N° ordre : 08/10

UNIVERSITÉ DE VALENCIENNES ET DU HAINAUT-CAMBRÉSIS

ECOLE DOCTORALE RÉGIONALE SPI - LILLE NORD-DE-FRANCE

THÈSE

présentée en vue d'obtenir le grade de

DOCTEUR

Specialité « Electronique »

par

François Septier

Méthodes séquentielles de Monte-Carlo pour les systèmes multiporteuses en présence de distorsions de phase.

Thèse soutenue le 14 mai 2008 devant le jury composé de :

M.	MANUEL DAVY	Chargé de Recherche CNRS, LAGIS	*Président du jury*
M.	THIERRY CHONAVEL	Professeur, TELECOM Bretagne	*Rapporteur*
M.	LUC DENEIRE	Maître de conférences HDR, Université de Nice Sophia Antipolis - I3S	*Rapporteur*
M.	SIMON GODSILL	Professeur, Université de Cambridge (UK)	*Examinateur*
M^{me}	ATIKA MENHAJ-RIVENQ	Professeur, Université de Valenciennes - IEMN-DOAE	*Directeur de thèse*
M.	YVES DELIGNON	Professeur, Institut TELECOM/TELECOM SudParis/TELECOM Lille1	*Directeur de thèse*

Thèse préparée à l'Institut d'Electronique, de Microélectronique et de Nanotechnologie (IEMN, UMR CNRS 8520) de Valenciennes et à TELECOM Lille1

REMERCIEMENTS

Je voudrais tout d'abord exprimer toute ma gratitude à Madame Atika Menhaj-Rivenq, Professeur à l'UVHC, et Monsieur Yves Delignon, Professeur à TELECOM Lille1, pour le co-encadrement de cette thèse et leurs conseils scientifiques. Merci pour la sympathie et la confiance qu'ils ont eus envers moi durant ces trois années de thèse.

Je remercie Messieurs Thierry Chonavel, Professeur à TELECOM Bretagne, et Luc Deneire, Maître de Conférences-HDR à l'université de Sophia Antipolis, rapporteurs de ce travail de thèse, pour l'intérêt porté à mes travaux et la lecture détaillée de ce mémoire de thèse. Leurs remarques pertinentes ont permis d'améliorer le manuscript final et également de construire un exposé de thèse clair et concis. Je remercie également Messieurs Manuel Davy, Chargé de Recherche au LAGIS, et Simon Godsill, Professeur à l'université de Cambridge, de participer à ce jury de thèse.

Je remercie Christophe Séguinot, responsable du département Systèmes de Communications à TELECOM Lille 1, ainsi que Bertrand Nongaillard, Responsable du département OAE à l'IEMN de Valenciennes, pour m'avoir permis d'accomplir ces travaux dans d'excellentes conditions. J'étends bien évidemment ces remerciements aux autres membres du département ainsi qu'à tout le personnel de TELECOM Lille1 et de l'IEMN-DOAE pour la qualité de leur accueil.

Ensuite, je tiens à remercier chaleureusement les nombreuses personnes que j'ai eu la chance de rencontrer pendant cette thèse. Je pense en premier lieu aux compagnons de route du GRIF avec qui j'ai passé de merveilleux moments, humainement inoubliables. Merci donc Rodrigue, Adrien, Matthieu, Manu, Karim et Hassan pour votre générosité, votre gentillesse, votre écoute, votre disponibilité et votre soutien... Je n'oublierai jamais les nombreuses discussions et moments de partage que l'on a eus souvent autour d'un bon café. Je tiens également à remercier entre autres Arnaud, Chafik, Julien, Nizar et David.

Enfin, je remercie ma mère, Jacqueline, et mon père, Alain, dont leur amour et leur disponibilité me rendent fier d'être leur fils, ma sœur Aurélia ainsi que son époux Sébastien, ma filleule Roxane et mon neveu Tristan pour m'avoir entouré. Je conclue en remerciant tout particulièrement ma fiancée, Ambre, pour avoir toujours cru en moi et pour m'avoir accompagné et soutenu par son amour. Sans toi cette thèse ne serait pas ce qu'elle est, je t'aime...

Valenciennes, le 15 juillet 2008.

CONTENTS

LIST OF FIGURES

NOTATIONS

$\mathbb{E}[.]$	Expectation
$\mathcal{N}(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma})$	Real Gaussian random vectors with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$
$\mathcal{N}_c(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma})$	Circularly symmetric complex Gaussian random vectors with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$
$\mathbf{I}_n$	$n \times n$ identity matrix
$\mathbf{0}_{n\times m}$	$n \times m$ matrix of zeros
x	Scalar
$\mathbf{x}$	Vector
$\mathbf{X}$	Matrix
$(.)^*$	Conjugate
$(.)^T$	Transpose
$(.)^H$	Hermitian transpose
$\mathrm{tr}[.]$	Matrix trace
$\Re(.)$	Real part
$\Im(.)$	Imaginary part

N	Number of subcarriers
N_{cp}	Cyclic prefix length
T	Multicarrier symbol duration excluding the cyclic prefix
L_p	Number of channel paths
M	Number of particles in SMC algorithms

ABBREVIATIONS

AR	Autoregressive
AWGN	Additive white Gaussian noise
BER	Bit error rate
BPSK	Binary phase shift keying
CDF	Cumulative distribution function
CDMA	Code division multiple access
CFO	Carrier frequency offset
CIR	Channel impulse response
CPE	Common phase error
DFT	Discrete Fourier transform
e.g.	Latin expression *exempli gratia* which means for example
EM	Expectation maximization
EKF	Extended Kalman filter
FDMA	Frequency division multiple access
FFT	Fast Fourier transform
HMM	Hidden Markov model
i.e.	Latin expression *id est* which means that is/in other words
i.i.d.	Independent and identically distributed
ICI	Inter-carrier interference
IDFT	Inverse DFT
IFFT	Inverse FFT
ISI	Inter-symbol interference
KL	Kullback-Leibler
MAI	Multiple access interference
MAP	Maximum a posteriori
MC	Monte Carlo
MC-CDMA	Muticarrier CDMA
MC-DS-CDMA	Multicarrier direct sequence CDMA
ML	Maximum likelihood
MCEM	Monte Carlo EM
MCMC	Markov chain Monte Carlo
MMSE	Minimum mean square error

MSE	Mean square error
OEM	On-line EM
OFDM	Orthogonal frequency division multiplexing
p.d.f.	Probability density function
PHN	Phase noise
PSEM	Particle stochastic EM
QAM	Quadrature amplitude modulation
QPSK	Quaternary phase shift keying
SEM	Stochastic EM
SIS	Sequential importance sampling
SMC	Sequential Monte Carlo
SNR	Signal-to-noise ratio
TDMA	Time division multiple access
WLAN	Wireless local area network
WMAN	Wireless metropolitan area network

GENERAL INTRODUCTION

THE need for mobile broadband communications has increased rapidly in recent years placing new demands for wireless networks. The demand for high data rate services with a high *quality of services* (QoS) is continuing to increase without slowdown in sight. These services require a very reliable data transmission over severe environments. Indeed, most of communication systems, including both wired and wireless systems, experience many degradations such as noise, attenuation, multi-path effects, interference and nonlinearities.

Multicarrier transmission systems have aroused great interest in recent years as a potential solution to the problem of transmitting high data rate over a frequency selective fading channel [Bingham 90]. Nowadays, multicarrier modulation is being selected as the transmission scheme for the majority of new communication systems [Wang *et al.* 00]. Examples include *Digital Subscriber Line* (DSL), *European Digital Video Broadcast* (DVB), *Digital Audio Broadcast* (DAB), *Wireless Local Area Networks* (WLAN) standards (IEEE 802.11, MMAC and Hiperlan/2) and *Wireless Metropolitan Area Network* (WMAN) standards (IEEE 802.16). Unlike traditional single-carrier systems, the used frequency band is subdivided into a large set of narrow subbands, or subcarriers, which are spaced very closely together. Each such subband carries an independent low-rate data stream. Because of the large number of such subbands, the aggregate data rate is high. Besides the resulting high spectral efficiency, its immunity to *Inter-Symbol Interference* (ISI) is another important benefit of multicarrier systems.

Unfortunately, multicarrier systems are very sensitive to *phase noise* (PHN) and *carrier frequency offset* (CFO) caused by the oscillator instabilities [Pollet *et al.* 95, Costa *et al.* 02, Tomba 98, Tomba *et al.* 99, Steedam *et al.* 01, Garnier *et al.* 02]. Indeed, random time-varying phase distortions destroy the orthogonality of subcarriers and lead after the discrete Fourier transform (DFT) both to rotation of every subcarrier by a random phase, called *common phase error* (CPE), and to *inter-carrier interference* (ICI). These phase impairments reduces drastically the system performance unless efficient compensation techniques are implemented.

In literature, many approaches have been proposed to perform data detection in OFDM systems by compensating PHN, or both PHN and residual CFO [Robertson *et al.* 95, Yee *et al.* 05, Petrovic *et al.* 04a, Wu *et al.* 04, Petrovic *et al.* 04b, Wu *et al.* 03a, Nikitopoulos *et al.* 05, Casas *et al.* 02, Lin *et al.* 05, Lin *et al.* 07]. Nevertheless, the channel impulse response (CIR) is always assumed perfectly known prior to phase distortions mitigation. In fact, as in many standard based on multicarrier modulation such as Hiperlan/2 or IEEE 802.11a, training OFDM symbols which are

known by the receiver are firstly used for channel estimation before the data transmission. Indeed, since the CIR changes slowly with respect to the OFDM symbol rate, the channel is thus only estimated at the beginning of a frame. Then, its estimate is used for data detection in the payload section. OFDM channel estimator in the presence of both PHN and CFO has been proposed in [Wu *et al.* 03b, Lin *et al.* 06].

However, these existing approaches suffer from several drawbacks. On the one hand, for the channel estimation problem, both the AWGN and PHN powers are assumed known to the receiver in existing schemes, which is not a realistic assumption. On the other hand, for data detection, efficient algorithms are based on a decision-directed scheme which consists in making a tentative decision over the transmitted signal still corrupted by phase imperfections. Consequently for significant phase distortions, noise-induced symbol decision errors may propagate through the feedback loop, leading to poor estimator performance.

As a consequence, in this thesis, we focus on the design of an efficient multicarrier receiver based on Bayesian inference in order to improve the performance and reliability of multicarrier transmission in the presence of phase distortions. This task proved to be especially difficult due to the nonlinear behavior of phase distortions but can be greatly facilitated by the use of an efficient signal processing technique such as the Sequential Monte Carlo (SMC), also called particle filtering, which recently emerged in the field of statistics and engineering [Doucet *et al.* 01, Djuric *et al.* 03].

The Sequential Monte Carlo methods are a set of powerful simulation-based algorithms to perform optimal state estimation in nonlinear non-Gaussian state space models. The approach has recently received a lot of interest since it allows a large number of challenging non-linear estimation tasks to be addressed in an efficient online manner [Doucet *et al.* 01]. The idea is to approximate the posterior distribution of interest by a set of weighted points in the sample space, called particles, which evolve randomly in time according to a simulation-based rule, and either give birth to offspring particles or die according to their ability to represent the different zones of interest of the state space.

By using the packet structure of existing standards like Hiperlan/2 or IEEE802.11a, we propose a multicarrier receiver based on SMC methodology which consists in two separate estimation problems. More precisely, we firstly propose an OFDM channel estimator in the presence of phase distortions from a training multicarrier symbol which also deals with the unknowledge of both PHN and AWGN powers. Finally, a joint multicarrier signal and phase distortions estimator is proposed. Nevertheless, although SMC methods are well adapted to problems of estimating a sequence of potentially quickly varying distributions whose dimension is increasing over time, parameter estimation using particle filtering is still a major issue. Since the estimation problem we deal with consists in estimating both dynamic states and static parameters, this task has to be considered with care. After reviewing existing solutions, we propose an original parameter estimation strategy which possesses both good stability property and

efficient convergence rate due to the use of a deterministic method for each particle trajectory.

The rest of this document is organized as follows. Chapter 1 provides an introduction to multicarrier systems. The traditional OFDM system with cyclic prefix is presented before introducing multiple access strategy in OFDM-based systems. Some real characteristics of real systems are also discussed and a specific example of an OFDM system (Hiperlan/2) is provided. Then after introducing the phase distortions including both CFO and PHN, we study the phase impairment impact on multicarrier system performance. These performance results clearly highlight the importance of having a phase distortion compensation scheme in multicarrier receiver. Finally, existing algorithms dealing with phase impairments in OFDM systems are reviewed.

Chapter 2 is devoted to Bayesian inference. The mathematical foundation of the Bayesian theory is introduced. Then, some deterministic methods for state inference are presented. A particular emphasis is placed on the Sequential Monte Carlo approach, and the basic ideas of particle filtering techniques and some strategies for their efficient implementation are discussed.

Chapter 3 addresses the parameter estimation problem using SMC methods. After reviewing off-line parameter estimation methods based on SMC methods, we describe the existing approaches for on-line parameter estimation. Then, we introduce the proposed strategy. Finally, simulation results are presented in two different models and performance comparisons between the proposed method and the on-line Expectation maximization [Andrieu *et al.* 03a, Andrieu *et al.* 05] are assessed.

In Chapter 4, we focus on the first step of the proposed multicarrier receiver in the presence of both PHN and CFO which consists in estimating the channel impulse response from a training sequence. Unlike existing works, we consider that both AWGN and PHN powers are assumed unknown at the receiver side. After introducing the dynamic state-space model, the proposed SMC filter is then described. The efficiency of the proposed algorithms is assessed and strategies for parameter estimation are compared.

In Chapter 5, we deal with the major problem of data detection in multicarrier systems in the presence of PHN and CFO. After a brief description of the observed signal, the unknown multicarrier signal is then statistically studied leading to its original autoregressive modeling. All available information concerning the states of interest is finally summed up in a DSS representation. The proposed SMC filter is described and the posterior Cramér-Rao bound is derived. Numerical results are given to demonstrate the validity of our approach. Performances of the proposed particle filter algorithm are assessed in different system configurations and are compared to those of existing schemes.

We finish this manuscript by drawing some concluding remarks and by providing an overview of some research directions that should be pursued in order to complete the work presented in this thesis.

INTRODUCTION GÉNÉRALE EN FRANÇAIS

LES communications tendent de plus en plus vers une globalisation des services (vidéoconférence, transfert de fichiers volumineux, applications, internet, ...). Dans ce contexte, la demande de services à hauts débits avec une grande qualité de services ne cesse d'augmenter. Ces services demandent une transmission fiable des données à travers un environnement souvent sévère. En effet, la plupart des systèmes de communication filaire ou sans fil sont soumis à de nombreuses dégradations telles que par exemple le bruit, l'atténuation, l'effet multi-trajets, les interférences et les non linéarités.

Les systèmes de transmission multiporteuses ont suscité un grand intérêt dans la communauté scientifique depuis ces dernières années afin de résoudre le problème de transmission de données à haut débit à travers un canal sélectif en fréquence [Bingham 90]. De nos jours, la technique multiporteuses a été selectionnée dans la majorité des nouveaux systèmes de communication [Wang *et al.* 00] comme la diffusion du son numérique sans fil (*Digital Audio Broadcast* - DAB), la télévision numérique sans fil (*European Digital Video Broadcast* - DVB), les communications numériques hauts débits filaires (*Digital Subscriber Line* - DSL) et les réseaux sans fil (Hiperlan/2, WiFi IEEE 802.11, MMAC et WiMAX IEEE 802.16). Contrairement aux traditionnels systèmes monoporteuse, la bande de fréquence allouée est divisée en un grand nombre de sous bandes étroites centrées sur des sous porteuses très proches. Chaque sous bande transporte un flux d'information à bas débit. Grâce à ce grand nombre de sous bandes, le débit global est ainsi élevé. Hormis la grande efficacité spectrale résultante du principe du système, son immunité face aux *Interferences Entre Symboles* (IES) est un des autres principaux avantages des système multiporteuses.

Malheureusement, ces systèmes multiporteuses sont extrêmement sensibles aux distorsions de phase, comme le *bruit de phase* et le *décalage fréquentiel de la porteuse*, engendrées par l'instabilité des oscillateurs locaux présent dans la chaîne de transmission [Pollet *et al.* 95, Costa *et al.* 02, Tomba 98, Tomba *et al.* 99, Steedam *et al.* 01, Garnier *et al.* 02]. En effet, ces distorsions de phase aléatoires et variant dans le temps détruisent l'orthogonalité des sous porteuses, entraînant après la transformation de Fourier du récepteur à la fois une rotation de phase commune à toutes les porteuses ainsi qu'une interférence entre porteuses. Ces imperfections de phase réduisent sévèrement les performances du système à moins qu'une technique efficace de compensation soit implémentée.

Dans la littérature scientifique, de nombreuses approches ont été proposées afin d'accomplir la détection des données dans les systèmes multiporteuses en compensant le bruit de phase, ou à la fois le bruit de phase et le décalage résiduel de la porteuse [Robertson *et al.* 95, Yee *et al.* 05, Petrovic *et al.* 04a, Wu *et al.* 04, Petrovic *et al.* 04b, Wu *et al.* 03a, Nikitopoulos *et al.* 05, Casas *et al.* 02, Lin *et al.* 05, Lin *et al.* 07]. Néanmoins, la réponse impulsionnelle du canal de propagation est toujours supposée parfaitement connue par l'algorithme destiné à la compensation des distorsions de phase. En fait, comme dans la plupart des standards actuels basés sur les systèmes multiporteuses tels que Hiperlan/2 ou encore IEEE 802.11a, une séquence d'apprentissage, constituée d'un ou plusieurs symboles OFDM, connue par le récepteur est dans un premier temps utilisée pour l'estimation du canal avant la transmission des données. En effet, comme la réponse impulsionnelle du canal varie lentement en comparaison au débit d'un symbole OFDM, le canal peut donc être seulement estimé en début de trame. Ensuite, son estimée est utilisée pour la détection des données. Dans cet esprit, un estimateur de canal pour un système OFDM en présence à la fois de bruit de phase et de décalage fréquentiel de la porteuse a ainsi été proposé dans [Wu *et al.* 03b, Lin *et al.* 06].

