequential Importance Sampling (SIS) algorithm is the basis for most sequential MC filters. The SIS algorithm have two important operations : generation of new particles from an importance function $\pi(z_k | z_{0:k-1}^{(j)}, r_{0:k})$ and computation of the particle weights as follows :

$$w_k^{(j)} \propto \frac{p(r_k | z_{0:k}^{(j)}, r_{0:k-1}) p(z_k^{(j)} | z_{0:k-1}^{(j)})}{\pi(z_k^{(j)} | z_{0:k-1}^{(j)}, r_{0:k})} w_{k-1}^{(j)} \quad (16)$$

After a few iterations, the SIS algorithm is known to suffer from degeneracy problems. To reduce these problems, the Sequential Importance Resampling (SIR) integrates a resampling step to select particles for new generations in proportion to the importance weights [15]. Liu and Chen [17] have introduced a measure known as the *effective sample size*, $N_{eff} = \left[\sum_{m=1}^M (\widetilde{w}_k^{(m)})^2\right]^{-1}$, and have proposed to apply the resampling procedure whenever N_{eff} is below a predefined threshold. For the resampling step, we use the Systematic Resampling scheme described in [18]. This scheme outperforms the simple random sampling scheme with a small Monte Carlo variance and a favorable computational time [19], [20].

SMC methods are well adapted to problems such as optimal filtering where we are interested in estimating a sequence of potentially quickly varying distributions whose dimension is increasing over

time. However, parameter estimation using particle filtering is still a major issue. Since the problem we deal with consists in estimating both one dynamic state and four static parameters, this task has to be considered with care.

First, if the general model contains a linear Gaussian sub-structure, it can be exploited to derive an accurate estimator with a lower computational demand. Using both the Bayes' theorem and the DSS model, the linear gaussian variables are marginalized out and estimated by a Kalman filter, which is the optimal filter in this case. Also called Rao-Blackwellized or marginalized particle filter [15], [21]–[23], this technique is mostly used for time-varying estimation but has also been achieved for fixed parameter estimation [17], [24].

Secondly, we consider the case where no linear Gaussian sub-structure is contained in the DSS. In this context, two different methodologies can be employed: the point estimation and the filtering method. The first strategy aims to give point estimate of the parameter λ from $\{r_{0:k}\}$ without a preliminary approximation of the posterior distribution $p(\lambda|r_{0:k})$. As an example, maximum likelihood estimation based on Expectation-Maximization (EM) technique using particle filtering has been proposed in [25]–[29]. However, the slow convergence of the EM is known as a major drawback [30]. The second strategy is filtering method which consists in approximating the joint posterior distribution of both the dynamic and static states $p(z_{0:k}, \lambda|r_{0:k})$, and thus $p(\lambda|r_{0:k})$ in order to deduce an estimation of λ . As a consequence, particle method is required in the parameter space. A classical approach consists in exploring the parameter space at the initialization step by setting up a prior distribution for the unknown parameters. However this scheme is known to be inefficient since only one value of the parameter will survive after several resampling steps. To address this problem, kernel smoothing techniques [31], [32], artificial evolution of parameters [32]–[34] and Markov chain Monte Carlo (MCMC) steps [35], [36] have been proposed. However, such solutions do not solve the fixed-parameter estimation problem. In [37], Papavasiliou proposes an adaptive particle filter which is a combination of the interacting particle filter and the Monte-Carlo filter used respectively for the dynamic states and the static parameters. It consists in running one particle filter for each Monte-Carlo sample of the static parameter. Uniform convergence of this algorithm has been demonstrated. The only major disadvantage is its high complexity. In this paper, we propose a new strategy for parameter estimation using filtering methods. In this filtering context, the particle approximation of the posterior distribution is given by :

$$\hat{p}(z_{0:k}, \lambda|r_{0:k}) = \sum_{j=1}^M \delta(z_{0:k} - z_{0:k}^{(j)}; \lambda - \lambda^{(j)}) \tilde{w}_k^{(j)} \quad (17)$$

with $\delta(\cdot; \cdot)$ the two-dimensional Dirac delta function and where $\lambda^{(j)}$ denotes the particles in the parameter space. In our proposition, $\lambda^{(j)}$ are originally set to the maximum of the posterior distribution associated to each particle trajectory, i.e. $\lambda^{(j)} = \arg \max_{\lambda} \left\{ p \left(z_{0:k}^{(j)}, \lambda | r_{0:k} \right) \right\}$.

Consequently, the proposed scheme has the advantage of adding diversity in the search space in comparison with point estimation method and especially of concentrating particles in regions with high posterior probability.

IV. CIR, CFO, PHN AND NOISE VARIANCE ESTIMATION BASED ON SAEM AND JCPCE

A. Introduction

In this section, the off-line estimator is derived. Since an off-line maximum a posteriori estimator of CIR, PHN and CFO assuming the noise variances known is described in [9], we propose a new strategy in order to estimate them before the use of the JCPCE. The SAEM algorithm aims to iteratively estimate noise variances and CFO from the received OFDM symbol $r_{0:N-1}$ as shown in Fig. 1. The estimate value of the CFO $\tilde{\epsilon}$ is used as initial value of the search method proposed in [9] for finding $\hat{\epsilon}$ the optimal solution of ϵ using the maximum a posteriori criterion.

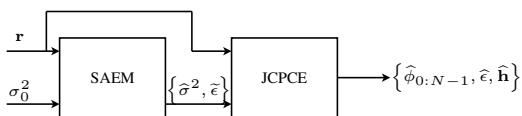


Fig. 1. Structure of the proposed SAEM-JCPCE

B. The proposed SAEM algorithm

Given an estimate of the noise variance obtained at the $(n - 1)$ -th iteration and denoted by λ_{n-1} , the SAEM requires the approximation of the posterior distribution $p(\boldsymbol{\phi}, \mathbf{h} | \mathbf{r}, \lambda_{n-1})$ using MC methods with $\boldsymbol{\phi} = [\phi_{N-1} \ \cdots \ \phi_0]^T$ and $\mathbf{r} = [r_{N-1} \ \cdots \ r_0]^T$. As detailed in Appendix A, the estimate of the posterior distribution of interest obtained by the Monte Carlo filter is given by :

$$\hat{p}(\boldsymbol{\phi}, \mathbf{h} | \mathbf{r}, \lambda_{n-1}) = \sum_{j=1}^M \tilde{w}^{(j),n} \delta(\boldsymbol{\phi} - \boldsymbol{\phi}^{(j),n}) p(\mathbf{h} | \boldsymbol{\phi}^{(j),n}, \mathbf{r}, \lambda_{n-1}) \quad (18)$$

with $p(\mathbf{h} | \boldsymbol{\phi}^{(j),n}, \mathbf{r}, \lambda_{n-1})$ defined in (59).

Once the posterior distribution of interest has been approximated and given an estimate λ_{n-1} of the unknown parameters as described previously, the aim is to find an estimate of the noise variances and the CFO at the current iteration, i.e.

$$\lambda_n = \arg \max_{\lambda} \left\{ \widehat{Q}_n(\lambda) \right\} \quad (19)$$

where using (14) and (18),

$$\begin{aligned} \widehat{Q}_n(\lambda) = & (1 - \gamma_n) \widehat{Q}_{n-1}(\lambda) \\ & + \gamma_n \sum_{j=1}^M \widetilde{w}^{(j),n} \int \log p(\boldsymbol{\phi}^{(j),n}, \mathbf{h}, \mathbf{r} | \lambda) p(\mathbf{h} | \boldsymbol{\phi}^{(j),n}, \mathbf{r}, \lambda_{n-1}) d\mathbf{h} \end{aligned} \quad (20)$$

which can also be written as :

$$\widehat{Q}_n(\lambda) = (1 - \gamma_n) \widehat{Q}_{n-1}(\lambda) + \gamma_n \left(A + \widehat{Q}'(\lambda, \lambda_{n-1}) \right) \quad (21)$$

where A is a constant. Since $p(\boldsymbol{\phi}^{(j),n}, \mathbf{h}, \mathbf{r} | \lambda) = p(\mathbf{r} | \boldsymbol{\phi}^{(j),n}, \mathbf{h}, \sigma_b^2) p(\boldsymbol{\phi}^{(j),n} | \epsilon, \sigma_v^2) p(\mathbf{h})$, $\widehat{Q}'(\lambda, \lambda_{n-1})$ becomes :

$$\begin{aligned} \widehat{Q}'(\lambda, \lambda_{n-1}) = & \sum_{j=1}^M \widetilde{w}^{(j),n} \left[\log p(\boldsymbol{\phi}^{(j),n} | \epsilon, \sigma_v^2) \right. \\ & + \log \left(\frac{1}{(2\pi\sigma_b^2)^{N/2}} \right) \\ & - \frac{1}{2\sigma_b^2} \left\{ (\mathbf{r} - \mathbf{C}^{(j),n} \mathbf{h}^{(j),n})^H (\mathbf{r} - \mathbf{C}^{(j),n} \mathbf{h}^{(j),n}) \right. \\ & \left. \left. + \text{trace} \left[\mathbf{C}^{(j),n} \boldsymbol{\Sigma}^{(j),n} (\mathbf{C}^{(j),n})^H \right] \right\} \right] \end{aligned} \quad (22)$$

Finally, the maximization of $\widehat{Q}_n(\lambda)$ gives the following update equation of the CFO, PHN and AWGN powers :

$$\begin{aligned} \epsilon_n = & (1 - \gamma_n) \epsilon_{n-1} \\ & + \gamma_n \sum_{j=1}^M \widetilde{w}^{(j),n} \frac{\boldsymbol{\mu}^T \boldsymbol{\Theta}^{-1} \boldsymbol{\phi}^{(j),n} + (\boldsymbol{\phi}^{(j),n})^T \boldsymbol{\Theta}^{-1} \boldsymbol{\mu}}{2\boldsymbol{\mu}^T \boldsymbol{\Theta}^{-1} \boldsymbol{\mu}} \end{aligned} \quad (23)$$

$$\begin{aligned}
\sigma_{v,n}^2 &= \frac{1}{N} \sum_{i=1}^{n-1} \gamma_i \prod_{k=i+1}^n (1 - \gamma_k) \\
&\quad \times \sum_{j=1}^M \tilde{w}^{(j),i} \left(\boldsymbol{\phi}^{(j),i} - \epsilon_n \boldsymbol{\mu} \right)^T \boldsymbol{\Theta}^{-1} \left(\boldsymbol{\phi}^{(j),i} - \epsilon_n \boldsymbol{\mu} \right) \\
&\quad + \gamma_n \sum_{j=1}^M \tilde{w}^{(j),n} \left(\boldsymbol{\phi}^{(j),n} - \epsilon_n \boldsymbol{\mu} \right)^T \boldsymbol{\Theta}^{-1} \left(\boldsymbol{\phi}^{(j),n} - \epsilon_n \boldsymbol{\mu} \right) \quad (24)
\end{aligned}$$

$$\begin{aligned}
\sigma_{b,n}^2 &= (1 - \gamma_n) \sigma_{b,n-1}^2 + \frac{\gamma_n}{N} \sum_{j=1}^M \tilde{w}^{(j),n} \left((\mathbf{r} - \mathbf{C}^{(j),n} \mathbf{h}^{(j),n})^H \right. \\
&\quad \left. (\mathbf{r} - \mathbf{C}^{(j),n} \mathbf{h}^{(j),n}) + \text{trace} \left[\mathbf{C}^{(j),n} \boldsymbol{\Sigma}^{(j),n} (\mathbf{C}^{(j),n})^H \right] \right) \quad (25)
\end{aligned}$$

The proposed SAEM algorithm is summed up in Table I.

TABLE I
THE PROPOSED SAEM ALGORITHM

<p>Initialization : Set $\{\epsilon_0; \sigma_{v,0}^2; \sigma_{b,0}^2\} = \{0; 0.1; 1\}$ For $n = 1 \dots I$ For $m = 1 \dots M$ Sample $\boldsymbol{\phi}^{(m),n}$ from the prior distribution using (51) Evaluate the <i>a posteriori</i> pdf of the CIR $p(\mathbf{h} \boldsymbol{\phi}^{(m),n}, \sigma_{b,n-1}^2, r_{0:k})$ using (60) and (61) Evaluate the corresponding importance weights using (53) end for Normalize importance weights : $\tilde{w}^{(j),n} = w^{(j),n} / \sum_{m=1}^M w^{(m),n}$ Estimation of ϵ_n, $\sigma_{v,n}^2$ and $\sigma_{b,n}^2$ using respectively (23), (24) and (25) end for Set $\{\tilde{\epsilon}; \tilde{\sigma}^2\} = \{\epsilon_I; \sigma_I^2\}$</p>
--

V. SMC METHODS FOR JOINT CIR, CFO, PHN AND NOISE VARIANCE ON-LINE ESTIMATION

We describe here in detail all the steps required in the implementation of algorithms for on-line estimation of the unknown states. To allow this on-line estimation, SMC methods are proposed. Like in the SAEM, the Rao-Blackwellization and importance sampling are respectively used for the channel and

the phase distortions but in a sequential manner. Now, concerning both the noise variances and the CFO, λ , two different strategies, namely point estimation or filtering methods, are derived in this section.

So, in a first time, we propose to describe the Rao-Blackwellization for the CIR and the SIS scheme for the phase distortions which is common to these two on-line estimation strategies. Then, we describe the proposed filtering scheme and a point-estimation method for parameter estimation using SMC algorithms, respectively denoted in the following by the PMAP-SMC (Particle Maximum A Posteriori using Sequential Monte Carlo) and the OEM-SMC (On-line Expectation Maximization using Sequential Monte Carlo) algorithms¹.

A. Rao-Blackwellization for the CIR estimation

The posterior distribution $p(\mathbf{h}|\phi_{0:k}^{(j)}, \lambda^{(j)}, r_{0:k})$ is obtained in a same way as in Appendix A-2 by introducing the time domain index k since a sequential method is considered in this section. This posterior pdf is $\mathcal{N}_c(\mathbf{h}; \mathbf{h}_k^{(j)}, \Sigma_k^{(j)})$ with :

$$\mathbf{h}_k^{(j)} = \frac{1}{L} \left(\mathbf{C}_k^{(j)} \right)^H \left(\frac{1}{L} \mathbf{C}_k^{(j)} \left(\mathbf{C}_k^{(j)} \right)^H + \sigma_b^{2(j)} \mathbf{I}_{k+1} \right)^{-1} \mathbf{r}_k \quad (26)$$

and,

$$\Sigma_k^{(j)} = \frac{1}{L} \mathbf{I}_L - \frac{1}{L^2} \left(\mathbf{C}_k^{(j)} \right)^H \left(\frac{1}{L} \mathbf{C}_k^{(j)} \left(\mathbf{C}_k^{(j)} \right)^H + \sigma_b^{2(j)} \mathbf{I}_{k+1} \right)^{-1} \mathbf{C}_k^{(j)} \quad (27)$$

where $\mathbf{r}_k = [r_k \ \dots \ r_0]^T$ and

$$\mathbf{C}_k^{(j)} = \begin{bmatrix} e^{j\phi_k^{(j)}} & & 0 \\ & \ddots & \\ 0 & & e^{j\phi_0^{(j)}} \end{bmatrix} \begin{bmatrix} \mathbf{s}_k^T \\ \vdots \\ \mathbf{s}_0^T \end{bmatrix}$$

B. Phase distortions importance sampling

The choice of the importance function is essential because it determines the efficiency as well as the complexity of the particle filtering algorithm. In this paper, we consider the optimal importance function

¹Let us remark that the Rao-Blackwellization and the SIS steps are derived in the following with particle index (j) for the parameter λ . However, contrary to PMAP-SMC, OEM-SMC does not require the use of particles in the parameter space so the adaptation of the two steps for the OEM-SMC is obtained by simply setting $\left\{ \lambda^{(j)} \right\}_{j=1}^M = \lambda_{k-1}$.

for phase distortions which minimizes the variance of the importance weights conditional upon the particle trajectories and the observations [38]. In our context, it is expressed as :

$$\pi(\phi_k | \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k}) = p(\phi_k | \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k}) \quad (28)$$