Cependant, toutes ces approches existantes souffrent de plusieurs inconvénients. D'une part, pour le problème de l'estimation du canal de propagation, les puissances du bruit additif Gaussien et du bruit de phase sont supposées parfaitement connues du récepteur, ce qui ne s'avère pas être une hypothèse très réaliste. D'autre part, pour la détection des données, les algorithmes fournissant les meilleures performances sont basés sur un schéma consistant à effectuer une tentative de décision sur le signal toujours corrompu par les imperfections de phase. En conséquence, pour des distorsions de phase significatives, les erreurs de décision peuvent se propager par la boucle de retour, entraînant ainsi de faibles performances de l'algorithme.

Dans cette thèse, notre objectif est donc de concevoir un récepteur, basé sur l'inférence Bayésienne, capable de compenser efficacement les effets néfastes de ces distortions de phase. Ce problème est particulièrement difficile due à la nature non linéaire de ces imperfections de phase mais celui-ci peut être grandement facilité grâce à l'utilisation de techniques avancées de traitement statistique du signal comme les méthodes séquentielles de Monte-Carlo (SMC), appelées également filtrage particulaire, qui ont récemment émergées dans le domaine des statistiques et de l'ingénierie [Doucet *et al.* 01, Djuric *et al.* 03].

Les méthodes séquentielles de Monte-Carlo sont de puissants algorithmes de simulation stochastique destinés à l'estimation optimale de processus dans un modèle d'état non linéaire et non Gaussien. Cette technique a récemment suscité un large intérêt puisqu'elle permet de résoudre, efficacement et de manière séquentielle, de nombreux problèmes difficiles d'estimation [Doucet *et al.* 01]. L'idée principale est d'approcher la distribution *a posteriori* de l'état à estimer par un ensemble de points pondérés dans l'espace d'état, appelés particules, évoluant aléatoirement dans le temps selon des règles de simulation. Ces particules peuvent soit donner naissance à d'autres particules

ou mourir selon leur capacité à représenter les différentes zones d'intérêt de l'espace d'état.

En utilisant la structure d'une trame des standards existants comme l'Hiperlan/2 et l'IEEE 802.11a, nous proposons ainsi un récepteur pour systèmes multiporteuses utilisant le principe des méthodes séquentielles de Monte-Carlo et consistant en deux problèmes distincts d'estimation. Plus précisément, dans un premier temps, on propose un estimateur du canal OFDM en présence à la fois de bruit de phase et de décalage fréquentiel de la porteuse en s'appuyant sur une séquence d'apprentissage. Mais contrairement aux méthodes existantes, nous supposons que les puissances à la fois du bruit additif Gaussien et du bruit de phase ne sont pas connues du récepteur. Finalement, un estimateur conjoint du signal multiporteuses et des distorsions de phase est proposé. Bien que les méthodes séquentielle de Monte-Carlo sont bien adaptées aux problèmes d'estimation de séquence variant rapidement dans le temps dont la dimension augmente également avec le temps, l'estimation de paramètres du modèle par le filtrage particulaire est encore un problème majeur. Comme le problème que nous traitons consiste à la fois en l'estimation de variables dynamiques et statiques, ce problème doit être considéré avec attention. Après la description des solutions existantes pour ce problème, nous proposons une stratégie originale d'estimation de paramètre possédant de bonnes propriétés à la fois de stabilité et de rapidité de convergence grâce notamment à l'utilisation d'une méthode déterministe associée à chaque particule du filtre.

Le reste du manuscrit est organisé comme suit. Le Chapitre 1 fournit une introduction aux systèmes multiporteuses. Dans un premier temps, le système OFDM traditionnel avec préfixe cyclique est présenté. Ensuite, nous décrivons les différentes stratégies permettant le multiple accès dans les systèmes multiporteuses. Quelques caractéristiques de systèmes réels sont également données et un exemple spécifique d'un système OFDM (Hiperlan/2) est également fournit. Puis, après avoir introduit à la fois le bruit de phase et le décalage fréquentiel de la porteuse, nous étudions l'impact de ces perturbations de phase sur les performances globales d'un système multiporteuses. Ces résultats mettent clairement en évidence l'importance d'avoir un algorithme de compensation de ces imperfections de phase au niveau du récepteur pour permettre une transmission fiable des données. Finalement, les algorithmes existants traitant le problème des distorsions de phase sont décrit.

Le Chapitre 2 est consacré à l'inférence Bayésienne. Le fondement mathématique de la théorie de Bayes est premièrement introduit. Puis, quelques méthodes déterministes pour l'inférence de l'état sont présentées. Une accentuation particulière est ensuite placée sur les méthodes séquentielles de Monte-Carlo où le concept du filtrage particulaire ainsi que quelques stratégies pour une implémentation efficace sont discutés.

Le Chapitre 3 traite du problème d'estimation des paramètres du modèle avec les méthodes séquentielles de Monte-Carlo. Après avoir décrit les méthodes exis-

tantes permettant l'estimation de paramètres de manière "off-line", nous présentons les méthodes "on-line". Puis, nous introduisons notre méthode basée sur l'association d'une méthode déterministe par particules. Finalement, des résultats de simulations sont présentés dans deux modèles différents nous permettant ainsi de comparer les performances de notre proposition avec celles de l'algorithme "on-line" d'espérance maximisation [Andrieu *et al.* 03a, Andrieu *et al.* 05].

Dans le Chapitre 4, nous nous focalisons sur la première étape de notre récepteur qui consiste en l'estimation du canal de propagation en présence du bruit de phase et du décalage fréquentiel de la porteuse depuis une séquence d'apprentissage. Contrairement aux travaux existants, nous considérons que les puissances du bruit additif Gaussien et du bruit de phase ne sont pas connues du récepteur. Après avoir introduit le modèle dynamique d'état, nous décrivons le filtre proposé. L'efficacité de notre proposition est évaluée et les différentes stratégies pour l'estimation des paramètres sont comparées.

Dans le Chapitre 5, nous traitons le problème de l'estimation des données dans les systèmes multiporteuses en présence de distorsions de phase. Après une brève description du signal observé à la réception, le signal multiporteuses dans le domaine temporel est ensuite étudié d'un point de vue statistique. Nous proposons ainsi sa modélisation par un modèle autorégressif. Toute l'information disponible est ensuite résumée dans le modèle dynamique d'état. Le filtre proposé est ensuite décrit et la borne de Cramér-Rao est fournit. Des résultats sont ensuite donnés afin d'illustrer la validité de notre approche. De nombreuses simulations sont effectuées dans des configurations différentes de système et les performances de notre algorithme sont comparées avec celles des méthodes existantes.

Finalement, nous concluons ce manuscrit et nous discutons de quelques perspectives pour de possibles futures directions de recherches.

1

MULTICARRIER SYSTEMS IN THE PRESENCE OF PHASE DISTORTIONS

IN the context of an increasing demand for higher data rate communication, multicarrier systems have received a considerable attention as they combine a high spectral efficiency with an immunity to channel dispersion.

Multicarrier systems are based on the well-studied *orthogonal frequency division multiplexing* (OFDM) modulation. The fundamental principle of OFDM originates from Chang in 1966 [Chang 66]. Nevertheless, at that time, it didn't have a particular success because of the high implementation complexity due to the use of analogical devices. In 1971, Weinstein and Ebert overcame the problem by proposing the use of the IDFT/DFT for multicarrier systems [Weinstein *et al.* 71]. Subsequently, the principle of the multicarrier modulation became the foundation of most current industry standards and in the coming broadband communication systems. Specifically, it has been chosen as solution for European digital audio and video broadcasting (DAB [ETSI Normalization Committee 97], DVB [ETSI Normalization Committee 96]). It has also been exploited for broadband wired applications : *asynchronous digital subscriber lines* (ADSL) [ANSI T1E1.4 Committee Contribution 93], *high-bit-rate digital subscriber lines* (HDSL) and the most recent *very-high-speed digital subscriber lines* (VDSL). Finally, it has been adopted for *wireless metropolitan area network* standards (WMAN - WiMAX IEEE 802.16) and *wireless local area network* (WLAN) standards in Europe (Hiperlan/2 [ETSI Normalization Committee 99]), in North America (WiFi IEEE 802.11{a,g,n}) and in Asia (MMAC).

However, multicarrier systems are very sensitive to phase noise (PHN) and carrier frequency offset (CFO) caused by oscillator instabilities [Pollet *et al.* 95, Costa *et al.* 02, Tomba 98, Tomba *et al.* 99, Steedam *et al.* 01, Garnier *et al.* 02]. Indeed, random time-varying phase distortions destroy the orthogonality of subcarriers and lead after the discrete Fourier transform (DFT) both to rotation of every subcarrier by a random phase, called *common phase error* (CPE), and to *inter-carrier interference* (ICI). These phase impairments reduces drastically the system performance unless efficient compensation techniques are implemented.

The rest of this chapter is organized as follows. The first section describes the principle of OFDM modulation and its digital implementation. Then, we discuss about

multiple access scheme in OFDM-based systems. Practical multicarrier packet structure, as used in most of existing standards, is also provided in this first section. Section 1.2 presents the phase distortions and illustrates the sensitivity of multicarrier systems to these phase impairments. Existing schemes for phase distortions compensation in multicarrier systems are reviewed in Section 1.3 and their limitations are also discussed. Finally, some concluding remarks are given.

1.1 MULTICARRIER SYSTEMS

1.1.1 OFDM system

1.1.1.1 The OFDM principle

In the OFDM scheme, the serial data stream is passed through a serial-to-parallel convertor which splits the data into a number of parallel channels. The data in each channel is applied to a modulator such that for N channels there are N modulators whose carrier frequencies are $f_0, f_1, \cdots, f_{N-1}$. OFDM uses a spacing between the frequencies equal to the Nyquist bandwidth of the parallel channel, *i.e.* $\Delta f_{\text{subcarrier}} = 1/T$ with T the duration of the OFDM symbol. Hence, the frequencies of the N subcarriers in the baseband equivalent form are given for $n = 0, \cdots, N-1$ by :

$$f_n = f_0 + n\Delta f_{\text{subcarrier}} = f_0 + n/T \tag{1.1}$$

By denoting the symbol on subcarrier k of the n-th OFDM symbol by $d_{n,k}$ and assuming a rectangular pulse shaping, the modulated OFDM signal is given by the following expression :

$$s(t) = \frac{1}{\sqrt{T}} \sum_{k=0}^{N-1} \sum_{n=-\infty}^{+\infty} d_{n,k} \text{rect}_T(t - nT) e^{j2\pi \frac{n}{T}(t-nT)} \tag{1.2}$$

with $\text{rect}_T(.)$ the rectangular function. Consequently, the power density spectrum of the OFDM signal is obtained as :

$$S(f) = \frac{1}{T} \sum_{k=0}^{N-1} |\text{sinc}(fT - k)|^2 \tag{1.3}$$

The power density spectra of the different subcarriers are depicted in Fig. 1.1. We can remark that there is no inter-carrier interference (ICI) since the power density spectrum of each subcarrier has zero at the maxima of all others subcarriers. $S(f)$ as the sum of these spectra decreases with $1/f^2$ outside the interval $[-N/(2T); N/(2T)]$. Hence, the bandwidth of the OFDM signal is given as :

$$B_{\text{OFDM}} = (N+1)/T \approx N/T \tag{1.4}$$

The approximation holds for a large number of subcarriers. OFDM has the nice property that the digital realization is very simple, since the algorithms for *Discrete Fourier Transform* (DFT) and its inverse (IDFT) can be employed [Weinstein *et al.* 71].

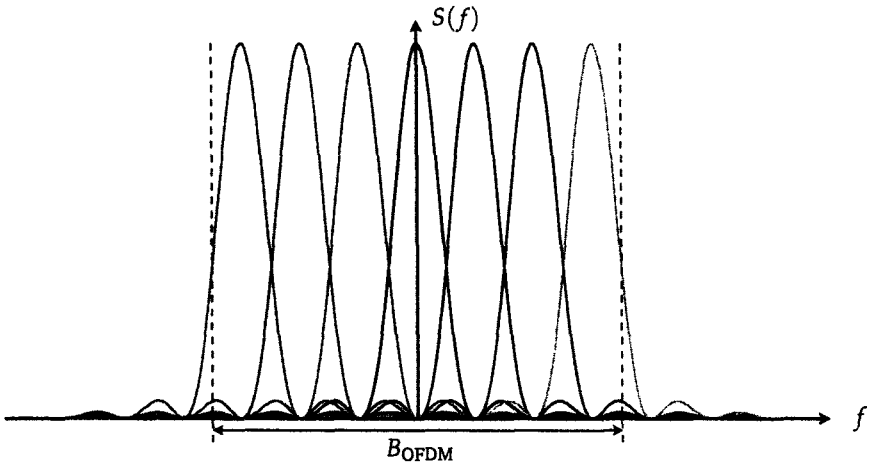

Figure 1.1 – *Power density spectrum of OFDM.*

For a further reduction of complexity, the algorithms *Fast Fourier Transform* (FFT) and *Inverse FFT* (IFFT) are generally used in practice. The IFFT eliminates the use of N oscillators and renders the OFDM transmitter implementationally attractive [Cimini 85].

1.1.1.2 *Digital implementation of an OFDM system*

As can be seen in Fig. 1.2, the serial data stream is mapped to data symbols with a symbol rate of $D = 1/T_s = N/T$. The symbol mapping function transforms the coded bits into complex-valued data symbols taking values in a discrete alphabet of finite size $\mathcal{A}$. The number of bits per symbol is thus given by :

$$\beta = \log_2(\mathcal{A}) \tag{1.5}$$

A constellation diagram is often conveniently used for representing the alphabet in the complex plane, as illustrated in Figure 1.3. The x and y coordinates of a data symbol correspond to its real and imaginary parts, respectively. There exist different kinds of alphabets, the most common ones are *Amplitude Shift Keying* (ASK), *Phase Shift Keying* (PSK), and *Quadrature Amplitude Modulation* (QAM) [Proakis 95]. The symbol mapping assigns bits or group of bits to the different constellation points. Within this thesis, we always use the depicted Gray mapping, where neighbour points in the constellation only differ in a single bit.

For performance comparison, it is important to be aware of the power spent for the transmission of an information bit. The energy of a data symbol is given by its square amplitude. Since the amplitude varies for QAM constellations, the mean energy of a data symbol E_s is defined as its mean squared amplitude. The energy per information bit is obtained by taking into account 1.5 as

$$E_b = \frac{E_s}{\beta} \tag{1.6}$$

The resulting symbol stream is demultiplexed into a vector $\mathbf{d}_n$ of N data symbols,

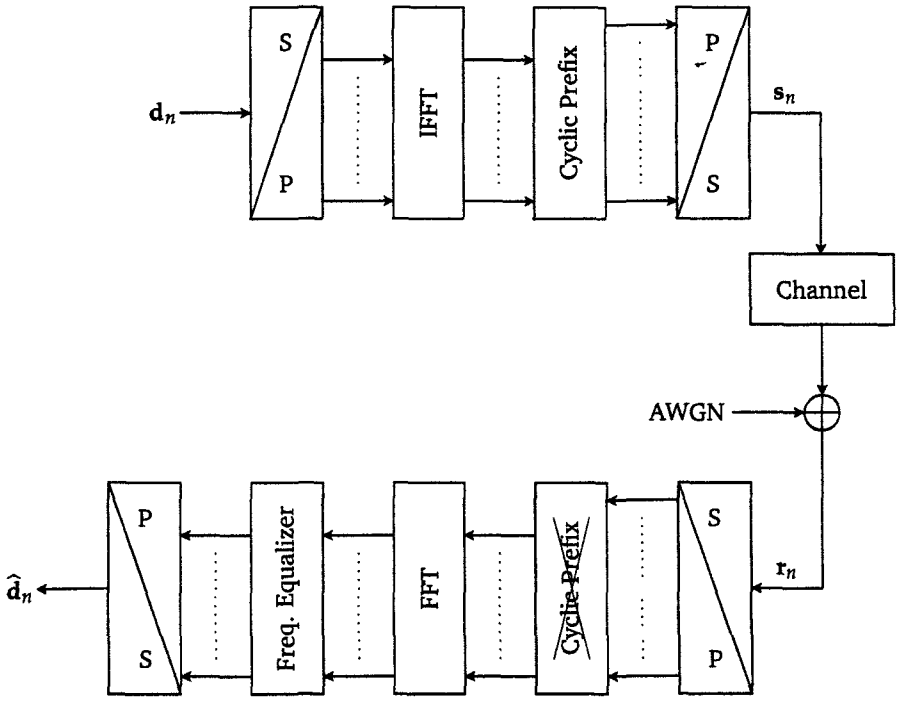

Figure 1.2 – *Block diagram of an OFDM transmission system.*

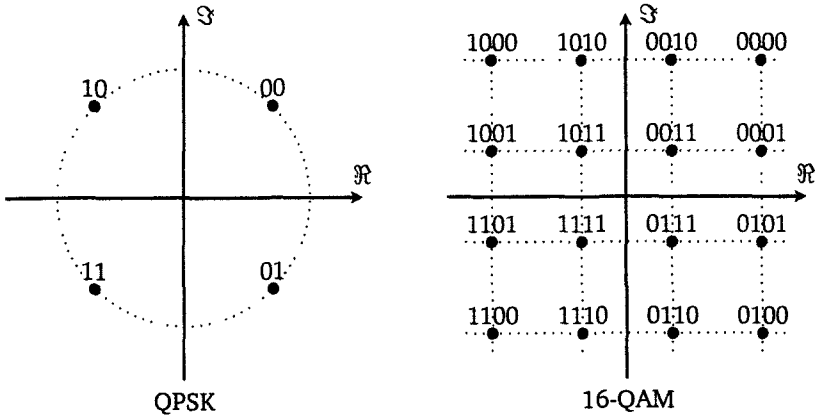

Figure 1.3 – *QPSK and 16-QAM constellations with Gray mapping.*

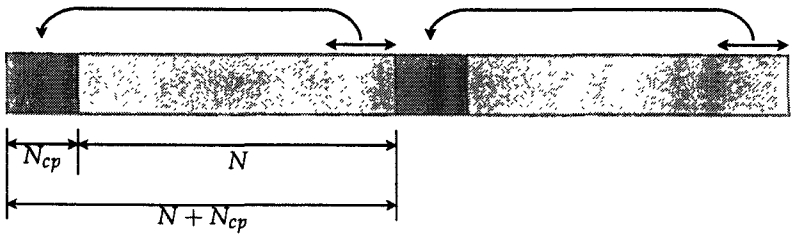

Figure 1.4 – *Insertion of a guard interval.*

$\mathbf{d}_n = \begin{bmatrix} d_{n,0} & \cdots & d_{n,N-1} \end{bmatrix}^T$. The parallel data symbol duration is N times longer than the serial symbol duration T_s, *i.e.* $T = NT_s$. The IFFT of the data symbol vector is computed and the coefficients $\{s_{n,k}\}_{k=0}^{N-1}$ constitute an OFDM symbol. $s_{n,k}$ is the k-th time domain sample of the n-th OFDM symbol and is defined as :

$$s_{n,k} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} d_{n,i} e^{j2\pi ik/N} \tag{1.7}$$

The use of the longer symbol period with OFDM already reduces the impact of the *inter-symbol interference* (ISI) arising from the delay spread of the channel. To completely cancel the ISI in the system, a *guard interval* between the consecutive OFDM symbols is inserted, as illustrated in Fig. 1.4. The transmitted signal obtained after the insertion of the cyclic prefix is a vector of $N + N_{cp}$ data and is denoted by $\mathbf{s}_n$. The duration of the guard interval is denoted by T_{cp}. The cyclic prefix is simply the last N_{cp} samples of the modulated OFDM signal as depicted in Fig. 1.4. The new duration of the overall OFDM symbol is defined as $T' = T + T_{cp}$. In order to completely cancel the ISI introduced by the delay spread of the channel, the duration of the cyclic prefix has to be longer than the maximum delay of the channel,

$$T_{cp} > \tau_{max} \tag{1.8}$$

At the OFDM receiver, the reverse action takes place. The down-converted received signal is sampled at a rate of N/T and converted into a parallel stream of $N + N_{cp}$ elements, denoted by the vector $\mathbf{r}_n$. In this thesis, the time varying frequency selective channel $h(t, \tau)$ is assumed to be static over one OFDM frame, with L_p independent propagation paths. Assuming perfect frequency, phase and timing synchronization, the n-th received OFDM symbol can be expressed as :

$$\underbrace{\begin{bmatrix} r_{n,N+N_{cp}-1} \\ r_{n,N+N_{cp}-2} \\ \vdots \\ r_{n,N_{cp}} \\ r_{n,N_{cp}-1} \\ \vdots \\ r_{n,0} \end{bmatrix}}_{\mathbf{r}_n} = \underbrace{\begin{bmatrix} h_{n,0} & h_{n,1} & \cdots & h_{n,L_p-1} & 0 & \cdots & 0 \\ 0 & h_{n,0} & h_{n,1} & \cdots & h_{n,L_p-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & \ddots & 0 \\ 0 & \cdots & 0 & h_{n,0} & h_{n,1} & \cdots & h_{n,L_p-1} \end{bmatrix}}_{\mathbf{H}_n} \underbrace{\begin{bmatrix} s_{n,N-1} \\ s_{n,N-2} \\ \vdots \\ s_{n,0} \\ s_{n,N-1} \\ \vdots \\ s_{n,N-N_{cp}} \\ s_{n-1,N-1} \\ \vdots \\ s_{n,N-L+1} \end{bmatrix}}_{\mathbf{s}_n} + \mathbf{b}_n \tag{1.9}$$

where vectors $\mathbf{r_n}$, $\mathbf{s_n}$, $\mathbf{b}_n$ and matrix $\mathbf{H_n}$ have the following respective size $(N + N_{cp}) \times 1$, $(N + N_{cp} + L_p - 1) \times 1$, $(N + N_{cp}) \times 1$ and $(N + N_{cp}) \times (N + N_{cp} + L_p - 1)$. The vector $\mathbf{b_n}$ corresponds to the additive white Gaussian noise :

$$\mathbf{b}_n = \begin{bmatrix} b_{n,N+N_{cp}-1} & \cdots & b_{n,0} \end{bmatrix}^T$$

where each element $b_{n,l}$ is a circular white Gaussian noise with power σ_b^2.

Then by assuming $N_{cp} \geqslant L_p$ and after discarding the cyclic prefix, the FFT is applied on the remaining N samples (*i.e.* $\{r_{n,k}\}_{k=N_{cp}}^{N+N_{cp}-1}$) in order to recover the desired frequency components, *i.e.* the useful symbols $d_{n,k}$ multiplied by the corresponding frequency response of the channel $\widetilde{h}_{n,k}$. For $k = 0 \ldots N-1$, we have :

$$\widetilde{r}_{n,k} = d_{n,k}\widetilde{h}_{n,k} + \widetilde{b}_{n,k} \tag{1.10}$$

where

$$\widetilde{h}_{n,k} = \sum_{p=0}^{L-1} h_{n,p} e^{-j2\pi kp/N} \tag{1.11}$$

is the channel frequency response and $\widetilde{b}_{n,k}$ is the transformed white noise which is still a circular Gaussian random variable with zero mean and variance σ_b^2.

In order to recover the desired data symbols from $\{\widetilde{r}_{n,k}\}_{k=0}^{N-1}$, an equalization step is required. Indeed, the equalization aims to compensate the distortions induced by the frequency-selectivity of the channel over each subcarrier.

From (1.10), it can be denoted that in the frequency domain only a one-tap equalizer is required. The most commonly used criterion for equalization is the *minimum mean square error* (MMSE). According to MMSE criterion, the coefficients of the equalizer are chosen to minimize the *mean square error* (MSE) between the transmitted data symbol and the equalized received symbol at the k-th subcarrier :

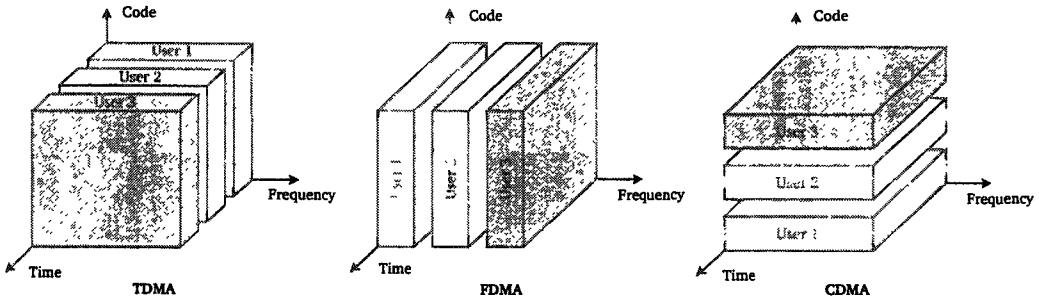

Figure 1.5 – *Multiple access techniques*

$$Q_{n,k}^{\text{MMSE}} = \arg\min_{Q_{n,k}} \left\{ \mathbb{E}\left[|d_{n,k} - Q_{n,k}\widetilde{r}_{n,k}|^2 \right] \right\} \tag{1.12}$$

Assuming that the data symbol power is unity, the solution of this minimization problem is straightforwardly given using (1.10) by :

$$Q_{n,k}^{\text{MMSE}} = \left(|\widetilde{h}_{n,k}|^2 + \sigma_b^2 \right)^{-1} \widetilde{h}_{n,k}^* \widetilde{r}_{n,k} \tag{1.13}$$

The OFDM system is now completely described. However, to enable simultaneous communications by multiple users, some type of multiple access technique must be used in combination with the OFDM modulation.

1.1.2 Multi-access techniques

Modern cellular systems are intended to enable simultaneous communications of multiple users. These users have to share the bandwidth available for a communication system. The aim of multiple access techniques is to efficiently use the bandwidth, while ensuring good transmission quality for all active users.

1.1.2.1 Some generalities

Fig. 1.5 illustrates three multi-access principles [Rappaport 02]. The three dimension representing the system resources are frequency, time and spreading code. The code dimension may equivalently be seen as the dimension of the transmitted power. Consequently, these figures show also how the transmitted power of each user is distributed in time and in frequency.

When the users are separated in time, the scheme is called *Time Division Multiple Access* (TDMA). Several users share the same frequency channel by dividing the signal into different timeslots. The users transmit in rapid succession, one after the other, each using his own timeslot. When the users are separated in frequency, the scheme is called *Frequency Division Multiple Access* (FDMA). In a FDMA scheme, the given radio frequency bandwidth is divided into adjacent frequency segments. Each segment is provided with bandwidth to enable an associated communication signal to pass through a transmission environment with an acceptable level of interference

from communication signal in adjacent frequency segments. For transmitting the same amount of information, FDMA uses a narrow bandwidth and a long duration, while TDMA uses a large bandwidth and a short duration. The disadvantage of these techniques is the lack of flexibility. Indeed, if a user is idle, a part of the available time-frequency space is unused. The third scheme is called *Code Division Access Multiple Access* (CDMA). It allows multiple users to access the channel at the same time and on the same frequency by allocating a distinct spreading code to each of them. As consequence, it is more flexible than the previous techniques and ensures that the complete time-frequency space is always used for signal transmission. Codes are chosen with small cross-correlations in order to keep the *Multiple Access Interference* (MAI) low and, thus enable the receivers to extract their desired signals even in the presence of interfering signals.

The spreading code is given by a normalized code vector of length L_c composed of elements $\{c_k^u\}_{k=0}^{L_c-1}$, also called *chips*, and defined as follows :

$$\mathbf{c}^u = \begin{bmatrix} c_0^u & \cdots & c_{L_c-1}^u \end{bmatrix}^T \tag{1.14}$$

The spreading codes vector can be chosen from a variety of code family. The choice of the spreading code depends on several factors such as orthogonality and correlation properties of the codes, the synchronism of users and the implementation complexity. In the downlink, where the signals of the different users are synchronously transmitted, orthogonal code sets are generally used. Examples of such orthogonal codes are Walsh-Hadamard codes, Golay codes, orthogonal Gold codes [Popovic 99] and carrier interferometry codes [Natarajan *et al.* 04]. Walsh-Hadamard code sets [Schnell 94] are the most common orthogonal codes for the downlink and are simple to generate. When the spreading factor is a power of two[1], the Walsh-Hadamard matrices are obtained by the following iterative rule,

$$\mathbf{C}_{L_c} = \begin{bmatrix} \mathbf{C}_{L_c/2} & \mathbf{C}_{L_c/2} \\ \mathbf{C}_{L_c/2} & -\mathbf{C}_{L_c/2} \end{bmatrix} \quad \forall L_c = 2^m \tag{1.15}$$

In the uplink, code orthogonality is less important than the previous case since the users' signals propagate through different asynchronous propagation channels, which breaks their orthogonality. In this context, codes with a good cross-correlation property, such as *Pseudo-Noise* (PN) sequences or Gold codes, are generally employed [Proakis 95].

1.1.2.2 *Multi-access in OFDM systems*

The combination of FDMA and TDMA, respectively, with OFDM modulation is referred to as OFDM-FDMA and OFDM-TDMA. In an OFDM-FDMA system, also called OFDMA (*Orthogonal Frequency-Division Multiple Access*), the data symbols for different users are transmitted on different subcarriers. One or several subcarriers are exclusively

[1]Walsh-Hadamard matrices also exist for spreading factor others than power of two. It has be shown that for $L_c < 428$, Hadamard matrices exist for all L_c divisible by 4 [van Lint *et al.* 93].

allocated to a user for transmission [Reiners *et al.* 94, Hara *et al.* 96, Kaiser 96]. An advantage of OFDMA compared to FDMA is that an optimal bandwidth utilization of the available bandwidth is guaranteed without *inter-carrier interference* (ICI). However, the performance of an uncoded OFDMA system is poor in a mobile radio channel since the frequency non-selectivity per subchannel causes significant performance degradations if a subcarrier is located in a deep fade. In the case of OFDM-TDMA, the whole bandwidth is allocated to a single user for a certain number of OFDM symbols [Rohling *et al.* 96a, Rohling *et al.* 96b, Kaiser 96]. The complexity of conventional TDMA systems is determined by the amount of ISI caused by the mobile radio channel. In OFDM-TDMA systems, the long OFDM symbol duration can drastically reduce the amount of ISI which can be completely be avoided by the use of the guard interval.

One of the most promising techniques for achieving high data rate transmission and high multiple access capability is the combination of OFDM modulation and CDMA spectrum spreading method, *i.e.* the multicarrier CDMA techniques.

1.1.3 Multicarrier CDMA techniques

Multiple access systems based on DS-CDMA with OFDM modulation were published in 1993. Two combinations were mainly proposed, namely, MC-CDMA proposed in [Yee *et al.* 93, Fazel *et al.* 93, Chouly *et al.* 93] and MC-DS-CDMA proposed in [DaSilva *et al.* 93, Kondo *et al.* 93]. All these combinations use CDMA principle in the sense that the signals of the different users occupy the total bandwidth simultaneously. The general concept of these multicarrier CDMA techniques are presented hereafter.

1.1.3.1 MC-CDMA

The principle of MC-CDMA was to perform spreading in the frequency domain [Yee *et al.* 93, Fazel *et al.* 93, Chouly *et al.* 93]. The high rate DS spread data stream is modulated using the OFDM principle in such a way that the L_c chips of a spread data symbol are transmitted in parallel on different subcarriers. Thus, the assigned data symbol is simultaneously transmitted on L_c subcarriers. If the spreading code length L_c is chosen to be equal to the number of subcarriers N, MC-CDMA requires the total bandwidth for the transmission of a single data symbol. Fig. 1.6 shows the principle of an equivalent realization of the serial concatenation of DS spreading and OFDM modulation with $L_c = N$. Each data symbol is copied into L_c substreams before multiplication with one chip of the spreading code per substream. In the case of N equals to L_c, the data symbol duration T_s and the MC-CDMA symbol duration T are identical. Fig. 1.6 reflects the fact that MC-CDMA has a spreading code grouped in the frequency domain.

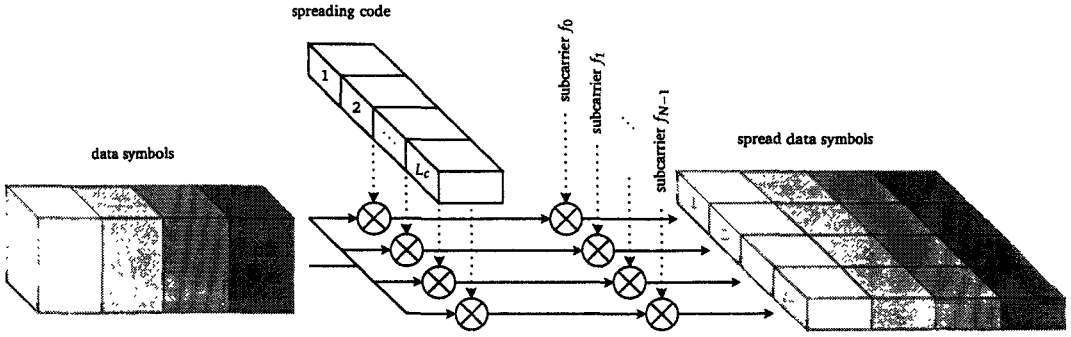

Figure 1.6 – *Principle of data spreading by MC-CDMA for a single user and in the case of* $L_c = N$.

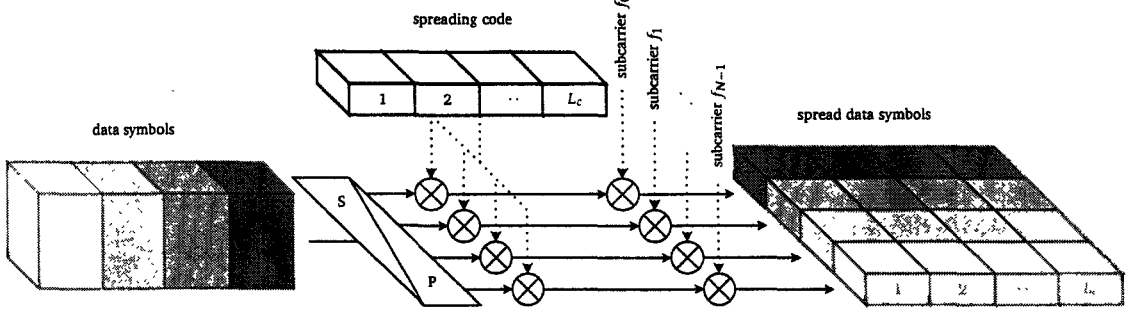

Figure 1.7 – *Principle of data spreading by MC-DS-CDMA for a single user and in the case of* $L_c = N$.

1.1.3.2 *MC-DS-CDMA system*

Similarly to the DS-CDMA, the principle of *Multicarrier Direct Sequence CDMA* (MC-DS-CDMA) was to perform spreading along the time domain [DaSilva *et al.* 93, Kondo *et al.* 93]. It can be seen as parallel DS-CDMA streams of long chip duration and, consequently, of moderate bandwidth.

With MC-DS-CDMA, the serial data stream is first converted into parallel low-rate substreams before applying DS spreading on each substream as illustrated in Fig. 1.7. When setting the number of subcarriers N to one, MC-DS-CDMA becomes identical to DS-CDMA. Unlike with MC-CDMA, if the spreading code length L_c is equal to the number of subcarriers N, as illustrated in Fig. 1.7, or less than N, a single data symbol is not spread in bandwidth with MC-DS-CDMA. Instead, is is extended in the time. The design of MC-DS-CDMA systems with a high number of subcarriers, where each subcarrier can be considered as frequency non-selective, is advantageous in the sense that the system exploits the time diversity.

1.1.4 Practical multicarrier symbol

Practical multicarrier systems are generally not fully loaded in order to avoid aliasing. In this case, some of the subcarriers at the edges of the multicarrier block are not modulated. These N_g subcarriers are referred to as virtual subcarriers (VSCs). This

number is dictated by system design requirements and is, in general, about 10 % of subcarriers. We denote by Ω_g the set of virtual subcarriers. Moreover, to stay in a general multicarrier transmission, let us introduce a set of P pilot subcarriers located on Ω_p, with $\{\Omega_p \cap \Omega_g\} = \emptyset$ and $\Omega_p \cup \Omega_g = \Omega$. These considerations can be summed up as :

$$d_{n,i} = \begin{cases} \text{Information Data} & i \notin \Omega \\ \text{Pilot} & i \in \Omega_p \\ 0 & i \in \Omega_g \end{cases} \tag{1.16}$$

Let us illustrate the presence of both pilot and null subcarriers in a multicarrier symbol by a specific example of a real OFDM system. Indeed, the OFDM transmission used in the European standard for WLAN in the 5 GHz band, Hiperlan/2, has been specified with $N = 64$ subcarriers of which 12 are set to zero and a cyclic prefix length $N_{cp} = 16$. Hence, 80 samples are transmitted for each OFDM symbol. Among the 52 non null subcarriers, $P = 4$ are pilots carrying known data symbols. The system is designed to provide various data rates and hence both several constellations (from BPSK up to 64-QAM) and several coding rates are specified in the standard. Main parameters of the physical layer of Hiperlan/2 are summarized in Table 1.1.