The sampling of ϕ_k from (28) requires the analytical expression of the optimal importance function. However, this pdf is analytically intractable. To derive an approximate optimal importance function, we propose to linearize only the term with the phase distortions driving noise, such as $e^{jv_k} = 1 + jv_k$ where v_k is defined in (6). Generally, the usual linearization of the PHN term $e^{j\theta_k} = 1 + j\theta_k$ is employed but is less accurate than the proposed one. As detailed in Appendix B, this pdf can thus be approximated by :

$$p(\phi_k | \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k}) \approx \mathcal{N}(\phi_k; \vartheta_k^{(j)}, \Lambda_k^{(j)}) \quad (29)$$

where $\Lambda_k^{(j)} = \frac{\chi_k^{(j)} \sigma_v^{2(j)}}{|\Gamma_k^{(j)}|^2 \sigma_v^{2(j)} + \chi_k^{(j)}}$ and

$$\vartheta_k^{(j)} = \begin{cases} \Upsilon_0^{(j)}, & \text{for } k = 0 \\ \Upsilon_k^{(j)} + \phi_{k-1}^{(j)} + 2\pi\epsilon^{(j)}/N, & \text{for } k > 0 \end{cases} \quad (30)$$

with $\chi_k^{(j)} = \mathbf{s}_k^T \Sigma_{k-1}^{(j)} \mathbf{s}_k^* + \sigma_b^{2(j)}$, $\Upsilon_k^{(j)} = \frac{\Im(\Gamma_k^{(j)*} r_k) \sigma_v^{2(j)}}{|\Gamma_k^{(j)}|^2 \sigma_v^{2(j)} + \chi_k^{(j)}}$ (with $\Im(\cdot)$ the imaginary part) and

$$\Gamma_k^{(j)} = \begin{cases} \mathbf{s}_0^T \mathbf{h}_{-1}^{(j)}, & \text{for } k = 0 \\ e^{j(\phi_{k-1}^{(j)} + 2\pi\epsilon^{(j)}/N)} \mathbf{s}_k^T \mathbf{h}_{k-1}^{(j)}, & \text{for } k > 0 \end{cases} \quad (31)$$

A Gaussian approximation of the optimal importance function is proposed in this paper for the phase distortions. However since the likelihood function is multimodal in our model, the suitability of this Gaussian approximation has to be considered with care [39]. As shown in Appendix C, the Gaussian approximation is accurate in our context. Indeed, the prior distribution of the phase distortions is narrow enough to ensure that only one mode is significant.

The sampling distribution of ϕ_k is now identified by (29). The corresponding unnormalized weights are then computed by :

$$w_k^{(j)} \propto w_{k-1}^{(j)} \frac{p(r_k | \phi_{0:k}^{(j)}, \lambda^{(j)}, r_{0:k-1}) p(\phi_k^{(j)} | \phi_{0:k-1}^{(j)}, \lambda^{(j)})}{\widehat{p}(\phi_k^{(j)} | \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k})} \quad (32)$$

with $\widehat{p}(\phi_k^{(j)} | \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k})$ the approximate optimal importance function given by (29),

$$p(\phi_k | \phi_{0:k-1}^{(j)}, \lambda^{(j)}) = \begin{cases} \mathcal{N}(\phi_0; 0, \sigma_v^2), & \text{for } k = 0 \\ \mathcal{N}(\phi_k; \phi_{k-1}^{(j)} + \frac{2\pi\epsilon^{(j)}}{N}, \sigma_v^2), & \text{for } k > 0 \end{cases} \quad (33)$$

$$p(r_k | \phi_{0:k}^{(j)}, \lambda^{(j)}, r_{0:k-1}) = \mathcal{N}_c(r_k; \rho_k^{(j)}, \chi_k^{(j)}) \quad (34)$$

and $\rho_k^{(j)} = e^{j\phi_k^{(j)}} \mathbf{s}_k^T \mathbf{h}_{k-1}^{(j)}$.

C. Estimation of CFO and noise variances

In this section, the proposed scheme for parameter estimation based on the maximization of the posterior distribution is described. For a fair comparison between our approach and a point estimation method, an on-line adaptation of the expectation-maximization algorithm using particle filtering proposed in [25] is derived.

1) *The Particle Maximum A Posteriori (PMAP) Estimator* : The main idea of our approach consists in finding the most adapted noise variance value to each particle trajectory. In other words, our strategy aims to concentrate particles in regions with high posterior probability. Therefore, the noise variance value of the j -th particle is set to the maximum of the joint posterior distribution associated to the j particle trajectory, i.e. :

$$\lambda^{(j)} = \arg \max_{\lambda} \left\{ p \left(x_{0:k}^{(j)}, \lambda | r_{0:k} \right) \right\} \quad (35)$$

Under the assumption of non-informative a priori distribution of λ and independence of noises in both state and observation equations, the posterior distribution involved in (35) can be written as :

$$p \left(x_{0:k}^{(j)}, \lambda | r_{0:k} \right) = \prod_{n=0}^k p \left(r_n | \phi_{0:n}^{(j)}, \mathbf{h}_k^{(j)}, \lambda, r_{0:n-1} \right) \times p \left(\phi_n^{(j)} | \phi_{n-1}^{(j)}, \lambda \right) \quad (36)$$

where the pdfs involved in (36) are given respectively by (33) and (66) in Appendix B. Using (35) and (66), the AWGN variance estimate, at time k and for the j -th particle, is given by :

$$\sigma_b^{2(j)} = \frac{1}{k+1} \sum_{n=0}^k \left| r_n - e^{j\phi_n^{(j)}} \mathbf{s}_n^T \mathbf{h}_k^{(j)} \right|^2 \quad (37)$$

Since both $\phi_k^{(j)}$ and $\mathbf{h}_k^{(j)}$ are required in (37), the AWGN variance estimation should be performed after the PHN sampling and the Rao-Blackwellization of the CIR. However, as shown by Eqs. (26), (27) and (29), $\sigma_b^{2(j)}$ is needed for these two steps. Consequently, both the Rao-Blackwellization and the

importance sampling can be improved by using an AWGN variance estimate at time k which takes into account the new observation r_k . By replacing $\{\phi_k^{(j)}, \mathbf{h}_k^{(j)}\}$ with $\{\phi_{k-1}^{(j)} + 2\pi\epsilon^{(j)}/N, \mathbf{h}_{k-1}^{(j)}\}$ in (37), the observation noise variance update is finally expressed as :

$$\begin{aligned} \sigma_b^{2(j)} = & \frac{1}{k+1} \left(\sum_{n=0}^{k-1} \left| r_n - e^{j\phi_n^{(j)}} \mathbf{s}_n^T \mathbf{h}_{k-1}^{(j)} \right|^2 \right. \\ & \left. + \left| r_k - e^{j(\phi_{k-1}^{(j)} + 2\pi\epsilon^{(j)}/N)} \mathbf{s}_k^T \mathbf{h}_{k-1}^{(j)} \right|^2 \right) \end{aligned} \quad (38)$$

Concerning the CFO and the PHN variance, their estimates are straightforwardly obtained using (33) and (35) :

$$\epsilon^{(j)} = \frac{N \sum_{n=1}^k \phi_n^{(j)} - \phi_{n-1}^{(j)}}{2k\pi} \quad (39)$$

$$\sigma_v^{2(j)} = \frac{1}{k+1} \left(\sum_{n=1}^k \left| \phi_n^{(j)} - \phi_{n-1}^{(j)} + 2\pi\epsilon^{(j)}/N \right|^2 + |\phi_0|^2 \right) \quad (40)$$

The PMAP-SMC is summed up in Table II.

2) *The on-line EM estimator:* In [25], Cappé and Moulines investigate the maximum likelihood estimation of parameters based on the EM algorithm using particle filtering. Since they consider batch estimation, we derive below an on-line adaptation of this approach for a fair comparison with the previous filtering method. In this context, the maximum likelihood estimate of λ (e.g. λ_k at the k -th time domain sample of the OFDM symbol) can be obtained by maximizing the following quantity

$$Q(\lambda, \lambda_{k-1}) = \int \log p(x_{0:k}, r_{0:k} | \lambda) p(x_{0:k} | r_{0:k}, \lambda_{k-1}) dx_{0:k} \quad (41)$$

where $p(x_{0:k} | r_{0:k}, \lambda_{k-1})$ is approximated by the SMC principle. Using the particle approximation of the posterior distribution $p(x_{0:k} | r_{0:k}, \lambda_{k-1})$, the estimation of λ is updated after each new received time sample of the OFDM symbol and is given by :

$$\epsilon_k = \frac{N}{2k\pi} \sum_{j=1}^M \tilde{w}_k^{(j)} \sum_{n=1}^k \left(\phi_n^{(j)} - \phi_{n-1}^{(j)} \right) \quad (42)$$

$$\begin{aligned} \sigma_{v,k}^2 = & \frac{1}{k+1} \sum_{j=1}^M \tilde{w}_k^{(j)} \left(\sum_{n=1}^k \left| \phi_n^{(j)} - \phi_{n-1}^{(j)} - 2\pi\epsilon_k/N \right|^2 \right. \\ & \left. + \left| \phi_0^{(j)} \right|^2 \right) \end{aligned} \quad (43)$$

TABLE II
PARTICLE MAXIMUM A POSTERIORI USING SEQUENTIAL MONTE CARLO ALGORITHM (PMAP-SMC)

<p>Initialization :</p> $\left\{ \theta_{-1}^{(m)}; \mathbf{h}_{-1}^{(m)}; \boldsymbol{\Sigma}_{-1}^{(m)}; \tilde{w}_k^{(m)} \right\}_{m=1}^M = \left\{ 0; \mathbf{0}_{L \times 1}; \frac{1}{L} \mathbf{I}_L; 1/M \right\}$ $\left\{ \lambda^m \right\}_{m=1}^M = \left\{ \tilde{\sigma}_v^{2(m)}; \tilde{\sigma}_b^{2(m)}; \epsilon^m \right\}_{m=1}^M = \{0.01; 1; 0\}$ <p>For $k = 0 \dots N - 1$</p> <p style="padding-left: 20px;">For $m = 1 \dots M$</p> <p style="padding-left: 40px;">Estimation of the observation noise statistics using (38)</p> <p style="padding-left: 40px;">Sample $\phi_k^{(m)}$ from the optimal importance function using (29)</p> <p style="padding-left: 40px;">Evaluate the <i>a posteriori</i> pdf of the CIR $p(\mathbf{h} \phi_{0:k}^{(m)}, \lambda^{(m)}, r_{0:k})$ using (26) and (27)</p> <p style="padding-left: 40px;">Estimation of the CFO and PHN noise variance using respectively (39) and (40)</p> <p style="padding-left: 40px;">Evaluate the corresponding importance weights using (32)</p> <p style="padding-left: 20px;">end for</p> <p style="padding-left: 40px;">Normalize importance weights : $\tilde{w}_k^{(j)} = w_k^{(j)} / \sum_{m=1}^M w_k^{(m)}$</p> <p style="padding-left: 20px;">If $N_{eff} < N_{seuil}$ then</p> <p style="padding-left: 40px;">Resample particle trajectories using systematic resampling</p> <p style="padding-left: 20px;">end if</p> <p>end for</p> <p>Evaluate the final smoothed estimates $\hat{x}_{0:N-1}$ using :</p> $\hat{x}_{0:N-1} = \sum_{j=1}^M \tilde{w}_{N-1}^{(j)} x_{0:N-1}^{(j)}$ <p>with $x_{0:N-1}^{(j)} = \left\{ \phi_{0:N-1}^{(j)}, \epsilon^{(j)}, \mathbf{h}^{(j)}, \sigma^{2(j)} \right\}$</p>
--

$$\sigma_{b,k}^2 = \frac{1}{k+1} \sum_{j=1}^M \tilde{w}_k^{(j)} \left(\sum_{n=0}^k \left| r_n - e^{j\phi_n^{(j)}} \mathbf{s}_n^T \mathbf{h}_k^{(j)} \right|^2 + \text{trace} \left[\mathbf{C}_k^{(j)} \boldsymbol{\Sigma}^{(j)} (\mathbf{C}_k^{(j)})^H \right] \right) \quad (44)$$

Since the importance weights $\left\{ \tilde{w}_k^{(j)} \right\}_{j=1}^M$ are required in the on-line EM, it is obvious that this on-line expectation maximization estimation is performed after both the phase distortions sampling and the Rao-Blackwellization for the CIR. The OEM-SMC is summed up in Table III.

VI. THE POSTERIOR CRAMÉR-RAO BOUND

In order to study the efficiency of an estimation method, it is of great interest to compute the variance bounds on the estimation errors and to compare them to the lowest bounds corresponding to the optimal estimator. For time-invariant statistical models, a commonly used lower bound is the Cramér-Rao bound

TABLE III
ON-LINE EXPECTATION-MAXIMIZATION USING SEQUENTIAL MONTE CARLO ALGORITHM (OEM-SMC)

Initialization :

$$\left\{ \theta_{-1}^{(m)}; \mathbf{h}_{-1}^{(m)}; \boldsymbol{\Sigma}_{-1}^{(m)}; \tilde{w}_k^{(m)} \right\}_{m=1}^M = \left\{ 0; \mathbf{0}_{L \times 1}; \frac{1}{L} \mathbf{I}_L; 1/M \right\}$$

$$\{\lambda_{-1}\} = \left\{ \hat{\sigma}_{v,-1}^2; \hat{\sigma}_{b,-1}^2; \epsilon_{-1} \right\} = \{0.01; 1; 0\}$$

For $k = 0 \dots N - 1$

For $m = 1 \dots M$

Sample $\phi_k^{(m)}$ from the optimal importance function using (29)

Evaluate the *a posteriori* pdf of the CIR $p(\mathbf{h} | \phi_{0:k}^{(m)}, \lambda_{k-1}, r_{0:k})$ using (26) and (27)

Evaluate the corresponding importance weights using (32)

end for

Normalize importance weights : $\tilde{w}_k^{(j)} = w_k^{(j)} / \sum_{m=1}^M w_k^{(m)}$

On-line EM estimation of CFO and unknown noise statistics using (42), (43) and (44)

If $N_{eff} < N_{seuil}$ **then**

Resample particle trajectories using systematic resampling

end if

end for

Evaluate the final smoothed estimates $\hat{x}_{0:N-1}$ using

$$\hat{x}_{0:N-1} = \sum_{j=1}^M \tilde{w}_{N-1}^{(j)} x_{0:N-1}^{(j)}$$

with $x_{0:N-1}^{(j)} = \left\{ \phi_{0:N-1}^{(j)}, \mathbf{h}^{(j)} \right\}$

(CRB), given by the inverse of the Fisher information matrix. In a time-varying context as we deal with here, a lower bound analogous to the CRB for random parameters has been derived in [40]; this bound is usually referred to as the Van Trees version of the CRB, or posterior CRB (PCRB) [41].

In this topic, the PCRB for the joint estimation of $\{\phi, \mathbf{h}, \epsilon, \sigma^2\}$ is analytically intractable. Since the CIR is the main quantity of interest, we derive in this paper the PCRB on the estimation error of \mathbf{h} in the absence of both CFO and PHN and where σ_b^2 is assumed perfectly known. In this context, the DSS model is linear and Gaussian since the likelihood function of interest is given by :

$$p(\mathbf{r} | \mathbf{h}) = \mathcal{N}_c(\mathbf{r}; \mathbf{S}\mathbf{h}, \sigma_b^2 \mathbf{I}_N) \quad (45)$$

with

$$\mathbf{S} = \begin{bmatrix} \mathbf{s}_{N-1}^T \\ \vdots \\ \mathbf{s}_0^T \end{bmatrix}$$

As a consequence, the Cramér-Rao bound for a given OFDM symbol is equal to the covariance matrix Σ of the posterior pdf $p(\mathbf{h}|\mathbf{r})$ and is defined using (45) and (55) by :

$$\Sigma = \frac{1}{L}\mathbf{I}_L - \frac{1}{L^2}\mathbf{S}^H \left(\frac{1}{L}\mathbf{S}\mathbf{S}^H + \sigma_b^2\mathbf{I}_N \right)^{-1} \mathbf{S} \quad (46)$$

where the elements of \mathbf{S} correspond to the OFDM training signal. The PCRb is thus obtained through numerical simulations by the expectation of (46) over the OFDM symbol \mathbf{S} :

$$\text{PCRb} = \mathbb{E} [\text{trace} [\Sigma]] \quad (47)$$

VII. SIMULATIONS

In this section, performance of the proposed strategies is studied through numerical simulations. The following system parameters are considered : a Rayleigh multipath channel with a delay of L taps and a uniform power delay profile, an OFDM training symbol with $N = 64$ subcarriers and a cyclic prefix of $N_{cp} = 16$ samples. Each subcarrier is arbitrarily modulated in quaternary phase-shift keying (QPSK). For each OFDM symbol, the CFO ϵ is drawn from a uniform distribution in $[-0.5; 0.5]$ and both SAEM and SMC algorithms are implemented with $M = 100$ particles. Since the unresolvable residual common phase rotation described in section II-C is not the subject of this paper, we assume that it can be perfectly corrected to facilitate the analysis of channel estimation as in [9].