Sampling rate, $f_s = 1/T_s$	20 MHz
Useful OFDM symbol part duration, T	$64T_s = 3.2\mu$s
Cyclic prefix duration, T_{cp}	$16T_s = 0.8\mu$s
Subcarrier spacing, $\Delta f_{\text{subcarrier}}$	0.3125 MHz
Number of pilot subcarriers, P	4
Number of null subcarriers, N_g	12
Number of data subcarriers	$N - P - N_g = 48$

Table 1.1 – *Main parameters of the OFDM system in Hiperlan/2.*

The information data in Eq. (1.16) is a function of user data symbols depending on the system configuration. In OFDM systems, $d_{n,i}$, for $i \notin \Omega$ is simply the data symbols.

However, in MC-CDMA and in MC-DS-CDMA this information data depends on the spreading code, the user data symbols and the set of pilots and null-subcarriers as illustrated in Figs. 1.8 and 1.9 where $\left\{d^u_{n,i}\right\}_{i=0}^{N-1}$ correspond to the information data of the u-th user. Let us note that the spreading code length must be chosen such as $L_c \leq N - P - N_g$.

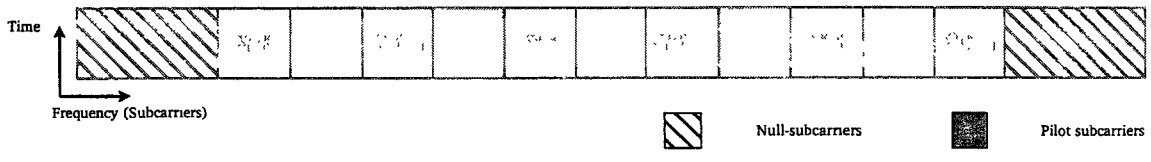

Figure 1.8 – *Example of a MC-CDMA block symbol corresponding to the u-th user in the presence of both null and pilot subcarriers.*

If a downlink multicarrier system without pilot and null subcarriers is considered,

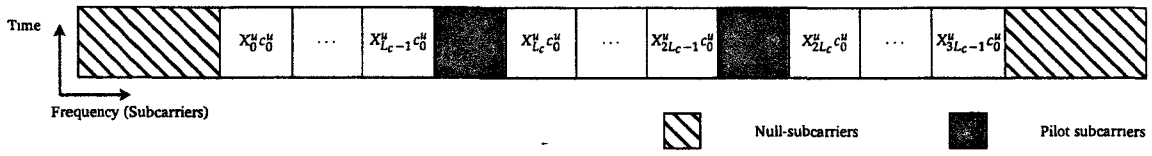

Figure 1.9 – *Example of a MC-DS-CDMA block symbol corresponding to the u-th user in the presence of both null and pilot subcarriers.*

$\{d_{n,i}\}_{i=0}^{N-1}$ is simply given by :

$$d_{n,i} = \begin{cases} X_{nN+i} & \text{for OFDM system} \\ \sum_{u=1}^{N_u} X_n^u c_i^u & \text{for MC-CDMA system} \\ \sum_{u=1}^{N_u} X_{\lfloor \frac{n}{N} \rfloor N+i}^u c_{n \bmod L_c}^u & \text{for MC-DS-CDMA systems} \end{cases} \quad (1.17)$$

where X_n^u correspond to the data symbol associated to the u-th user.

As a consequence, the received signal defined in Eq. (1.9), with $s_{n,k}$ given by (1.7), holds whatever both the multicarrier system and the number of pilot and null subcarriers are. The only difference is in the definition of $\{d_{n,i}\}_{i=0}^{N-1}$. In this section, imperfections induced by the system components have not been introduced which is consequently not a realistic case. Since multicarrier systems are very sensitive to phase distortions, the next section is devoted to their description and to their impact on system performances.

1.2 MULTICARRIER RECEPTION IN THE PRESENCE OF PHASE DISTORTIONS

1.2.1 Oscillator instabilities

One of the most critical components of any wireless system, especially in multicarrier systems, which may ultimately limit the performance of the communication is the receiver and/or transmitter front-end. Indeed, analog front-end impairments, such as noise, distortion and mismatch, may affect performance by degrading the integrity of the desired signal.

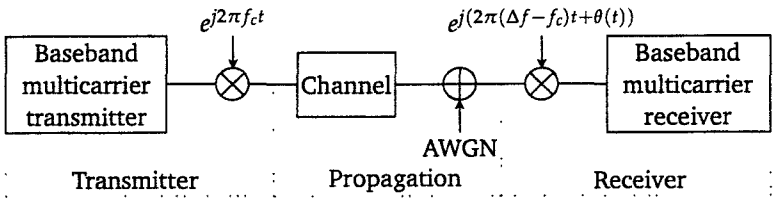

Figure 1.10 – *Multicarrier system with oscillator instabilities at the receiver side.*

As illustrated in Fig. 1.10, we focus on carrier frequency offset and phase noise which corrupt the received multicarrier signal.

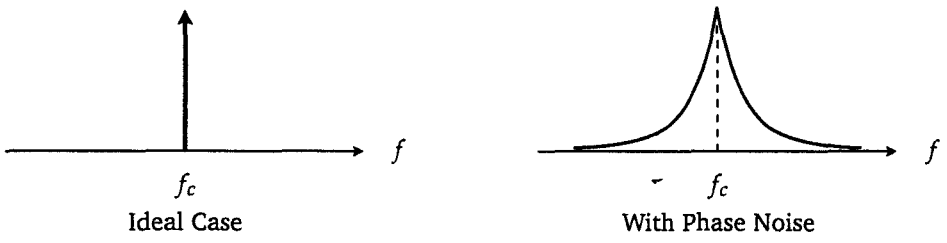

Figure 1.11 – *Representation of frequency spectra of oscillator signal.*

1.2.1.1 *Carrier frequency offset*

The local oscillators in the transmitter and receiver are based on accurate frequency references. However, due to their physical separation, different frequency references are used in the transmitter and receiver which lead to a frequency offset between the local oscillators.

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

$$p(t) = \exp(j2\pi\Delta f t) \tag{1.18}$$

where Δf corresponds to the carrier frequency offset. At the sampling rate of the receiver N/T, the discrete form of (1.18) is given by :

$$p(\frac{kT}{N}) = \exp(j2\pi k\epsilon/N) \tag{1.19}$$

where $\epsilon = \Delta f T$ is the normalized CFO with respect to the subcarrier spacing $1/T$.

Consequently, CFO results in a spinning constellation. If the CFO is sufficiently small relative to the signal bandwidth, a differentially-encoded signaling scheme such as differential phase shift-keying (DPSK) can be used [Sheng *et al.* 98]. In this case, the data is encoded in the transmission between symbols so that only the phase difference between successive symbols is needed for signal modulation. Unfortunately, differentially-encoded modulation schemes require a larger signal-to-noise ratio (SNR) for the same bit error rate (BER) performance. If coherent demodulation is preferred or if the frequency offset is large relative to the signal bandwidth, then frequency offset compensation is required.

1.2.1.2 *Phase noise*

Ideally, the local oscillator used in a communication system delivers a pure sine wave carrier. This would be represented in the frequency domain by a single line. However, all real sources have unwanted amplitude or phase modulated noise components. These phase modulated components are known as phase noise. As illustrated in Fig. 1.11, for an ideal oscillator operating at f_c, the spectrum assumes the shape of an impulse whereas for an actual oscillator, the spectrum exhibits "skirts" around the carrier frequency.

In a baseband complex equivalent form, the signal delivered by the noisy oscillator

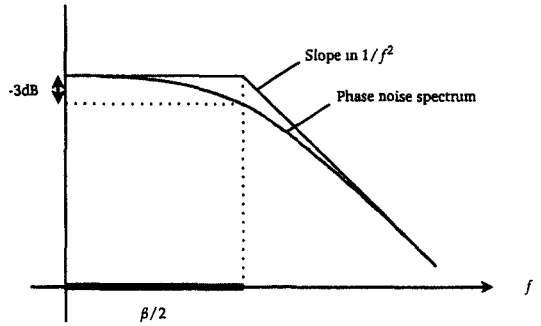

Figure 1.12 – *Representation of the power spectral density of $p(t)$ in the case of Wiener PHN.*

can be modelled as :

$$p(t) = \exp(j\theta(t)) \tag{1.20}$$

where $\theta(t)$ represents the phase noise process. Two different models of PHN are available in the literature [Piazzo *et al.* 02]. The first one models a free-running oscillator and assumes the PHN process to be a Wiener process that is nonstationary and whose power grows with time. The second one models an oscillator controlled by a phase-locked loop (PLL) and approximates the PHN process as a zero-mean colored Gaussian process that is wide-sense stationary (WSS) and has finite power. In this thesis, we focus on Wiener PHN and therefore assume the presence of a free-running oscillator.

The PHN modelled as a Wiener process [Pollet *et al.* 95, Tomba 98] is defined as :

$$\theta(t) = \int_0^t v(\tau)d\tau \tag{1.21}$$

where $v(\tau)$ is zero-mean white Gaussian noise process. The power spectral density of $p(t)$ has a Lorentzian shape controlled by the parameter β representing the two-sided 3 dB bandwidth as illustrated in Fig. 1.12. This model produces a $1/f^2$ type noise power behavior that agrees with experimental measurements carried out on real RF oscillators. The phase noise rate is characterized by the bandwidth β normalized with respect to the subcarrier spacing $1/T$, namely by the parameter βT. The discrete-time samples of $\theta(t)$ form a random-walk process :

$$\theta_{n,k} = \theta_{n,k-1} + v_{n,k} \tag{1.22}$$

where n and k denote respectively the n-th multicarrier symbol and the k-th time domain sample; $v_{n,k}$ is an independent and identically distributed (*i.i.d.*) zero mean Gaussian variable with variance $2\pi\beta T/N$. Moreover, if perfect phase synchronization at the beginning of each mulicarrier symbol is assumed as in [Tomba 98, Lin *et al.* 07], the following relation holds :

$$\theta_{n,-1} = 0 \tag{1.23}$$

1.2.2 On the effect of phase distortions in multicarrier systems

In this section, we study the phase impairment impact on multicarrier system performance. In this section, the multicarrier system is assumed fully loaded with information data (*i.e.* $\Omega = \emptyset$). After discarding the cyclic prefix at the receiver side and by assuming $N_{cp} \geqslant L_p$, the remaining N samples of the received signal corrupted by the phase distortions, including both PHN and CFO, are given for $k = 0, ..., N-1$:

$$r_{n,N_{cp}+k} = e^{j\theta_{n,k}} e^{j\frac{2\pi k\epsilon}{N}} \sum_{l=0}^{L_p-1} h_{n,l} s_{n,k-l} + b_{n,N_{cp}+k} \tag{1.24}$$

Then, by performing discrete Fourier transform (DFT) on these N samples, the demodulated signal at subcarrier k ($s = 0, 1, ..., N-1$) of the n-th OFDM symbol is given as :

$$\begin{aligned} \widetilde{r}_{n,k} &= d_{n,k}\widetilde{h}_{n,k} \underbrace{I_n(0)}_{\text{CPE}} \\ &\quad + \underbrace{\sum_{\substack{i=0 \\ i\neq k}}^{N-1} d_{n,i}\widetilde{h}_{n,i} I_n(k-i)}_{\text{ICI}} + \widetilde{b}_{n,k} \end{aligned} \tag{1.25}$$

where

$$I_n(m) = \frac{1}{N} \sum_{k=0}^{N-1} e^{j\frac{2\pi k\epsilon}{N}} e^{j\theta_{n,k}} e^{-j\frac{2\pi km}{N}} \tag{1.26}$$

and,

$$\widetilde{h}_{n,i} = \sum_{p=0}^{L-1} h_{n,p} e^{-j2\pi ip/N} \tag{1.27}$$

with $\widetilde{b}_{n,k}$ the transformed white noise which is still a circular gaussian random variable with zero mean and variance σ_b^2. In Eq. (1.25), the multiplicative term $I_n(0)$ is common to all subcarriers of one OFDM symbol and correspond to the common phase error (CPE). The terms $I_n(m)$ is simply the DFT of the sequence $\left\{ e^{j\frac{2\pi k\epsilon}{N}} e^{j\theta_{n,k}} \right\}_{k=0}^{N-1}$. Phase distortions introduce also inter-carrier interference (ICI) defined in (1.25). To understand ICI, we need to characterize the coefficients $I_n(m)$ where its cross-correlation is calculated as :

$$\begin{aligned} \mathbb{E}\left[I_n(m) I_n^*(p)\right] &= \frac{1}{N^2} \mathbb{E}\left[\sum_{k=0}^{N-1} \sum_{l=0}^{N-1} e^{j(\theta_{n,k}-\theta_{n,l})} e^{j\frac{2\pi(k-l)\epsilon}{N}} e^{-j\frac{2\pi}{N}(mk-pl)} \right] \\ &= \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \mathbb{E}\left[e^{j\Delta_{n,kl}}\right] \mathbb{E}\left[e^{j\frac{2\pi(k-l)\epsilon}{N}}\right] e^{-j\frac{2\pi}{N}(mk-pl)} \end{aligned} \tag{1.28}$$

where $\Delta_{n,kl}$ denotes the cumulative phase noise increment between the l-th and k-th samples of the received signal. From Eq. (1.22), the increments of the phase noise between two consecutive samples are *i.i.d.* zero mean Gaussian random variables of

variance $\sigma_v^2 = 2\pi\beta T/N$. Consequently, $\Delta_{n,kl}$, as a sum of *i.i.d.* zero mean gaussian random variables, is also a Gaussian random variable defined as follows :

$$\Delta_{n,kl} \sim \mathcal{N}(0, |k-l|\sigma_v^2) \tag{1.29}$$

In order to evaluate Eq. (1.28), the expectations $\mathbb{E}\left[e^{j\Delta_{n,kl}}\right]$ and $\mathbb{E}\left[e^{j\frac{2\pi(k-l)\epsilon}{N}}\right]$ has to be calculated for each $k,l = 0,1,...,N-1$. However, these expressions can be easily obtained by using the definition of characteristic function. Indeed, the characteristic function of the random variable X is defined as :

$$\psi(t) = \mathbb{E}\left[e^{jtX}\right] \tag{1.30}$$

Since $\Delta_{n,kl}$ is a Gaussian random variable, it follows that :

$$\mathbb{E}\left[e^{j\Delta_{n,kl}}\right] = \psi_1(1) = e^{-\frac{|k-l|\sigma_v^2}{2}} \tag{1.31}$$

Now, if we assume that ϵ follows an uniform distribution on the interval $[-\Delta_{\text{CFO}}; \Delta_{\text{CFO}}]$, the expectation of the CFO term can be written as :

$$\mathbb{E}\left[e^{j\frac{2\pi(k-l)\epsilon}{N}}\right] = \psi_2(k-l) = \text{sinc}\left(\frac{2\pi(k-l)\Delta_{\text{CFO}}}{N}\right) \tag{1.32}$$

where the sinc function is defined by $\text{sinc}(x) = \sin(x)/x$. Finally, the cross-correlation of coefficients $I_n(m)$ is then given by :

$$\mathbb{E}\left[I_n(m)I_n^*(p)\right] = \frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\text{sinc}\left(\frac{2\pi(k-l)\Delta_{\text{CFO}}}{N}\right)e^{-\frac{|k-l|\sigma_v^2}{2}}e^{-j\frac{2\pi}{N}(mk-pl)} \tag{1.33}$$

Considering the ICI term in (1.25), the correlation between the $I_n(m)$ is destroyed due to the randomization by data and channel coefficients. The total ICI power induced by both PHN and CFO is given by :

$$\sigma_{ICI}^2 = \mathbb{E}\left[\left|\sum_{\substack{v=0\\v\neq k}}^{N-1} d_{n,k-v}\tilde{h}_{n,k-v}I_n(v)\right|^2\right] \tag{1.34}$$

Since in general multicarrier systems, the information data $d_{n,k}$ is a function of user symbols and spreading codes as described in Section 1.1.4 and by assuming $\mathbb{E}[X_k^u X_l^{*v}] = \delta_{k,l}\delta_{u,v}$, $\mathbb{E}[c_k^u c_l^v] = \delta_{k,l}\delta_{u,v}$ and $\mathbb{E}\left[|d_{n,k}|^2\right] = 1$, the cross-correlation of the information data involved in this equation is given by :

$$\mathbb{E}[d_{n,k}d_{n',k'}^*] = \delta_{n,n'}\delta_{k,k'} \tag{1.35}$$

Using (1.35), the total ICI power is therefore obtained by :

$$\sigma^2_{ICI} = \sum_{\substack{\nu=0 \\ \nu \neq k}}^{N-1} \mathbb{E}\left[|d_{n,k-\nu}|^2\right] \mathbb{E}\left[\left|\widetilde{h}_{n,k-\nu}\right|^2\right] \mathbb{E}\left[|I_n(\nu)|^2\right]$$
$$= \sum_{\substack{\nu=0 \\ \nu \neq k}}^{N-1} \mathbb{E}\left[|\vec{I}_n(\nu)|^2\right] \tag{1.36}$$

where it is assumed that $\mathbb{E}\left[\left|\widetilde{h}_{n,k-\nu}\right|^2\right] = 1$.

Figures 1.13 and 1.14 show the evolution of the ICI power versus respectively the PHN rate βT and the CFO interval. Since these characterization parameters of both PHN and CFO are normalized with respect to the subcarrier spacing, the ICI power does not depend on the number of subcarriers. It can be logically remarked that the ICI power increases with the PHN rate, βT, and the upper bound of the CFO interval, Δ_{CFO}.

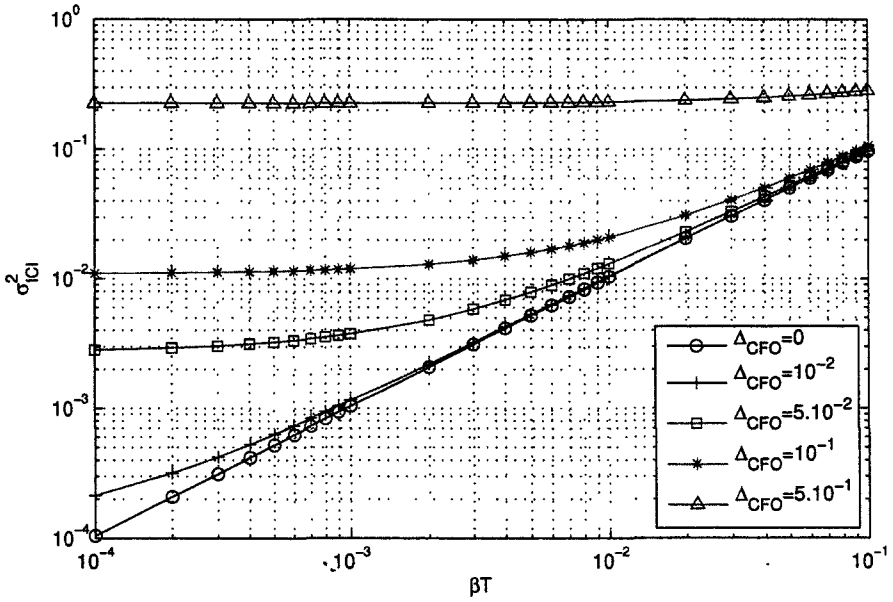

Figure 1.13 – *Evolution of the ICI power versus the PHN rate βT for different CFO intervals.*

In order to evaluate more precisely the performance degradations due to phase distortions, let us introduced the signal-to-interference-plus-noise ratio (SINR) which is given from (1.25) by :

$$\text{SINR} = \frac{\mathbb{E}\left[|I_n(0)|^2\right]}{\sigma^2_{ICI} + \sigma^2_b} \tag{1.37}$$

As shown by this equation, the SINR takes into account the power of both the CPE and ICI. The SINR is then a very important indicator of system performance. The evolution of the SINR versus the SNR is depicted in Fig. 1.15 for different PHN rates

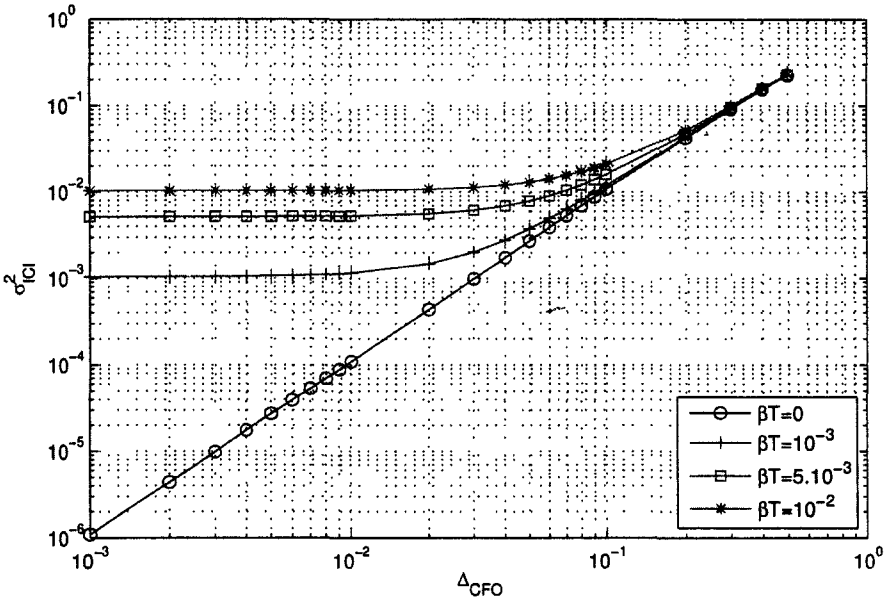

Figure 1.14 – *Evolution of the ICI power versus the CFO interval Δ_{CFO} for different PHN rates βT.*

and CFO intervals. Let us recall that the SNR is simply equal to $1/\sigma_b^2$. As shown in these figures, the larger βT or Δ_{CFO} is, the worse the SINR. Moreover, it can be seen that when phase distortions corrupts the received signal and the SNR increases, the SINR tends toward a value which depends on the phase distortions characteristics. Indeed, this value is given directly from Eq. (1.37) :

$$\text{SINR}|_{\sigma_b^2 \to 0} = \frac{\mathbb{E}\left[|I_n(0)|^2\right]}{\sigma^2_{ICI}} \tag{1.38}$$

The value of $\text{SINR}|_{\sigma_b^2 \to 0}$ depicted versus the PHN rate in Fig. 1.16 clearly shows the high performance degradations induced by the phase distortions.

Now by using unnormalized phase distortions characteristics, *i.e.* β and Δ_{CFO}/T both in Hz, we study the impact of the number of subcarriers on SINR in OFDM systems with an equivalent symbol rate of $D = N/T = 10^6$ symbols/s and thus the same bandwidth (Eq. (1.4)). As shown in Fig. 1.17, the larger number of subcarrier N is, the worse SINR is. Indeed, a larger number of subcarriers leads to a severe performance degradation due to the shorter subcarrier spacing distance which is consequently more sensitive to phase distortions. Furthermore, we remark that the SINR degrades as a logarithmically linear function of β for $\beta \geq 100$ Hz and $\beta \geq 1$ kHz respectively with Δ_{CFO}/T equal to 0 Hz and 1 kHz.

Finally, Fig. 1.18 depicts the OFDM BER performance in the presence of phase distortions. The system parameters are $N = 64$, $N_{cp} = 8$. A 16-QAM modulation and a Rayleigh channel with $L_p = 4$ paths and a uniform power delay profile are used for these numerical simulations. As expected by the previous results, a severe performance degradation is observed. Moreover, it can be denoted that the BER curves

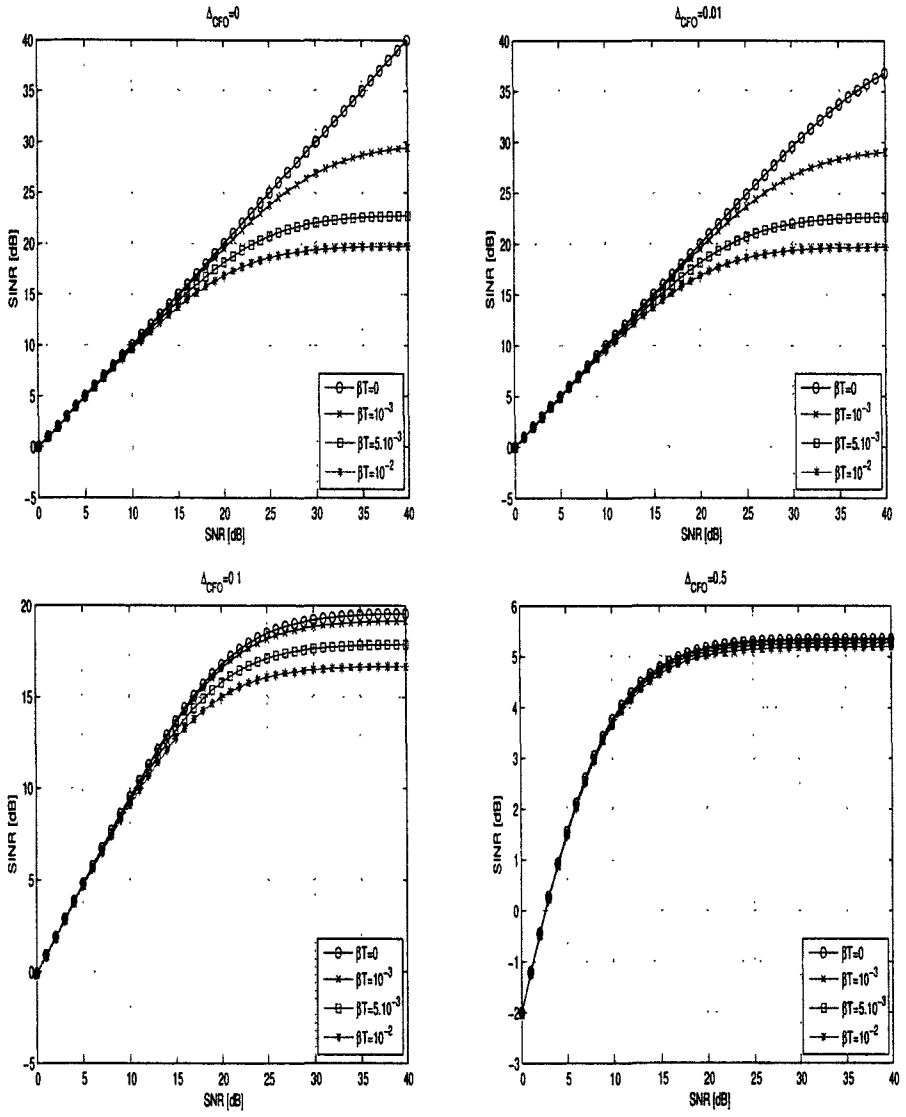

Figure 1.15 – *Evolution of the SINR versus the SNR for different PHN rates and CFO intervals.*

tends toward a threshold which depends on the ICI power and therefore on the phase distortions characteristics. BER performances of other multicarrier systems are also significantly degraded in the presence of phase distortions as shown in [Pollet *et al.* 95, Costa *et al.* 02, Tomba 98, Tomba *et al.* 99, Steedam *et al.* 01, Garnier *et al.* 02].

From these different results, it can be concluded that it is essential to take into account these distortions in multicarrier receiver in order to improve system performances. In the next section, existing strategies for multicarrier receivers in the presence of phase impairments are reviewed.

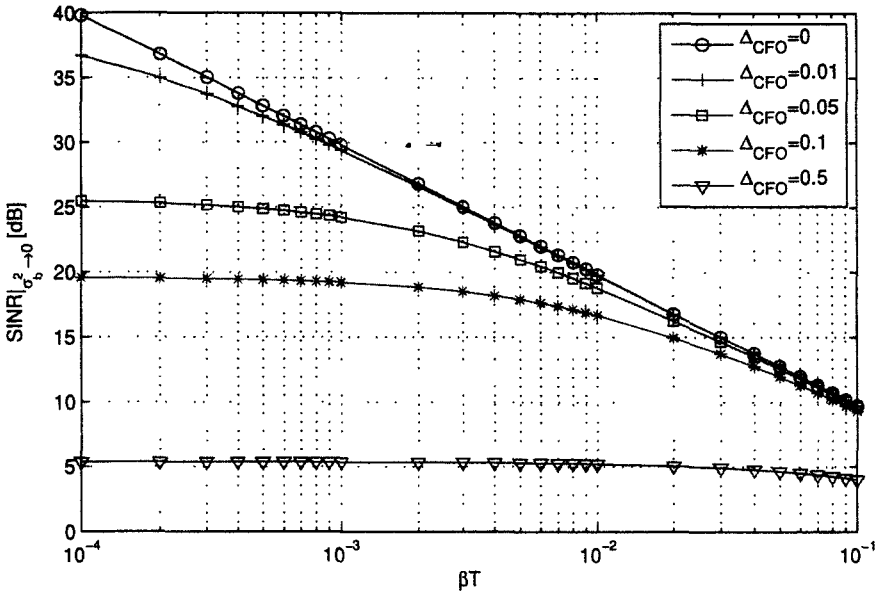

Figure 1.16 – *Evolution of the* $SINR|_{\sigma_b^2 \to 0}$ *versus the PHN rate* βT *for different CFO intervals* Δ_{CFO}.

1.3 EXISTING MULTICARRIER RECEIVERS IN THE PRESENCE OF PHASE DISTORTIONS

In literature, many approaches have been proposed to data detection in OFDM systems by compensating PHN, or both PHN and residual CFO[2]. Neverthless, the channel impulse response (CIR) is always assumed perfectly known prior to phase distortions mitigation.

In fact, as in many standard based on multicarrier modulation such as Hiperlan/2 or IEEE 802.11a, training OFDM symbols which are known to the receiver are used for channel estimation before the data transmission as illustrated by Fig. 1.19. Indeed, since the CIR changes slowly[3] with respect to the OFDM symbol rate, the channel is thus only estimated at the beginning of a frame. Then, its estimates is used for data detection in the payload section. However in the presence of phase impairments at the receiver side, a channel estimator which takes into account the presence of both PHN and CFO in the received signal are consequently required to achieve reliable transmission.

[2]When no PHN is present, a detailed overview of the CFO estimation problem in OFDM systems is given in [Ghogho *et al.* 04]. However, this is not the subject of this thesis since a random PHN is here considered.

[3]For example, in Hiperlan/2 standard specifications, the channel variations are supposed to correspond to terminal speeds $v \leq 3$m/s.

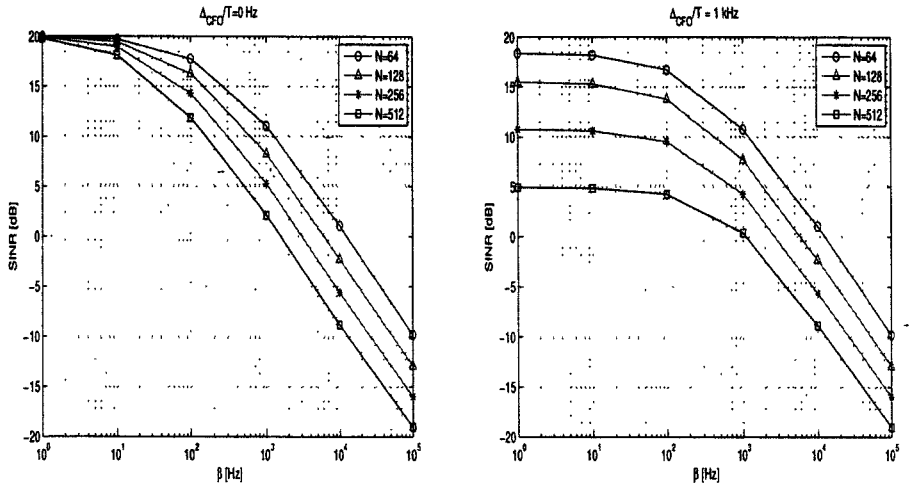

Figure 1.17 – *Effect of PHN linewidth β on SINR performance for different numbers of subcarriers, with SNR = 20 dB and $D = 10^6$ symbols/s.*

1.3.1 Channel estimation in the presence of phase distortions

Only two papers deal with channel estimation in the presence of both CFO and PHN. In [Wu *et al.* 03b], PHN was considered in the formulation of the channel estimation problem but was not directly used in the solution and thus the method is not statistically optimal. Recently, in [Lin *et al.* 06] an approximate maximum *a posteriori* estimator of CIR, PHN and CFO has been proposed. However in this proposed estimator, the PHN term is linearized in order to obtain an analytical solution of the estimation problem, *i.e.* $e^{j\theta_{n,k}} \approx 1 + j\theta_{n,k}$. Moreover both the AWGN and PHN powers, respectively denoted by σ_b^2 and σ_v^2, are assumed perfectly known by the receiver which is not a realistic assumption.

1.3.2 Data detection in the presence of phase distortions

In literature, various methods have been proposed for data detection in the presence of phase distortions. Theses techniques can be divided into frequency domain [Robertson *et al.* 95, Yee *et al.* 05, Petrovic *et al.* 04a, Wu *et al.* 04, Petrovic *et al.* 04b, Wu *et al.* 03a, Nikitopoulos *et al.* 05] and time domain approaches [Casas *et al.* 02, Lin *et al.* 05, Lin *et al.* 07]. For both of them, two different strategies based either on pilot or non-pilot aided algorithms has been employed. The latters are a challenging task since they have the advantage of being bandwidth more efficient as it does not require the transmission of pilots symbols.

1.3.2.1 Frequency domain approaches

The frequency domain approach consists in correcting CPE and ICI in the received signal obtained after the FFT, Eq. (1.25). The CPE correction method is the simplest algorithm and one of the earliest work to eliminate phase noise effects [Robertson *et al.* 95]. The CPE estimation is performed using the known data on pilot

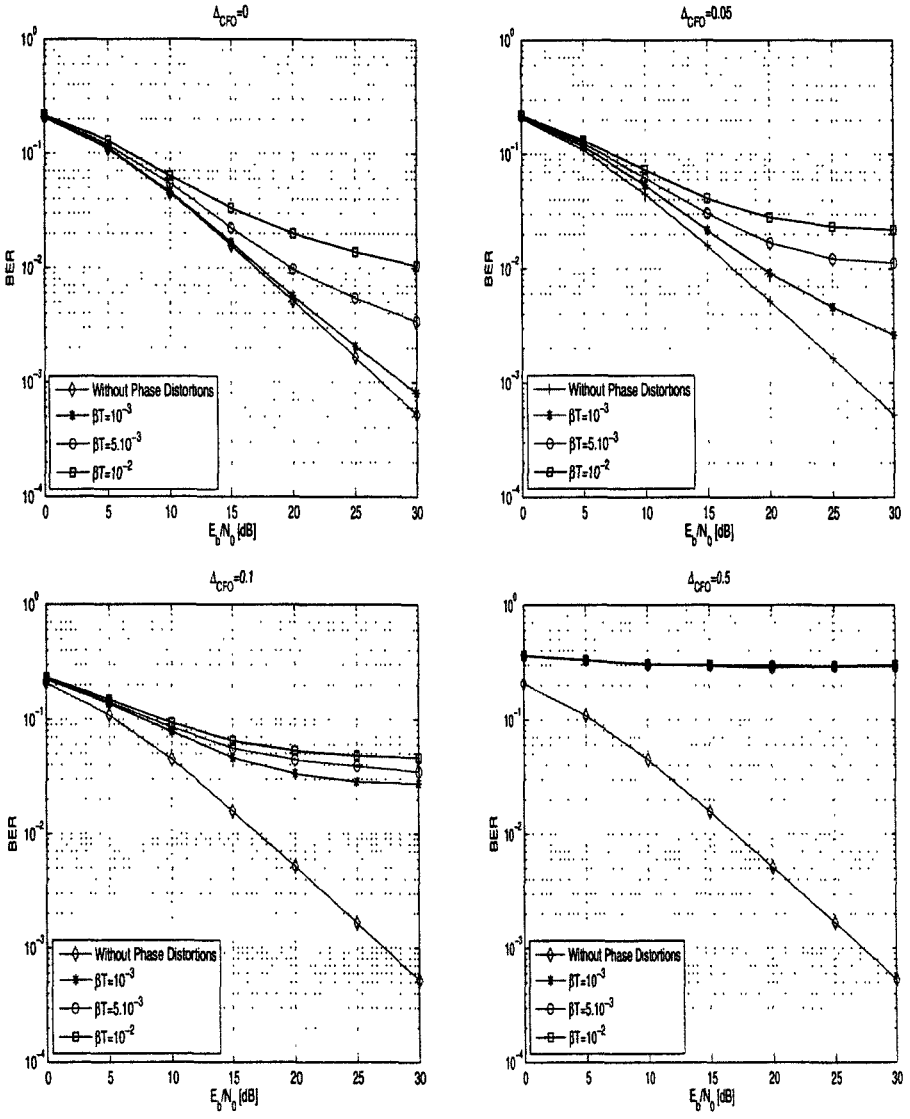

Figure 1.18 – *BER performance of OFDM systems in the presence of phase distortions.*

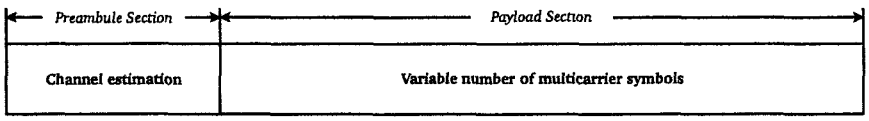

Figure 1.19 – *Multicarrier packet structure.*

subcarriers. However, the effectiveness of this algorithm is limited as it neglects ICI, which is more significant when the phase noise rate increases [Wu *et al.* 04]. Thus after modelling the ICI caused by the random phase noise as an extra additive noise, algorithms proposed in [Yee *et al.* 05, Petrovic *et al.* 04a, Wu *et al.* 04] improve performances of CPE estimation and correction. However, as illustrated by Fig. 1.21 where BER performance of a perfect CPE correction scheme is depicted, suitable system per-

formances cannot be reached with only a CPE correction, especially for significant phase distortions where ICI becomes predominant.