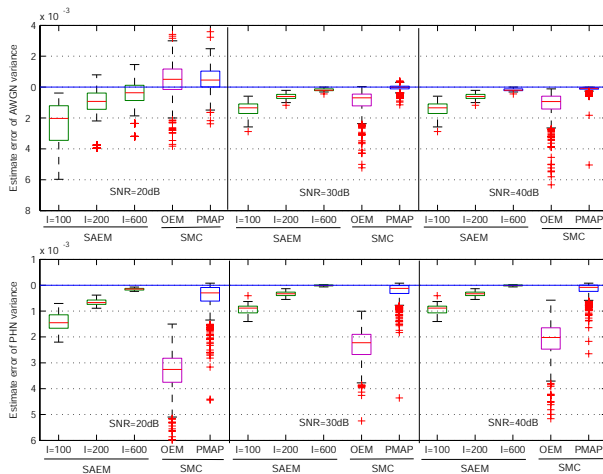


Fig. 2. Box-and-whiskers plots of the estimate error of AWGN (top) and PHN (bottom) power obtained by SAEM (with different numbers of iterations I), the PMAP-SMC and the OEM-SMC algorithms for several SNR values and with $\beta T = 10^{-3}$ and $L = 10$ paths.

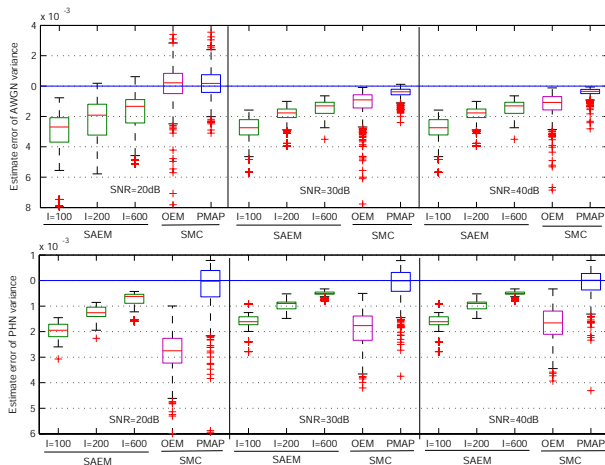


Fig. 3. Box-and-whiskers plots of the estimate error of AWGN (top) and PHN (bottom) power obtained by SAEM (with different numbers of iterations I), the PMAP-SMC and the OEM-SMC algorithms for several SNR values and with $\beta T = 10^{-2}$ and $L = 10$ paths.

For comparison purposes between parameter estimators, we have firstly studied the distribution of PHN and AWGN power estimate error. Box-and-whiskers plots of PHN and AWGN power estimate errors of the SAEM, OEM-SMC and PMAP-SMC are drawn in Fig. 2 with $\beta T = 10^{-3}$ and in Fig. 3 with $\beta T = 10^{-2}$. A Box-and-whiskers plot is a convenient way of graphically depicting the five-number summary including the smallest observation, the lower quartile, the median, the upper quartile and the largest observation. It enables to display different types of populations, without any assumption on the statistical distribution. The spacing between the different parts of the box helps to visualize variance, skew and outliers. It can be logically denoted that the SAEM algorithm gives a more accurate parameter estimate of both AWGN and PHN variances when the number of iterations increases. However, after 600 iterations, there is no more improvement in parameter estimation. Now concerning the on-line estimators based on SMC methods, box-and-whiskers plots show that estimation of both PHN and AWGN powers is better performed with the PMAP-SMC than with the OEM-SMC. Indeed, the bias obtained by the PMAP-SMC is smaller than the one obtained by the OEM-SMC. Moreover, the PMAP-SMC offers the smaller estimation variance. These remarks about bias and variance become more pronounced when the SNR increases. Therefore, it is obvious that the diversity introduced in the search space of λ in the PMAP-SMC significantly improves the precision of the final estimates for both PHN and AWGN powers. Finally, the PMAP-SMC is more accurate in the estimation of AWGN variance than the SAEM, especially

when both SNR and βT increase. Moreover, concerning the PHN variance, the SAEM achieves the best estimation accuracy for small phase noise rate but in the case of $\beta T = 10^{-2}$, the bias of the PMAP-SMC estimator is smaller than the SAEM one. It can be explained by the sequential nature of the PMAP-SMC which is more adapted to time-varying state estimation than MC methods. Indeed, the sampling step is improved in the PMAP-SMC since the phase distortions values are simulated sequentially by using an approximate optimal importance function which takes into account the observations.

The MSE of phase distortions (CFO + PHN) and channel estimation are respectively depicted in Figures 4 and 5 for several values of SNR and βT and for $L = 10$ paths. Performance of the JCPCE algorithm using noise variance estimates of the SAEM with 600 iterations is also plotted. The MSE of the channel estimation is also compared to the conditional posterior Cramér-Rao bound (PCRB). The MSE is obtained by $MSE = \frac{1}{Q} \sum_{n=1}^Q \text{trace} \left[\left(\mathbf{h}^n - \hat{\mathbf{h}}^n \right) \left(\mathbf{h}^n - \hat{\mathbf{h}}^n \right)^H \right]$ where Q is the number of OFDM symbols used in simulations and n denotes the n -th OFDM symbol ($Q = 10,000$). Firstly, the results point out the robustness of the proposed algorithms when the *a priori* statistics of the model are not perfectly known. As shown on this figure, for a SNR less than 20 dB, the proposed methods give results close to the optimal bound whatever the PHN power. Moreover, the MSE of the PMAP-SMC is consistently lower than the MSE of the OEM-SMC, especially for large SNR. Now concerning performance comparison between off-line and on-line estimators, even if the PMAP-SMC is more accurate in the noise variance estimation, the SAEM-JCPCE slightly outperforms the PMAP-SMC. It can be explained by the fact that at the beginning of the OFDM symbol, the PMAP-SMC parameter estimates are less accurate than at the end of the symbol due to the on-line nature of the proposed SMC methods. As a consequence, by taking into account worse noise variance estimates, the first samples of the phase distortions are less accurate than those obtained at the end of the OFDM symbol.

From these results, it can be concluded that PHN and AWGN power estimation has a significant impact for both phase distortion and CIR estimation. In on-line case, the diversity introduced in the search space of parameters clearly improves the parameter estimator performance, especially for a large SNR. The combination of the SAEM and the JCPCE outperforms the two on-line schemes using SMC methods but at the expense of higher complexity and memory requirements.

VIII. CONCLUSION

This paper deals with the major problem of OFDM channel estimation in the presence of CFO and PHN. Unlike [9], we have considered the difficult task where prior statistics of PHN and AWGN are assumed unknown at the receiver which is a more realistic case. We have proposed different strategies for

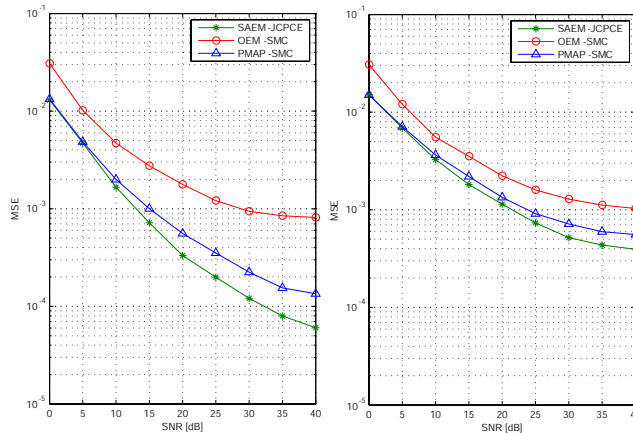


Fig. 4. MSE of phase distortions (CFO + PHN) vs. SNR with $L = 10$ paths and for phase noise rate $\beta T = 10^{-3}$ (left) and $\beta T = 10^{-2}$ (right).

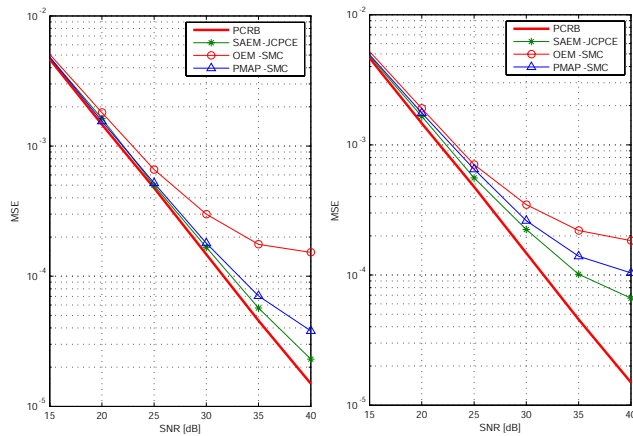


Fig. 5. MSE of channel estimation vs. SNR with $L = 10$ paths and for phase noise rate $\beta T = 10^{-3}$ (left) and $\beta T = 10^{-2}$ (right).

off-line and on-line estimation based on Monte Carlo methods. Firstly in the off-line case, we propose a combination of the SAEM and the JCPCE to perform a batch estimation of the quantities of interest. Secondly for the on-line estimation, we propose two different approaches to jointly and sequentially estimate the many unknowns using for parameters either point estimation or filtering method. In the former case, an on-line adaptation of the EM method using particle filtering proposed in [25], called the

OEM-SMC algorithm, is derived. In the latter case, an on-line maximum a posteriori update per particle for the noise variances has been proposed leading to an efficient parameter estimator, the PMAP-SMC.