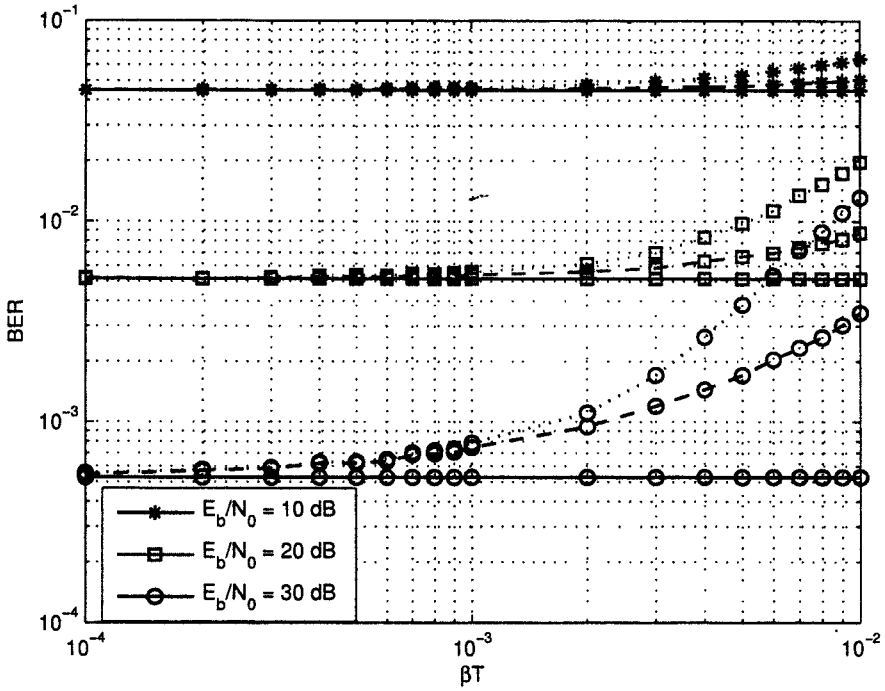

Figure 1.20 – *BER performance of OFDM systems without phase distortions (solid lines) and in the presence of only PHN (i.e. $\Delta_{CFO} = 0$) without any correction (dotted lines) and with a perfect CPE correction (dashed lines).*

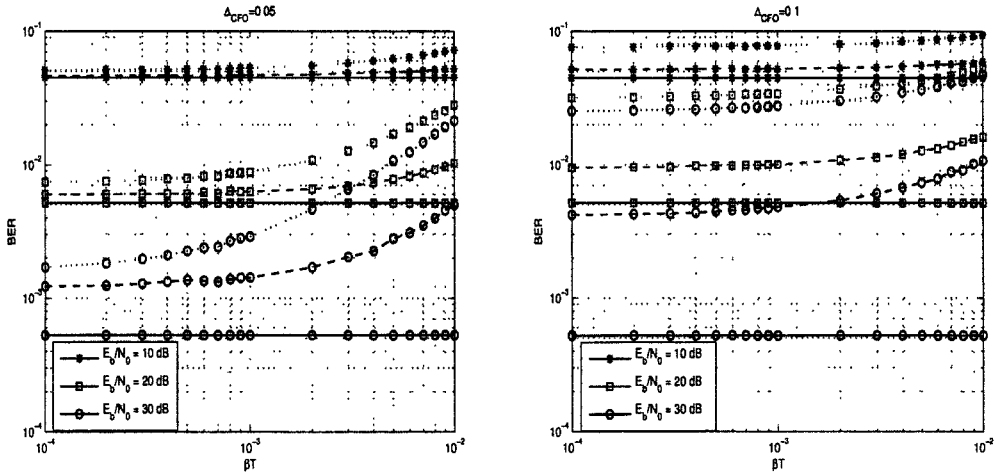

Figure 1.21 – *BER performance of OFDM systems without phase distortions (solid lines) and in the presence of both PHN and CFO without any correction (dotted lines) and with a perfect CPE correction (dashed lines).*

As a consequence, in [Petrovic *et al.* 04b, Wu *et al.* 03a], authors proposed new algorithms performing both CPE and ICI correction. In a first step, CPE estimation is

carried out with either a Kalman filter or a minimum mean square error pilot aided algorithm. Secondly, a decision feedback algorithm, using the demodulated symbols obtained after the CPE correction, is performed to obtain ICI estimate. However, noise-induced symbol decision errors may propagate through the feedback loop, leading to unreliable transmissions with error bursts. Moreover, the CPE and ICI correction algorithms are proposed for OFDM systems in the presence of only PHN (*i.e* without CFO).

Finally in [Nikitopoulos *et al.* 05], unlike the previous algorithms processing in the frequency domain, the proposed scheme does not require the presence of pilot subcarriers. A decision-directed scheme is used at the initialization step in order to make a tentative decision on a prespecified set of subcarriers without any phase distortion correction. Authors consequently assume that the phase distortion variations are slow enough to lead to reliable tentative decision used to CPE estimation. This CPE estimation is thus employed to correct the corrupted received signal. In [Nikitopoulos *et al.* 05], performances of the proposed method is studied in the presence of CFO which is however considered as residual (*i.e* $\Delta_{\text{CFO}} < 0.01$).

Using (1.25), we can denote that, in frequency domain, each observation over one multicarrier symbol depends on all phase noise states $\{\theta_{n,k}\}_{k=0}^{N-1}$ and all data information $\{d_{n,k}\}_{k=0}^{N-1}$. One disadvantage of this approach is that the *a priori* information about phase noise evolution in (1.22) cannot be taken into account properly for phase tracking. The time domain approach thus appears as an alternative for phase noise mitigation in OFDM systems.

1.3.2.2 Time domain approaches

In [Casas *et al.* 02], authors present an algorithm for estimating and cancelling the phase noise interference with the use of pilot subcarriers. From a linearized parametric model for phase noise (a sum of discrete time domain components), they obtain a least squares (LS) estimate of the transmitted symbol and indirectly eliminate the effect of the modeled components of the phase noise in the time domain.

Recently, an other approach is proposed in [Lin *et al.* 05, Lin *et al.* 07]. An approximate probabilistic inference technique associated with the minimization of variational free energy[4], is proposed to jointly estimate phase noise and transmitted symbols without any pilot subcarriers. However, in order to derive its algorithm, the phase distortions term is linearized by the assumption of small phase noise rate (i.e. $e^{j\theta(t)} \approx 1 + j\theta(t)$). Moreover, like in [Nikitopoulos *et al.* 05], a tentative symbol decision without any phase correction is made at initialization step and thus leads to a severe performance degradation when phase distortions are significant.

[4]Variational inference principle is briefly reviewed in Section 2.2.3

CONCLUSION

OFDM systems have been standardized in many recent applications due to its ability to combat multipath effects and make better use of the system available bandwidth. In this chapter, the traditional OFDM system with cyclic prefix have been firstly introduced. Then, multiple access schemes based on OFDM modulation have been discussed. More specifically, the benefits and success of OFDM modulation on the one hand, and the flexibility offered by spread spectrum techniques on the other, have, since 1993, motivated many researchers to investigate the combination of both techniques, known as multicarrier CDMA techniques. After a brief description of these combinations which benefit from the main advantages of both systems, some characteristics of real systems are also discussed, especially the presence of null and/or pilot subcarriers in the multicarrier symbol.

Unfortunately, multicarrier systems are very sensitive to phase noise (PHN) and carrier frequency offset (CFO) caused by oscillator instabilities. Indeed, after a brief statistical description of phase distortions, their effects have been analyzed with a closed form expression for the SINR, with which, system behavior can clearly be judged for any phase distortions levels.

In order to have a reliable multicarrier system, a phase distortion compensation scheme is consequently required. Existing schemes dealing with phase impairments in OFDM systems are briefly reviewed. These algorithms are based on two consecutive steps which consists in estimating successively the channel impulse response using a training multicarrier symbol and then the transmitted data symbols. However as seen in Section 1.3, these approaches suffer from several drawbacks. On the one hand, for the channel estimation problem, both the AWGN and PHN powers are assumed known to the receiver in existing schemes, which is not a realistic assumption. On the other hand, for data detection, efficient pilot or non-pilot aided algorithms are based on decision-directed scheme and thus lead to poor performance for high phase distortions levels. Moreover, these data estimator are only derived for OFDM systems.

As a consequence, in this thesis, we focus on the design of new multicarrier receiver based on Bayesian inference in order to improve the performance and reliability of multicarrier transmission in the presence of phase distortions. More precisely, we will firstly propose an OFDM channel estimator based on a training multicarrier symbol in the presence of phase distortions which deals with the unknowledge of both PHN and AWGN powers. Finally, a joint multicarrier signal and phase distortions estimator will be proposed.

BAYESIAN INFERENCE 2

MANY applications in science and engineering require the estimation of a signal of interest from a noisy observation sequence. Except for linear and Gaussian models where an optimal filtering solution is given by the Kalman filter, this problem is still a major issue. Indeed, in order to model more accurately dynamics of a physical system, it is becoming important to include non-linearity and non-Gaussianity. In these scenarios, various approximation methods have been developed among which the extended Kalman filter is the most commonly used. Since the nineties, sequential Monte Carlo (SMC) approaches, also known as particle filtering, have become a powerful methodology to cope with non-linear and non-Gaussian problems.

Firstly, we begin with the mathematical formulation of the problem statement and the concept of Bayesian filtering. Then, deterministic methods for Bayesian inference such as Kalman filter, its extended version and variational methods are briefly described. Finally, we focus on the SMC principle.

2.1 PROBLEM STATEMENT

2.1.1 State-space model

The state sequence $\{\mathbf{x}_k; k \in \mathbb{N}\}$, $\mathbf{x}_k \in \mathbb{R}^{n_\mathbf{x}}$, is assumed to be an unobserved (hidden) Markov process with initial distribution $p(\mathbf{x}_0)$ (which we subsequently denote as $p(\mathbf{x}_0|\mathbf{x}_{-1})$ for notational convenience) and transition distribution $p(\mathbf{x}_k|\mathbf{x}_{k-1})$, where $n_\mathbf{x}$ is the dimension of the state vector. The observations $\{\mathbf{y}_k; k \in \mathbb{N}\}$, $\mathbf{y}_k \in \mathbb{R}^{n_\mathbf{y}}$, are conditionally independent given the process $\{\mathbf{x}_k; k \in \mathbb{N}\}$ with distribution $p(\mathbf{y}_k|\mathbf{x}_k)$ and $n_\mathbf{y}$ is the dimension of the observation vector. Figure 2.1.1 represents the dependence structure of a hidden Markov model (HMM). As shown in this figure, the distribution of $\mathbf{x}_k$ conditional on the history of the process, $\mathbf{x}_{0:k-1}$, is determined by the value taken by the preceding one $\mathbf{x}_{k-1}$. This is called the Markov property. Likewise, the distribution of $\mathbf{y}_k$ conditionally on the past observations $\mathbf{y}_{0:k-1}$ and the past values of the state, $\mathbf{x}_{0:k-1}$ is determined by $\mathbf{x}_k$ only.

To sum up, the model is a HMM described by :

$$\text{A priori distribution :} \quad p(\mathbf{x}_k|\mathbf{x}_{k-1}) \text{ for } k \geq 0 \tag{2.1}$$

$$\text{Likelihood distribution :} \quad p(\mathbf{y}_k|\mathbf{x}_k) \text{ for } k \geq 0 \tag{2.2}$$

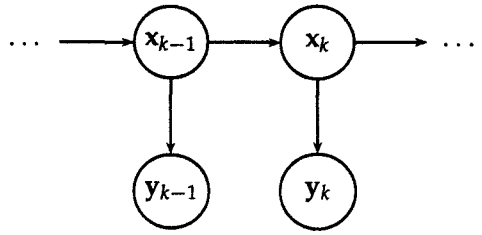

Figure 2.1 – *Graphical representation of the dependence structure of a hidden Markov model.*

Any HMM may be equivalently defined through a functional representation which consists in two equations, known as a state-space model,

$$\text{State Equation :} \quad \mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{v}_k) \tag{2.3}$$

$$\text{Observation Equation :} \quad \mathbf{y}_k = g(\mathbf{x}_k, \mathbf{u}_k) \tag{2.4}$$

where $f(.)$ is a state transition function, $g(.)$ is a measurement transition function, and $\mathbf{v}_k$ and $\mathbf{u}_k$ are independent noise vectors with known distributions. For simplicity, we assume that the analytical forms of the functions $f(.)$, $g(.)$ and the initial probability density of the state $p(\mathbf{x}_0)$ are known.

2.1.2 State Inference

The fundamental issue in hidden Markov modeling is given a fully specified model and some observations $\mathbf{y}_{0:\tau}$, what can be said about the corresponding unobserved state sequence $\mathbf{x}_{0:k}$? The Bayesian approach provides an elegant and consistent method of dealing with uncertainty. Indeed using the Bayes' theorem, the posterior distribution $p(\mathbf{x}_{0:k}|\mathbf{y}_{0:\tau})$ reflects all the information we have about the state of the system $\mathbf{x}_{0:k}$ contained in the observations $\mathbf{y}_{0:\tau}$ and the prior $p(\mathbf{x}_{0:k})$:

$$p(\mathbf{x}_{0:k}|\mathbf{y}_{0:\tau}) = \frac{p(\mathbf{y}_{0:\tau}|\mathbf{x}_{0:k})p(\mathbf{x}_{0:k})}{p(\mathbf{y}_{0:\tau})} \tag{2.5}$$

The generic estimation of the hidden state can be divided in three problems :

- for $\tau = k$, *Filtering*, which means the evaluation of the quantity of interest is achieved using observations up to and including time k.
- for $\tau > k$, *Smoothing*, which differs from filtering in that information about the quantity of interest not need be available at time k and data measured later than time k can be used in obtaining this information. This means that there is a delay in producing the result of interest. Since in the smoothing process more information in the observations is taken into account, we would expect it be more accurate in some sense than the filtering process.
- for $\tau < k$, *Prediction*, which the aim here is to evaluate the quantity of interest at some time k in the future by using only observations up to and including τ.

Without loss of generality, we will consider the filtering problem. When we are interested to estimate a finite sequence $\mathbf{x}_{0:k}$ of hidden states given the observations $\mathbf{y}_{0:k}$, the problem is referred to as *off-line filtering* and the density of interest is the following posterior distribution, $p(\mathbf{x}_{0:k}|\mathbf{y}_{0:k})$. The off-line filtering is used for systems where the data are grouped in batch. On the other hand, when we estimate the hidden state $\mathbf{x}_k$ given the observations $\mathbf{y}_{0:k}$, we make an *on-line filtering*. Depending on the requirements of a given application, we can be interested to consider as target distribution $\pi_k(\mathbf{x})$ the following marginal distribution, called (on-line) filtering distribution :

$$\pi_k(\mathbf{x}) = p(\mathbf{x}_k|\mathbf{y}_{0:k}) \tag{2.6}$$

or the following joint posterior distribution

$$\pi_k(\mathbf{x}) = p(\mathbf{x}_{0:k}|\mathbf{y}_{0:k}) \tag{2.7}$$

The on-line filtering is applied to systems where the observations arrive sequentially in time. The on-line filtering appears in the literature under many different names, including *Bayesian filtering, Bayesian inference* and *optimal filtering*.