Numerical simulations demonstrate the efficiency of the proposed algorithms for OFDM channel estimation in the presence of CFO and PHN when both PHN and AWGN powers are assumed unknown. We have seen that the SAEM-JCPCE outperforms the SMC algorithms in the estimation of both the channel impulse response and phase distortions. However compared to the PMAP-SMC, it should be denoted that only a slight improvement is achieved, especially at high SNR. Because of both the computational cost and the high memory requirements of the SAEM, the on-line SMC methods and particularly the PMAP-SMC consequently offer a good compromise between performances and complexity. More precisely, the proposed strategy in the PMAP-SMC algorithm for parameter estimation significantly outperforms the point estimation method of the OEM-SMC. Such an algorithm can thus be efficiently used for the design of OFDM receivers in wireline and wireless communication systems. Moreover in order to solve the challenging problem of both on-line static and dynamic state estimation, the proposed scheme for parameter estimation can be easily extended to many applications leading to efficient particle filters.

APPENDIX A

MONTE CARLO APPROXIMATION OF THE POSTERIOR DISTRIBUTION FOR THE SAEM ALGORITHM

In this appendix, we describe the Monte Carlo filter required in the SAEM in order to approximate the posterior distribution $p(\boldsymbol{\phi}, \mathbf{h}|\mathbf{r}, \lambda_{n-1})$. Using the Bayes' theorem, this target distribution can be decomposed as :

$$p(\boldsymbol{\phi}, \mathbf{h}|\mathbf{r}, \lambda_{n-1}) \propto p(\mathbf{h}|\boldsymbol{\phi}, \mathbf{r}, \lambda_{n-1})p(\boldsymbol{\phi}|\mathbf{r}, \lambda_{n-1}) \quad (48)$$

Since $p(\mathbf{h}|\boldsymbol{\phi}, \mathbf{r}, \lambda_{n-1})$ is analytically tractable, only the posterior distribution of the phase distortions is approximated by the principle of importance sampling.

1) *Importance sampling of the phase distortions:* The marginal posterior distribution $p(\boldsymbol{\phi}|\mathbf{r}, \lambda_{n-1})$ is approximated by Monte Carlo algorithm as :

$$\hat{p}(\boldsymbol{\phi}|\mathbf{r}, \lambda_{n-1}) = \sum_{j=1}^M \tilde{w}^{(j),n} \delta(\boldsymbol{\phi} - \boldsymbol{\phi}^{(j),n}) \quad (49)$$

The phase distortions values are sampled from its prior distribution,

$$\pi(\boldsymbol{\phi}|\mathbf{r}, \lambda_{n-1}) = p(\boldsymbol{\phi}|\lambda_{n-1}) \quad (50)$$

where $p(\boldsymbol{\phi}|\lambda_{n-1})$ is obtained from the state equation (6) and is equal to

$$p(\boldsymbol{\phi}|\lambda_{n-1}) = \mathcal{N}(\boldsymbol{\phi}; \epsilon_{n-1}\boldsymbol{\mu}, \sigma_{v,n-1}^2 \boldsymbol{\Theta}) \quad (51)$$

with $\boldsymbol{\mu} = [2\pi(N-1)/N \ \dots \ 2\pi/N \ 0]^T$ and

$$\boldsymbol{\Theta} = \begin{bmatrix} N & & 2 & 1 \\ & \ddots & \vdots & \vdots \\ 2 & \dots & 2 & 1 \\ 1 & \dots & 1 & 1 \end{bmatrix}$$

Using (13), the unnormalized importance weights are given by:

$$w^{(j),n} = \frac{p(\boldsymbol{\phi}^{(j),n}|\mathbf{r}, \lambda_{n-1})}{\pi(\boldsymbol{\phi}^{(j),n}|\mathbf{r}, \lambda_{n-1})} \quad (52)$$

Since the prior importance function has been considered for phase distortions sampling, it is obvious that the importance weights are obtained as follows :

$$w^{(j),n} \propto p(\mathbf{r}|\boldsymbol{\phi}^{(j),n}, \lambda_{n-1}) \quad (53)$$

The distribution $p(\mathbf{r}|\boldsymbol{\phi}^{(j),n}, \lambda_{n-1})$ is obtained by marginalizing $p(\mathbf{r}, \mathbf{h}|\boldsymbol{\phi}^{(j),n}, \lambda_{n-1})$. Because \mathbf{h} is independent of $\boldsymbol{\phi}$ and λ ,

$$p(\mathbf{r}, \mathbf{h}|\boldsymbol{\phi}^{(j),n}, \lambda_{n-1}) = p(\mathbf{r}|\mathbf{h}, \boldsymbol{\phi}^{(j),n}, \lambda_{n-1})p(\mathbf{h}) \quad (54)$$

where

$$p(\mathbf{h}) = \mathcal{N}_c\left(\mathbf{0}_{L \times 1}, \frac{1}{L}\mathbf{I}_L\right) \quad (55)$$

and the likelihood distribution is given by:

$$p(\mathbf{r}|\mathbf{h}, \boldsymbol{\phi}^{(j),n}, \lambda_{n-1}) = \mathcal{N}_c\left(\mathbf{r}; \mathbf{C}^{(j),n}\mathbf{h}, \sigma_{b,n-1}^2 \mathbf{I}_N\right) \quad (56)$$

with

$$\mathbf{C}^{(j),n} = \begin{bmatrix} e^{j\phi_{N-1}^{(j),n}} & & 0 \\ & \ddots & \\ 0 & & e^{j\phi_0^{(j),n}} \end{bmatrix} \begin{bmatrix} \mathbf{s}_{N-1}^T \\ \vdots \\ \mathbf{s}_0^T \end{bmatrix}$$

As a consequence, we obtain :

$$w^{(j)} \propto \mathcal{N}_c(\mathbf{r}; \mathbf{0}_{N \times 1}, \sigma_{b,n-1}^2 \mathbf{I}_N + \frac{\mathbf{C}^{(j),n} (\mathbf{C}^{(j),n})^H}{L}) \quad (57)$$

2) *Posterior distribution of the channel:* Since the CIR \mathbf{h} has a circular gaussian *a priori* pdf, the posterior pdf $p(\mathbf{h}|\boldsymbol{\phi}^{(j),n}, \mathbf{r}, \lambda_{n-1})$ of the j -th particle can be written as :

$$p(\mathbf{h}|\boldsymbol{\phi}^{(j),n}, \mathbf{r}, \lambda_{n-1}) \propto p(\mathbf{r}|\mathbf{h}, \boldsymbol{\phi}^{(j),n}, \lambda_{n-1})p(\mathbf{h}) \quad (58)$$

where $p(\mathbf{h})$ and $p(\mathbf{r}|\mathbf{h}, \boldsymbol{\phi}^{(j),n}, \lambda_{n-1})$ are respectively given by (55) and (56)

Consequently, the *a posteriori* pdf (58) of \mathbf{h} is :

$$p(\mathbf{h}|\boldsymbol{\phi}^{(j),n}, \mathbf{r}, \lambda_{n-1}) = \mathcal{N}_c \left(\mathbf{h}; \mathbf{h}^{(j),n}, \boldsymbol{\Sigma}^{(j),n} \right) \quad (59)$$

with :

$$\begin{aligned} \mathbf{h}^{(j),n} &= \frac{1}{L} \left(\mathbf{C}^{(j),n} \right)^H \\ &\times \left(\frac{1}{L} \mathbf{C}^{(j),n} \left(\mathbf{C}^{(j),n} \right)^H + \sigma_{b,n-1}^2 \mathbf{I}_N \right)^{-1} \mathbf{r} \end{aligned} \quad (60)$$

and,

$$\begin{aligned} \boldsymbol{\Sigma}^{(j)} &= \frac{1}{L} \mathbf{I}_L - \frac{1}{L^2} \left(\mathbf{C}^{(j),n} \right)^H \\ &\times \left(\frac{1}{L} \mathbf{C}^{(j),n} \left(\mathbf{C}^{(j),n} \right)^H + \sigma_{b,n-1}^2 \mathbf{I}_N \right)^{-1} \mathbf{C}^{(j),n} \end{aligned} \quad (61)$$

APPENDIX B

DERIVATION OF THE OPTIMAL IMPORTANCE FUNCTION FOR PHN SAMPLING

The optimal importance function for PHN sampling in the proposed particle filter requires the derivation of the pdf defined in (28). This pdf can be rewritten as :

$$\begin{aligned} p(\phi_k | \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k}) &= p(r_k | \phi_k, \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k-1}) \\ &\times \frac{p(\phi_k | \phi_{0:k-1}^{(j)}, \lambda^{(j)})}{p(r_k | \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k-1})} \end{aligned} \quad (62)$$

with

$$p(\phi_k | \phi_{0:k-1}^{(j)}, \lambda^{(j)}) = \begin{cases} \mathcal{N}(\phi_0; 0, \sigma_v^2), & \text{for } k = 0 \\ \mathcal{N}(\phi_k; \phi_{k-1}^{(j)} + 2\pi\epsilon^{(j)}/N, \sigma_v^2), & \text{for } k > 0 \end{cases} \quad (63)$$

and,

$$p(r_k | \phi_k, \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k-1}) = \int p(\mathbf{h} | \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k-1}) \times p(r_k | \mathbf{h}, \phi_k, \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k-1}) d\mathbf{h} \quad (64)$$

Finally, the normalization term which does not depend on ϕ_k can be written as :

$$p(r_k | \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k-1}) = \int p(\phi_k | \phi_{0:k-1}^{(j)}, \lambda^{(j)}) \times p(r_k | \phi_k, \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k-1}) d\theta_k \quad (65)$$

The summand in (64) is the product of the likelihood pdf and the posterior pdf of the CIR which are defined respectively by :

$$p(r_k | \mathbf{h}, \phi_k, \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k-1}) = \mathcal{N}_c \left(r_k; e^{j\phi_k} \mathbf{s}_k^T \mathbf{h}, \sigma_b^{2(j)} \right) \quad (66)$$

and,

$$p(\mathbf{h} | \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k-1}) = \mathcal{N}_c \left(\mathbf{h}; \mathbf{h}_{k-1}^{(j)}, \boldsymbol{\Sigma}_{k-1}^{(j)} \right) \quad (67)$$

Therefore using (66) and (67), it is straightforward to show that the expression described in (64) can be rewritten as :

$$p(r_k | \phi_k, \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k-1}) = \mathcal{N}_c(r_k; \rho_k^{(j)}, \chi_k^{(j)}) \quad (68)$$

where $\rho_k^{(j)} = e^{j\phi_k} \mathbf{s}_k^T \mathbf{h}_{k-1}^{(j)}$ and $\chi_k^{(j)} = \mathbf{s}_k^T \boldsymbol{\Sigma}_{k-1}^{(j)} \mathbf{s}_k + \sigma_b^{2(j)}$. According to (63)-(68), an analytical form of (62) remains untractable due to the double exponential in (68) by $\rho_k^{(j)}$. However, by linearizing the noise term v_k in (6), the mean of (68) is approximated by :

$$\rho_k^{(j)} \approx \begin{cases} (1 + jv_0) \mathbf{s}_0^T \mathbf{h}_1^{(j)}, & \text{for } k = 0 \\ (1 + jv_k) e^{j(\phi_{k-1}^{(j)} + 2\pi\epsilon^{(j)}/N)} \mathbf{s}_k^T \mathbf{h}_{k-1}^{(j)}, & \text{for } k > 0 \end{cases} \quad (69)$$

This approximation holds when the phase noise rate is small and is more accurate than the usual approximation $e^{j\theta_k} \approx 1 + j\theta_k$. Using (63), (68) and (69) and after several algebraic manipulations, (62) can be simplified as :

$$p(\phi_k | \phi_{0:k-1}^{(j)}, \lambda^{(j)}, r_{0:k}) \approx \mathcal{N}(\phi_k; \vartheta_k^{(j)}, \Lambda_k^{(j)}) \quad (70)$$

where $\Lambda_k^{(j)} = \frac{\chi_k^{(j)} \sigma_v^{2(j)}}{|\Gamma_k^{(j)}|^2 \sigma_v^{2(j)} + \chi_k^{(j)}}$ and

TABLE IV
STUDY OF THE GAUSSIAN APPROXIMATION OF THE OPTIMAL IMPORTANCE FUNCTION.

		10^{-3}			10^{-2}		
		SNR (dB)					
		0	20	40	0	20	40
Gaps $D_k^{(j)}$	min	0	0	0	0	0	0
	mean	6.3948e-7	1.27e-6	1.4e-6	6.62e-7	1.65e-6	1.848e-6
	max	1.15e-4	3.83e-4	3.94e-4	1.32e-4	3.86e-4	4.1e-4

$$\vartheta_k^{(j)} = \begin{cases} \Upsilon_0^{(j)}, & \text{for } k = 0 \\ \Upsilon_k^{(j)} + \phi_{k-1}^{(j)} + 2\pi\epsilon^{(j)}/N, & \text{for } k > 0 \end{cases} \quad (71)$$

with $\chi_k^{(j)} = \mathbf{s}_k^T \boldsymbol{\Sigma}_{k-1}^{(j)} \mathbf{s}_k^* + \sigma_b^{2(j)}$, $\Upsilon_k^{(j)} = \frac{\Im(\Gamma_k^{(j)*} r_k) \sigma_v^{2(j)}}{|\Gamma_k^{(j)}|^2 \sigma_v^{2(j)} + \chi_k^{(j)}}$ (with $\Im(\cdot)$ the imaginary part) and

$$\Gamma_k^{(j)} = \begin{cases} \mathbf{s}_0^T \mathbf{h}_{-1}^{(j)}, & \text{for } k = 0 \\ e^{j(\phi_{k-1}^{(j)} + 2\pi\epsilon^{(j)}/N)} \mathbf{s}_k^T \mathbf{h}_{k-1}^{(j)}, & \text{for } k > 0 \end{cases} \quad (72)$$

APPENDIX C

SUITABILITY OF THE GAUSSIAN APPROXIMATION OF THE IMPORTANCE FUNCTION

As remarked in Section V-B, the accuracy of the Gaussian approximation should be studied since the likelihood function is multimodal and the prior density has an infinite support [39]. In order to study precisely the validity of the Gaussian approximation, Table IV presents the gaps between the exact and the approximate cumulative distributions given for the k -th time sample and for the j -th particle by :

$$D_k^{(j)} = \sup_{\phi_k} \left| F(\phi_k | \phi_{k-1}^{(j)}, \lambda^{(j)}, r_k) - F^*(\phi_k | \phi_{k-1}^{(j)}, \lambda^{(j)}, r_k) \right| \quad (73)$$

where $F(\cdot)$ and $F^*(\cdot)$ are the cumulative distributions of the exact and the Gaussian approximate optimal importance function. The exact optimal importance function is obtained using Eq. (62) and a sufficiently dense discretization of the state space of ϕ_k .

From these gaps between the two cumulative distributions, we note that the proposed Gaussian approximation leads to an accurate approximation of the optimal importance function even if the optimal importance function is multimodal. Indeed, although the optimal importance function is multimodal, only one mode is significant due to the narrowness of the prior distribution of the phase distortions. As a consequence, the Gaussian approximation fits the optimal importance function.