Using the target distribution (either (2.6) or (2.7)), an estimate of the hidden state $\mathbf{x}_k$ is obtained using the *Minimum Mean Square Error* (MMSE) criterion:

$$\widehat{\mathbf{x}}_k^{\text{MMSE}} = \mathbb{E}_{\pi_k(\mathbf{x})}[\mathbf{x}_k] = \int \mathbf{x}_k \pi_k(\mathbf{x}) d\mathbf{x}_k \tag{2.8}$$

or using the *Maximum A Posteriori* (MAP) :

$$\widehat{\mathbf{x}}_k^{\text{MAP}} = \arg\max_{\mathbf{x}_k} \pi_k(\mathbf{x}) \tag{2.9}$$

2.1.3 Sequential evaluation of the target distribution

The target distribution required for the estimation of the hidden states can be evaluated sequentially. Indeed in off-line filtering, a recursive formula for the joint probability distribution can be obtained straightforwardly from 2.5 and the dependence structure of HMM :

$$\begin{aligned} p(\mathbf{x}_{0:k}|\mathbf{y}_{0:k}) &= \frac{p(\mathbf{x}_0)p(\mathbf{y}_0|\mathbf{x}_0)\prod_{n=1}^{k} p(\mathbf{y}_n|\mathbf{x}_n)p(\mathbf{x}_n|\mathbf{x}_{n-1})}{p(\mathbf{y}_{0:k})} \\ &= \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{k-1})}{p(\mathbf{y}_k|\mathbf{y}_{0:k-1})} p(\mathbf{x}_{0:k-1}|\mathbf{y}_{0:k-1}) \end{aligned} \tag{2.10}$$

The probability density of interest $p(\mathbf{x}_k|\mathbf{y}_{0:k})$ required in on-line filtering can be obtained by marginalization of (2.10). However the dimension of the integration grows as k increases. This can be avoided by using a sequential scheme since this marginal density also satisfies a recursion [Sorenson 88] :

$$p(\mathbf{x}_k|\mathbf{y}_{0:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1}) d\mathbf{x}_{k-1} \tag{2.11}$$

$$p(\mathbf{x}_k|\mathbf{y}_{0:k}) = \frac{p(\mathbf{y}_k|\mathbf{x}_k)}{p(\mathbf{y}_k|\mathbf{y}_{0:k-1})} p(\mathbf{x}_k|\mathbf{y}_{0:k-1}) \tag{2.12}$$

with

$$p(\mathbf{y}_k|\mathbf{y}_{0:k-1}) = \int p(\mathbf{y}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{y}_{0:k-1}) d\mathbf{x}_k \tag{2.13}$$

Equations (2.11) and (2.12) are respectively called *prediction* and *updating*. Unfortunately, the target distribution is usually intractable, except only in a few cases, based on various hypothesis to ensure the mathematical tractability. For example, the Kalman filter provides the optimal analytical solution for a linear Gaussian state-space model. In the most general case, the presence of non-Gaussianity and nonlinearity precludes an analytical solution.

2.2 DETERMINISTIC METHODS

For a linear and Gaussian model, Kalman and Bucy have noticed that Equations (2.11)-(2.12) could be solved analytically thus producing the Kalman filter [Kalman *et al.* 61]. The Kalman filter was then extended to consider more general non-linear cases. An alternative approach based on deterministic fixed point iterations is the variational method [Jordan 98]. These methods are briefly described in this section.

2.2.1 Kalman filter

Let us consider the following linear state-space model :

$$\mathbf{x}_k = \mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{C}_k \mathbf{v}_k \tag{2.14}$$

$$\mathbf{y}_k = \mathbf{B}_k \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k \tag{2.15}$$

where $\{\mathbf{v}_k\}_{k\geq 1}$ and $\{\mathbf{u}_k\}_{k\geq 1}$ are two mutually independent Gaussian random vectors with zero vector mean and identity covariance matrix and are independent of the initial state $\mathbf{x}_0$. The initial state $\mathbf{x}_0$ is assumed to be Gaussian random vectors with mean $\mathbf{x}_{0|0}$ and covariance matrix $\mathbf{\Sigma}_{0|0}$.

With these assumptions, the probability densities $p(\mathbf{x}_k|\mathbf{y}_{0:k-1})$ and $p(\mathbf{x}_k|\mathbf{y}_{0:k})$ are themselves Gaussian and are uniquely defined by their mean and covariance matrices, denoted respectively by $\{\mathbf{x}_{k|k-1}, \mathbf{\Sigma}_{k|k-1}\}$ and $\{\mathbf{x}_{k|k}, \mathbf{\Sigma}_{k|k}\}$. According to Eqs. (2.11)-(2.12), the Kalman filter proceeds as follows, for $k = 1, ..., T$:

Kalman predicted equations :

$$\mathbf{x}_{k|k-1} = \mathbf{A}_k \mathbf{x}_{k-1|k-1} \tag{2.16}$$

$$\mathbf{\Sigma}_{k|k-1} = \mathbf{A}_k \mathbf{\Sigma}_{k-1|k-1} \mathbf{A}_k^H + \mathbf{C}_k \mathbf{C}_k^H \tag{2.17}$$

Kalman filtered equations :

$$\widetilde{\mathbf{y}}_k = \mathbf{y}_k - \mathbf{B}_k \mathbf{x}_{k|k-1} \quad \text{Innovation} \tag{2.18}$$

$$\mathbf{S}_k = \mathbf{B}_k \boldsymbol{\Sigma}_{k|k-1} \mathbf{B}_k^H + \mathbf{D}_k \mathbf{D}_k^H \quad \text{Innovation covariance} \tag{2.19}$$

$$\mathbf{K}_k = \boldsymbol{\Sigma}_{k|k-1} \mathbf{B}_k^H \mathbf{S}_k^{-1} \quad \text{Kalman gain} \tag{2.20}$$

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{K}_k \widetilde{\mathbf{y}}_k \quad \text{Filter. state estim.} \tag{2.21}$$

$$\boldsymbol{\Sigma}_{k|k} = \boldsymbol{\Sigma}_{k|k-1} - \mathbf{K}_k \mathbf{B}_k \boldsymbol{\Sigma}_{k|k-1} \quad \text{Filter. Error Cov.} \tag{2.22}$$

2.2.2 Extended Kalman filter

Although the Kalman filter is optimal in the linear and Gaussian case, numerous approach of filtering were required to consider more general case. Therefore, various approximation methods have been developed among which the extended Kalman filter (EKF) is probably the most commonly used one. The idea of the EKF is to use a first-order Taylor expansion in order to linearize both the state and the observation equations, and thus approximate the system and noise distribution as Gaussians. The Kalman update equations can then be used to derive the solution for the considering system.

Let us consider the following non-linear and Gaussian state space model :

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}) + \mathbf{C}_k \mathbf{v}_k \tag{2.23}$$

$$\mathbf{y}_k = g(\mathbf{x}_k) + \mathbf{D}_k \mathbf{u}_k \tag{2.24}$$

where $\{\mathbf{v}_k\}_{k\geq 1}$ and $\{\mathbf{u}_k\}_{k\geq 1}$ are two mutually independent Gaussian random vectors with zero vector mean and identity matrix and independent of the initial state x_0. The initial state $\mathbf{x}_0$ is assumed to be Gaussian random vectors with mean $\mathbf{x}_{0|0}$ and covariance matrix $\boldsymbol{\Sigma}_{0|0}$.

Using a first-order Taylor expansion about the current estimate of the mean of the state, *i.e.* $\mathbf{x}_{k-1|k-1}$, we obtain respectively for the state and observation equations :

$$\mathbf{x}_k \approx f(\mathbf{x}_{k-1|k-1}) + \underbrace{\left.\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}_{k-1|k-1}}}_{\mathbf{A}_k} (\mathbf{x}_{k-1} - \mathbf{x}_{k-1|k-1}) + \mathbf{C}_k \mathbf{v}_k \tag{2.25}$$

$$\mathbf{y}_k \approx g(\mathbf{x}_{k|k-1}) + \underbrace{\left.\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}_{k|k-1}}}_{\mathbf{B}_k} (\mathbf{x}_k - \mathbf{x}_{k|k-1}) + \mathbf{D}_k \mathbf{u}_k \tag{2.26}$$

Applying the Kalman equations to these approximations gives the following EKF equations :

EKF predicted equations :

$$\mathbf{x}_{k|k-1} = f(\mathbf{x}_{k-1|k-1}) \tag{2.27}$$

$$\boldsymbol{\Sigma}_{k|k-1} = \mathbf{A}_k \boldsymbol{\Sigma}_{k-1|k-1} \mathbf{A}_k^H + \mathbf{C}_k \mathbf{C}_k^H \tag{2.28}$$

EKF filtered equations :

$$\widetilde{\mathbf{y}}_k = \mathbf{y}_k - g(\mathbf{x}_{k|k-1}) \quad \text{Innovation} \tag{2.29}$$

$$\mathbf{S}_k = \mathbf{B}_k \boldsymbol{\Sigma}_{k|k-1} \mathbf{B}_k^H + \mathbf{D}_k \mathbf{D}_k^H \quad \text{Innovation covariance} \tag{2.30}$$

$$\mathbf{K}_k = \boldsymbol{\Sigma}_{k|k-1} \mathbf{B}_k^H \mathbf{S}_k^{-1} \quad \text{Kalman gain} \tag{2.31}$$

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{K}_k \widetilde{\mathbf{y}}_k \quad \text{Filter. state estim.} \tag{2.32}$$

$$\boldsymbol{\Sigma}_{k|k} = \boldsymbol{\Sigma}_{k|k-1} - \mathbf{K}_k \mathbf{B}_k \boldsymbol{\Sigma}_{k|k-1} \quad \text{Filter. Error Cov.} \tag{2.33}$$

Unlike its linear counterpart, the EKF is not an optimal estimator. However, the filter works well for a weakly non-linear system. Divergence is a further problem with the EKF occurring due to model mis-specification. Since the model that the Kalman filter is solving is inaccurate, the EKF can often considerably underestimate the covariance of its error estimate of the state.

Remark : The unscented Kalman filter (UKF) [Julier *et al.* 97] is an other variant of the Kalman filter in the non-linear case. This method uses a deterministic sampling technique known as the unscented transform to pick a minimal set of sample points (called sigma points) around the mean. However, this approach remains an approximation of the Bayes' solution in the non-linear case and as the previous ones can not be used in more general models like for example in non-Gaussian models.

2.2.3 Variational Bayesian methods

Variational methods [Jordan 98] provide another approach to the design of approximate inference algorithms. Variational methodology yields deterministic approximation procedures that generally provide bounds on probabilities of interest. In the Bayesian framework, the aim of variational inference is to find a tractable and accurate posterior approximation to an intractable posterior distribution.

The variational Bayes method approximates the joint *p.d.f.* $p(\{\mathbf{x}_{0:k}^i\}_{i=1}^I | \mathbf{y}_{0:k})$ of $I \geq 2$ vector variables $\{\mathbf{x}_{0:k}^i\}_{i=1}^I$ by the following separable *p.d.f.* using the *mean field* approximation :

$$q(\left\{\mathbf{x}_{0:k}^i\right\}_{i=1}^I | \mathbf{y}_{0:k}) = \prod_{i=1}^I q(\mathbf{x}_{0:k}^i | \mathbf{y}_{0:k}) \tag{2.34}$$

The approximation forces posterior independence between subsets of states in a particular partition of $\mathbf{x}_{0:k}$ chosen by the designer. The optimal set of "approximate" marginal distributions are chosen by minimizing a particular measure of divergence between $q(\{\mathbf{x}_{0:k}^i\}_{i=1}^I | \mathbf{y}_{0:k})$ and $p(\{\mathbf{x}_{0:k}^i\}_{i=1}^I | \mathbf{y}_{0:k})$, namely Kullback-Leibler (KL) divergence defined by :

$$D_{\mathrm{KL}}(q||p) = \int q(\left\{\mathbf{x}_{0:k}^i\right\}_{i=1}^I | \mathbf{y}_{0:k}) \ln \frac{p(\{\mathbf{x}_{0:k}^i\}_{i=1}^I | \mathbf{y}_{0:k})}{q(\{\mathbf{x}_{0:k}^i\}_{i=1}^I | \mathbf{y}_{0:k})} d\left\{\mathbf{x}_{0:k}^i\right\}_{i=1}^I \tag{2.35}$$

The marginal distributions are therefore specified, up to a constant additive term,

by the following system of equations, for $i = 1, 2, ..., I$ [Smidl *et al.* 06]:

$$\begin{aligned} \ln q(\mathbf{x}^i_{0:k}|\mathbf{y}_{0:k}) &= \mathbb{E}_{\prod_{\substack{i=1 \\ \neq i}}^{I} q(\mathbf{x}^i_{0:k}|\mathbf{y}_{0:k})} \left[\ln p(\left\{ \mathbf{x}^i_{0:k} \right\}_{i=1}^{I} |\mathbf{y}_{0:k}) \right] \\ &= \ln p(\mathbf{x}^i_{0:k}|\mathbf{y}_{0:k}) + \mathbb{E}_{\prod_{\substack{i=1 \\ \neq i}}^{I} q(\mathbf{x}^i_{0:k}|\mathbf{y}_{0:k})} \left[\ln p(\left\{ \mathbf{x}^j_{0:k} \right\}_{\substack{j=1 \\ \neq i}}^{I} |\mathbf{y}_{0:k}) \right] \end{aligned} \quad (2.36)$$

The solution of the above system of equations provides the distribution $q(\{\mathbf{x}^i_{0:k}\}_{i=1}^{I}|\mathbf{y}_{0:k})$ that minimize the KL distance. When the distributions involved in $p(\left\{ \mathbf{x}^j_{0:k} \right\}_{\substack{j=1 \\ \neq i}}^{I} |\mathbf{y}_{0:k})$ belong to the class of exponential distributions and are conjugate to each other, the solution of the system can be found in a straightforward manner [Smidl *et al.* 06]. However in more general problems, a closed form solution is usually not available, so that the solution has to be found iteratively [Beal 03]. Indeed, even with this approach, the solution of the system may still remain intractable, due to the complex form of marginal and conditional distributions required in Eq. (2.36). In this case, the usual practice is to update each one of them in turn, while holding the others constant, a simple technique termed *coordinate descent* in the optimization literature [Nocedal *et al.* 99].

The algorithm is guaranteed to converge to a local minimum of the free energy expression [Neal *et al.* 98]. Indeed, by forcing the posterior distribution to be a product of *p.d.f.*, the variational method generally does not converge to the global maximum. Variational Bayesian methods nevertheless also suffer from slow convergence when the variables in the factored approximation are strongly correlated in the original model.

After the description of deterministic methods, we now describe in details in the next section the sequential Monte Carlo (SMC) which shows a great promising in providing a general methodology for non-linear non-Gaussian filtering. Indeed in complex models, the main advantages of stochastic methods, compared to deterministic ones, are its generality, robustness and attractive theorical properties.

2.3 Sequential Monte-Carlo methods

2.3.1 The Monte Carlo principle

Monte Carlo (MC) methods are commonly used for approximation of intractable integrals and rely on the ability to draw a random sample from the required probability distribution. The idea is to simulate M *independent and identically distributed* (*i.i.d.*) samples $\left\{ \mathbf{x}^{(j)}_{0:K} \right\}_{j=1}^{M}$ from the distribution of interest, which is the posterior distribution $p(\mathbf{x}_{0:K}|\mathbf{y}_{0:K})$ in the Bayesian framework, and use them to obtain an empirical estimate of the distribution by the following point-mass measure :

$$\widehat{p}(\mathbf{x}_{0:K}|\mathbf{y}_{0:K}) = \frac{1}{M} \sum_{j=1}^{M} \delta(\mathbf{x}_{0:K} - \mathbf{x}^{(j)}_{0:K}) \quad (2.37)$$

where $\delta(.)$ denotes the Dirac delta function.

Consequently, the expected value of any function $g(.)$ of $\mathbf{x}_{0:k}$ can be approximated by the following sum :

$$\mathbb{E}_{\widehat{p}(\mathbf{x}_{0:K}|\mathbf{y}_{0:k})}[g(\mathbf{x}_{0:K})] = \frac{1}{M}\sum_{j=1}^{M} g(\mathbf{x}_{0:K}^{(j)}) \tag{2.38}$$

The estimate (2.38) is unbiased with the variance proportional to $1/M$ for the finite variance of $g(z_{0:K})$ [Doucet *et al.* 01], and is easily obtained providing we can sample from $p(\mathbf{x}_{0:K}|\mathbf{y}_{0:K})$. It is possible to generate samples from most standard distributions (e.g., uniform, Gaussian, Gamma, Poisson, etc.) using standard techniques. Most of these techniques are based on the inverse cumulative distribution function (CDF) or on the acceptance-rejection method [Devroye 86]. Note that the inverse cdf method is only applicable to simple probability distributions that are known exactly rather than up to a normalizing constant. On the other hand, the acceptance-rejection method does not require knowledge of the normalizing constant but is inefficient for high dimensional distributions since the acceptance probability of the candidate is typically extremely small.

2.3.2 Importance sampling

A simple alternative to the acceptance-rejection method is the importance sampling. Suppose we cannot efficiently sample from $p(\mathbf{x}_{0:K}|\mathbf{y}_{0:K})$, but, there is another arbitrary convenient probability distribution $\pi(\mathbf{x}_{0:K}|\mathbf{y}_{0:K})$ which is easy to sample from. Assuming that $p(\mathbf{x}_{0:k}|\mathbf{y}_{0:K}) > 0$ implies $\pi(\mathbf{x}_{0:K}|\mathbf{y}_{0:K}) > 0$, then the following identity holds trivially

$$p(\mathbf{x}_{0:K}|\mathbf{y}_{0:K}) = \frac{w(\mathbf{x}_{0:K})\pi(\mathbf{x}_{0:K}|\mathbf{y}_{0:K})}{\int w(\mathbf{x}_{0:K})\pi(\mathbf{x}_{0:K}|\mathbf{y}_{0:K})d\mathbf{x}_{0:K}} \tag{2.39}$$

where $w(.)$ is the so-called unnormalized importance weight given by

$$w(\mathbf{x}_{0:K}) = \frac{p(\mathbf{x}_{0:K}|\mathbf{y}_{0:K})}{\pi(\mathbf{x}_{0:K}|\mathbf{y}_{0:K})} \tag{2.40}$$

This suggests that if M samples $\left\{\mathbf{x}_{0:K}^{(j)}\right\}_{j=1}^{M}$ from $\pi(\mathbf{x}_{0:K}|\mathbf{y}_{0:K})$ are available, then an approximation of the distribution is given by :

$$\widehat{p}(\mathbf{x}_{0:K}|\mathbf{y}_{0:K}) = \sum_{j=1}^{M} \widetilde{w}^{(j)}\delta(\mathbf{x}_{0:K} - \mathbf{x}_{0:K}^{(j)}) \tag{2.41}$$

where $\widetilde{w}^{(j)} = w(\mathbf{x}_{0:K}^{(j)})/\sum_{i=1}^{M} w(\mathbf{x}_{0:K}^{(i)})$ is the normalized importance weights.