REFERENCES

- [1] T. Pollet, M. V. Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun.*, vol. 43, pp. 191–193, Feb. 1995.
- [2] L. Tomba, "On the effect of Wiener phase noise in OFDM systems," *IEEE Trans. Commun.*, vol. 46, pp. 580–583, May 1998.
- [3] C. Garnier, L. Clavier, Y. Delignon, M. Loosvelt, and D. Boulinguez, "Multiple access for 60 GHz mobile ad hoc networks," in *Proc. IEEE VTC Spring*, vol. 3, May 2002, pp. 1517–1521.
- [4] S. Wu and Y. Bar-Ness, "A new phase noise mitigation method in OFDM systems with simultaneous CPE and ICI correction," in *Proc. MCSS*, Sep. 2003.
- [5] D. Petrovic, W. Rave, and G. Fettweis, "Inter-carrier interference due to phase noise in OFDM - estimation and suppression," in *Proc. IEEE VTC Fall*, Sep. 2004.
- [6] D. D. Lin and T. J. Lim, "The variational inference approach to joint data detection and phase noise estimation in OFDM," *IEEE Trans. Signal Process.*, vol. 55, pp. 1862–1874, May 2007.
- [7] K. Nikitopoulos and A. Polydoros, "Compensation schemes for phase noise and residual frequency offset in OFDM systems," in *Proc. IEEE GLOBECOM*, vol. 1, Nov. 2002, pp. 330–333.
- [8] S. Wu and Y. Bar-Ness, "OFDM channel estimation in the presence of frequency offset and phase noise," in *Proc. IEEE ICC*, vol. 5, May 2003, pp. 3366–3370.
- [9] D. Lin, R. Pacheco, T. Lim, and D. Hatzinakos, "Joint estimation of channel response, frequency offset and phase noise in OFDM," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3542–3554, Sep. 2006.
- [10] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc. B*, vol. 39, no. 1, pp. 1–38, 1977.
- [11] G. Wei and M. Tanner, "A Monte Carlo implementation of the em algorithm and the poor man's data augmentation algorithms," *J. Amer. Stat. Assoc.*, vol. 85, pp. 699–704, 1990.
- [12] C. Andrieu, N. de Freitas, A. Doucet, and M. Jordan, "An introduction to MCMC for machine learning," *Machine Learning*, vol. 50, pp. 5–43, 2003.
- [13] J. Booth, J. Hobert, and W. Jank, "A survey of monte carlo algorithms for maximizing the likelihood of a two-stage hierarchical model," *Stat. Modelling*, vol. 1, pp. 333–349, May 2001.
- [14] B. Delyon, M. Lavielle, and E. Moulines, "Convergence of a stochastic approximation version of the EM algorithm," *Ann. Stat.*, vol. 27, pp. 94–128, 1999.
- [15] A. Doucet, N. D. Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, 2001.
- [16] P. M. Djuric, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Míguez, "Particle filtering," *IEEE Signal Process. Mag.*, vol. 20, no. 5, pp. 19–38, Sep. 2003.
- [17] J. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *J. Amer. Stat. Assoc.*, vol. 93, no. 443, pp. 1032–1044, 1998.
- [18] M. Bolic, P. Djuric, and S. Hong, "New resampling algorithms for particle filters," in *Proc. IEEE ICASSP*, Apr. 2003.
- [19] J. Hol, T. Schon, and F. Gustafsson, "On resampling algorithms for particle filters," in *Nonlinear Statistical Signal Processing Workshop*, Cambridge, United-Kingdom, Sep. 2006.
- [20] M. Bolic, P. Djuric, and S. Hong, "Resampling algorithms for particle filters: A computational complexity perspective," *EURASIP J. Appl. Signal Process.*, vol. 15, pp. 2267–2277, 2004.
- [21] R. Chen and J. S. Liu, "Mixture Kalman filters," *J. Roy. Stat. Soc.*, vol. 62, pp. 493–508, 2000.

- [22] T. Schon, F. Gustafsson, and P. Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *IEEE Trans. Signal Process.*, vol. 53, pp. 2279–2289, Jul. 2005.
- [23] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statist. Comput.*, vol. 10, pp. 197–208, 2000.
- [24] A. Kong, J. Liu, and W. Wong, "Sequential imputations and Bayesian missing data problems," *J. Amer. Stat. Assoc.*, vol. 89, pp. 278–288, 1994.
- [25] O. Cappé and E. Moulines, "On the use of particle filtering for maximum likelihood parameter estimation," in *Proc. EUSIPCO*, Turkey, Sep. 2005.
- [26] A. Doucet and V. Tadic, "Parameter estimation in general state-space models using particle methods : New trends in statistical information processing," *Ann. Inst. Stat. Math.*, vol. 55, no. 2, pp. 409–422, 2003.
- [27] C. Andrieu, A. Doucet, and V. Tadic, "On-line parameter estimation in general state-space models," in *Proc. IEEE CDC-ECC*, Dec. 2005, pp. 332–337.
- [28] C. Andrieu and A. Doucet, "Online expectation-maximization type algorithms for parameter estimation in general state space models," in *Proc. IEEE ICASSP*, Apr. 2003.
- [29] F. Cérou, F. L. Gland, and N. Newton, "Stochastic particle methods for linear tangent filtering equations," in *Optimal Control and PDE's - Innovations and Applications. In honor of Alain Bensoussan on the occasion of his 60th birthday*, J.-L. Menaldi, E. Rofman, and A. Sulem, Eds. IOS Press, 2001, pp. 231–240.
- [30] R. Redner and H. Walker, "Mixture densities, maximum likelihood and the EM algorithm," *Siam Review*, vol. 26, no. 2, pp. 195–239, 1984.
- [31] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-gaussian Bayesian state estimation," *IEE Proceedings-F*, vol. 140, pp. 107–113, Apr. 1993.
- [32] J. Liu and M. West, "Combined parameter and state estimation in simulation-based filtering," in *Sequential Monte Carlo in Practice*, A. Doucet, J. de Freitas, and N. Gordon, Eds. New-York : Springer-Verlag, 2001, pp. 197–223.
- [33] C. Andrieu, A. Doucet, S. Singh, and V. Tadic, "Particle methods for change detection, system identification, and control," in *Proceedings of the IEEE*, vol. 92, Mar. 2004, pp. 423–438.
- [34] B.-N. Vo, B.-T. Vo, and S. Singh, "Sequential Monte Carlo methods for static parameter estimation in random set models," in *Proc. IEEE ISSNIP*, 2004, pp. 313–318.
- [35] N. Chopin, "A sequential particle filter method for static models," *Biometrika* 89, pp. 539–552, 2002.
- [36] C. Andrieu, N. D. Freitas, and A. Doucet, "Sequential MCMC for Bayesian model selection," in *Proc. IEEE SPW-HOS*, Jun. 1999, pp. 130–134.
- [37] A. Papavasiliou, "A uniformly convergent adaptive particle filter," *J. Appl. Probab.*, vol. 42, no. 4, pp. 1053–1068, 2005.
- [38] V. Zaritskii, V. Svetnik, and L. Shimelevich, "Monte Carlo technique in problems of optimal data processing," *Automation and remote control*, pp. 95–103, 1975.
- [39] N. Vaswani, "PF-EIS & PF-MT: New particle filtering algorithms for multimodal observation likelihoods and large dimensional state spaces," in *Proc. IEEE ICASSP*, Honolulu, Hawaii, Apr. 2007, pp. 1193–1196.
- [40] H. L. van Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968.
- [41] P. Tichavsky, "Posterior cramer-rao bounds for adaptive harmonic retrieval," *IEEE Trans. Signal Process.*, vol. 43, pp. 1299–1302, May 1995.



François Septier was born in Doullens, France, in 1981. He received the Engineer degree in electrical engineering and signal processing in 2004 from TELECOM Lille 1, France, and the M.S. degree in digital communications from the University of Valenciennes, France, in 2004.

Since 2004, he is a Ph.D. student at IEMN-DOAE, UMR 8520, University of Valenciennes. His research interests include statistical signal processing for communications. More precisely, his work is focused on the compensation of phase distortions in multicarrier systems using sequential Monte-Carlo methods.



Yves Delignon received in 1990 the M.S. degree from the Université de Technologies de Compiègne, France. He joined Telecom Bretagne, Brest, France during the same year for preparing a Ph.D. about statistical modelling of radar images and received the PhD degree from the Université de Rennes 1, France in 1993.

In 1992 he joined the engineering school TELECOM Lille 1. From 1993 to 2003 he was assistant Professor and conducted his research in the field of the statistical radar image processing. In 2002, he received the HdR from the University of Lille 1 and since 2004 he became full Professor at Telecom Lille 1. From that date, he focuses his research on statistical signal processing for digital communication. In 2005 he became dean of research at Telecom Lille 1. His teaching activity mainly concerns statistical signal processing and mobile communication.



Atika Menhaj-Rivenq was born in 1970 in Marrakech, Morocco. She received her Diploma of Engineering and the M.S. degree in 1993 and then her Ph.D. degree in 1996, from the University of Valenciennes (France).

She actually is Professor in electronics at this University. Her primary interest is in signal processing applied to intelligent transportation systems and telecommunication systems.



Christelle Garnier received the M.S. and the Ph.D. degrees in signal processing and telecommunications from the University of Rennes, France, in 1995 and 1999 respectively.

Since 2000, she has been with the Department of Communication Systems at the engineering school TELECOM Lille 1 as an Assistant Professor in mobile communications. Her research interests include multi-carrier and spread spectrum transmission schemes, in particular the effects of phase noise and the strategies for resource allocation.