Using importance sampling, the expected value of any function $g(.)$ of $\mathbf{x}_{0:K}$ can thus

be approximated by:

$$\mathbb{E}_{\widehat{p}(\mathbf{x}_{0:K}|\mathbf{y}_{0:K})}[g(\mathbf{x}_{0:K})] = \sum_{i=1}^{M} \widetilde{w}^{(j)} g(\mathbf{x}_{0:K}^{(i)}) \tag{2.42}$$

This estimate is biased since a ratio of estimates is involved in the normalization step of the importance weights. However, the bias is of the order $\mathcal{O}(1/M)$ and the estimate is also asymptotically consistent [Robert *et al.* 99].

2.3.3 Sequential importance sampling

The method described up to now is an off-line method and is not 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. A sequential method for importance sampling can be obtained by taking an importance function with the following form :

$$\pi(\mathbf{x}_{0:k}|\mathbf{y}_{0:k}) = \pi(\mathbf{x}_{0:k-1}|\mathbf{y}_{0:k-1})\pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{0:k}) \tag{2.43}$$

By substituting Eqs. (2.43) and (2.10) into (2.40), the following recursive expression for the computation of the importance weights can be obtained :

$$\begin{aligned} w_k^{(j)} &\propto \frac{p(\mathbf{x}_{0:k-1}|\mathbf{y}_{0:k-1})p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{y}_k|\mathbf{x}_k)}{\pi(\mathbf{x}_{0:k-1}|\mathbf{y}_{0:k-1})\pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{0:k})} \\ &\propto w_{k-1}^{(j)} \frac{p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{y}_k|\mathbf{x}_k)}{\pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{0:k})} \end{aligned} \tag{2.44}$$

The Sequential Importance sampling (SIS) algorithm shown in Table 2.1 approximates posterior *p.d.f.* $p(\mathbf{x}_{0:k}|\mathbf{y}_{0:k})$ by a set of M random samples with associated weights, denoted by $\{\mathbf{x}_{0:k}^{(j)}, \widetilde{w}_k^{(j)}\}_{j=1}^{M}$ [Doucet *et al.* 01] :

$$\widehat{p}(\mathbf{x}_{0:k}^{(j)}|\mathbf{y}_{0:k}) = \sum_{j=1}^{M} \delta(\mathbf{x}_{0:k} - \mathbf{x}_{0:k}^{(j)})\widetilde{w}_k^{(j)} \tag{2.45}$$

where $\widetilde{w}_k^{(j)} = w_k^{(j)} / \sum_{m=1}^{M} w_k^{(m)}$ is the normalized importance weight of the j-th particle.

Initialization Sample $\left\{\mathbf{x}_{-1}^{(j)}\right\}_{j=1}^{M} \sim p(\mathbf{x}_{-1})$ and set $\left\{\widetilde{w}_0^{(j)}\right\}_{j=1}^{M}$
For $k = 0...$
  For $j = 1...M$
    Sample $\mathbf{x}_k^{(j)} \sim \pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}^{(j)}, \mathbf{y}_{0:k})$ and set $\mathbf{x}_{0:k}^{(j)} = (\mathbf{x}_{0:k-1}^{(j)}, \mathbf{x}_k^{(j)})$
    Compute the importance weights using $w_k^{(j)} \propto w_{k-1}^{(j)} \frac{p(\mathbf{x}_k^{(j)}|\mathbf{x}_{k-1}^{(j)})p(\mathbf{y}_k|\mathbf{x}_k^{(j)})}{\pi(\mathbf{x}_k^{(j)}|\mathbf{x}_{0:k-1}^{(j)}, \mathbf{y}_{0:k})}$
  end for
  Normalize the importance weights : $\left\{\widetilde{w}_k^{(j)}\right\}_{j=1}^{M} = \left\{w_k^{(j)}\right\}_{j=1}^{M} / \sum_{m=1}^{M} w_k^{(m)}$
end for

Table 2.1 – *Sequential Importance Sampling*

Let us remark that the SIS algorithm has the main advantage of being parallelizable

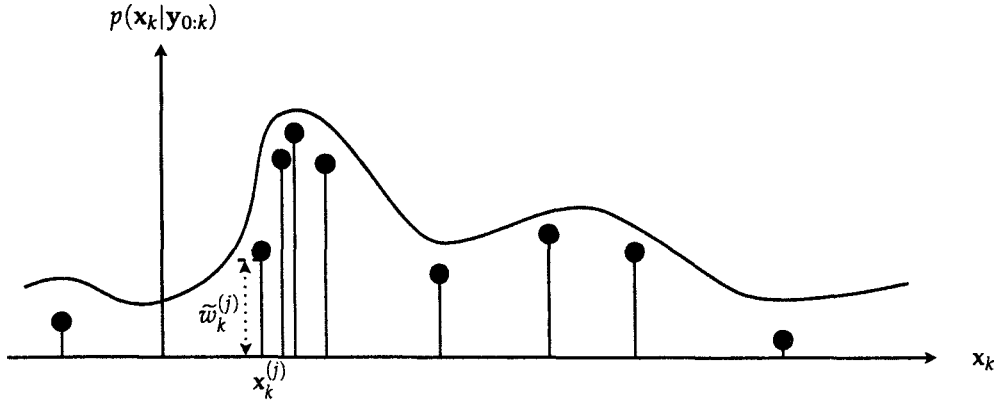

Figure 2.2 – *Approximation of the filtering distribution using the SIS*

and its numerical complexity is $O(M)$. Since it is necessary to keep all the simulated trajectories in order to obtain the approximation in (2.45), the memory requirements are $O((k+1)M)$. However, if we focus in the estimation of the filtering distribution $p(\mathbf{x}_k^{(j)}|\mathbf{y}_{0:k})$ as illustrated in Fig. 2.2 and the chosen importance function depend only on the previous particle value, the memory requirements reduces to $O(M)$.

2.3.4 Resampling step

Unfortunately, the SIS algorithm is known to suffer from degeneracy problems. Indeed after few iterations, the variance of the importance weights quickly increases, *i.e.* only a few particles are assigned non-negligible weights. As a consequence, since the approximation of the target distribution is made using a very small number of particles, the inference accuracy is degraded. To limit this effect, a resampling step is integrated in the SIS algorithm. The idea of resampling is to remove the particle trajectories with small weights and replicate the trajectories with large weights as illustrated by Fig. 2.3. The resampling step is therefore critical in every implementation of SMC filters.

In the literature, different resampling methods can be found. The most frequently encountered algorithms are multinomial resampling [Gordon *et al.* 93], stratified resampling [Doucet *et al.* 01], systematic resampling [Arulampalam *et al.* 02] and residual resampling [Liu *et al.* 98]. Convergence results have been derived for some of them, see e.g. [Crisan *et al.* 02, Douc. *et al.* 05].

Residual and systematic resampling are to be preferred, as they lead to smaller Monte Carlo errors [Hol *et al.* 06]. However, systematic resampling is the most commonly used since it is the fastest resampling algorithm for computer simulations [Hol *et al.* 06, Bolic *et al.* 04]. As a consequence, the systematic resampling described in Table 2.2, is therefore used for all SMC filters derived in this thesis.

It is now desirable to give guidance on when to do resampling. Indeed, resampling too many times often leads to severe sample impoverishment [Doucet *et al.* 00, Doucet *et al.* 01]. Another method consists in performing occasionally the resampling stage when it seems to be needed. Liu and Chen [Liu *et al.* 98] have introduced a measure known as the *effective sample size* $M/(1+\mathrm{var}(\widetilde{w}_k^{(j)}))$, whose an estimate is

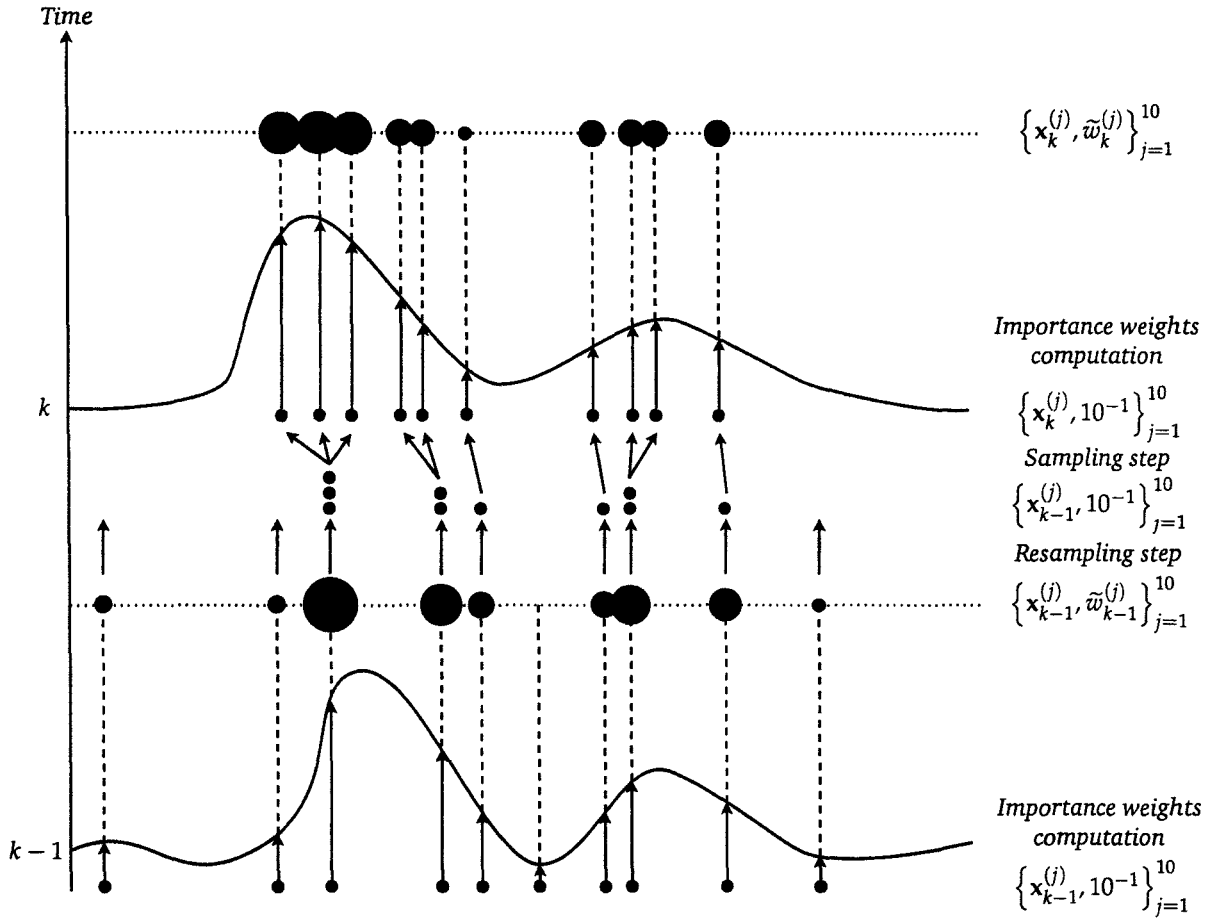

Figure 2.3 – *SIS with resampling strategy*

$N_{eff} = [\sum_{m=1}^{M}(\tilde{w}_k^{(m)})^2]^{-1}$. In this case, the resampling stage is performed whenever N_{eff} is below a predefined threshold.

For $j = 1...M$
Generate a number : $u_j = \frac{(j-1)+\tilde{u}}{M}$, with $\tilde{u} \sim \mathcal{U}[0,1)$
Select the new particle index according to : $I_j = F^{-1}(u_j)$
where F^{-1} denotes the inverse of the cumulative *p.d.f.* of the normalized importance weights.
end for
Set particles trajectories : $\left\{x_{0:k}^{(j)}, \tilde{w}_k^{(j)}\right\}_{j=1}^{M} = \left\{x_{0:k}^{I_j}, 1/M\right\}_{j=1}^{M}$

Table 2.2 – *Systematic Resampling*

2.3.5 Choice of the importance function

The choice of the importance function is essential because it determines the efficiency as well as the complexity of the SMC algorithm. The most popular choices are described below.

2.3.5.1 *Prior distribution*

The prior distribution is the most widely used importance density which is due to its implementation simplicity [Gordon *et al.* 93]. Using Eq. (2.44), if $\pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{0:k}) = p(\mathbf{x}_k|\mathbf{x}_{k-1})$, the importance weights are then computed as follows :

$$w_k^{(j)} \propto w_{k-1}^{(j)} p(\mathbf{y}_k|\mathbf{x}_k^{(j)}) \tag{2.46}$$

Since this distribution does not incorporate any information contained in the most recent observation, this choice may be inefficient and especially sensitive to outliers.

2.3.5.2 *Optimal distribution*

The optimal distribution was introduced in [Zaritskii *et al.* 75] and has the following form :

$$\pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{0:k}) = p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_k) \tag{2.47}$$

This choice is optimal in a sense that this distribution minimizes the variance of the importance weights conditional upon the particle trajectories and the observations [Doucet *et al.* 01]. Using Eq. (2.44), the weights in this case are obtained as follows :

$$w_k^{(j)} \propto w_{k-1}^{(j)} p(\mathbf{y}_k|\mathbf{x}_{k-1}^{(j)}) \tag{2.48}$$

Let us remark that the update of the importance weights does not depend on the current value of the state, *i.e.* $\mathbf{x}_k^{(j)}$. Unfortunately, for many models, the distributions $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_k)$ and $p(\mathbf{y}_k|\mathbf{x}_{k-1}^{(j)})$,

$$p(\mathbf{y}_k|\mathbf{x}_{k-1}^{(j)}) = \int p(\mathbf{y}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{x}_{k-1}) d\mathbf{x}_k \tag{2.49}$$

are often analytically untractable. In this case, we can resort to an approximation of this optimal distribution. On the one hand, in a similar way to the EKF, the optimal importance function can be approximated by a Gaussian distribution using a local linearization of the state-space model when the driving noise of both the state and measurement equations are two mutually *i.i.d.* Gaussian sequences [Doucet *et al.* 00]. On the other hand, when the a priori and/or the likelihood *p.d.f* are not Gaussian, we can resort to the local linearization of the optimal importance function instead of state space linearization leading generally to a Gaussian approximation [Doucet *et al.* 00]. However, the unimodality of true optimal importance function must be verified in order to obtain an accurate Gaussian approximation and thus an efficient SMC filter [Vaswani 07].

2.3.6 Rao-Blackwellization technique

As noted by [Doucet *et al.* 00, Liu *et al.* 98], a well known technique in mathematical statistics, the Rao-Blackwellization [Casella *et al.* 96], can be successfully applied to

particle filtering. A Rao-Blackwellized SMC filter, also known as mixture kalman filter [Chen *et al.* 00] or marginalized particle filter [Schon *et al.* 05], could be derived in order to marginalize the states appearing linearly in the dynamic state-space model.

The basic idea is to partition the state vector as $\mathbf{x}_k = \{\mathbf{x}_k^l, \mathbf{x}_k^n\}$ where $\mathbf{x}_k^l$ denotes the state variable having linear dynamics conditionally upon the nonlinear state $\mathbf{x}_k^n$. Then, using Bayes' theorem, the posterior density of interest can be decomposed as follows :

$$p(\mathbf{x}_k^l, \mathbf{x}_{0:k}^n|\mathbf{y}_{0:k}) = p(\mathbf{x}_k^l|\mathbf{x}_{0:k}^n, \mathbf{y}_{0:k})p(\mathbf{x}_{0:k}^n|\mathbf{y}_{0:k}) \tag{2.50}$$

where $p(\mathbf{x}_k^l|\mathbf{x}_{0:k}^n, \mathbf{y}_{0:k})$, unlike $p(\mathbf{x}_{0:k}^n|\mathbf{y}_{0:k})$, is analytically tractable and is given by the Kalman filter. The marginal posterior distribution $p(\mathbf{x}_{0:k}^n|\mathbf{y}_{0:k})$ is approximated using SMC methodology as :

$$p(\mathbf{x}_{0:k}^n|\mathbf{y}_{0:k}) \approx \sum_{j=1}^{M} \delta(\mathbf{x}_{0:k}^n - \mathbf{x}_{0:k}^{n,(j)})\widetilde{w}_k^{(j)} \tag{2.51}$$

Thus, by substituting (2.51) in (2.50), the posterior distribution is given by :

$$p(\mathbf{x}_k^l, \mathbf{x}_{0:k}^n|\mathbf{y}_{0:k}) \approx \sum_{j=1}^{M} p(\mathbf{x}_k^l|\mathbf{x}_{0:k}^{n,(j)}, \mathbf{y}_{0:k})\delta(\mathbf{x}_{0:k}^n - \mathbf{x}_{0:k}^{n,(j)})\widetilde{w}_k^{(j)} \tag{2.52}$$

where $p(\mathbf{x}_k^l|\mathbf{x}_{0:k}^{n,(j)}, \mathbf{y}_{0:k})$ is a multivariate Gaussian probability density function with mean $\widehat{\mathbf{x}}_{k|k}^{l,(j)}$ and covariance matrix $\mathbf{\Sigma}_{k|k}^{(j)}$ which are both obtained using the Kalman filtering equations (2.16)-(2.22). This technique provides a reduced variance of the estimates of interest [Doucet *et al.* 00] leading to a more efficient SMC filter[1] since the particles have to explore a lower state-space dimension (*i.e.* $n_{\mathbf{x}^n} < n_{\mathbf{x}}$).

CONCLUSION

In the Bayesian framework, the posterior distribution is the main quantity of interest since it reflects all the information we have about the state of the system contained in the observations. In linear and Gaussian model, a deterministic method, namely the Kalman filter, provides the optimal solution to the Bayesian filtering problem. However in more general and complex models, optimal solution is analytically intractable. Various deterministic schemes , such as EKF or variational methods, have therefore been proposed to approximate the distribution of interest. Variational methods are interesting technique but they do not ensure a convergence to the global maximum due to the *mean field* approximation required to derive the algorithm.

Simulation-based methods represent another class of algorithms used to approximate a target distribution. The aim of stochastic methods is to generate *i.i.d.* samples from the distribution of interest. The main advantage of stochastic methods, compared

[1]Other techniques, such as the Auxiliary particle filter [Pitt *et al.* 99] and the Regularized particle filter [Musso *et al.* 01], have been proposed in the literature to improve the efficiency of the classical SIS algorithm with resampling.

to deterministic ones, is its generality, robustness and attractive theorical properties. Indeed, they can yield approximations to an arbitrary level of accuracy, but at the expense of higher complexity. In particular, amongst sampling methods, the sequential Monte Carlo approach appears as a powerful tools for tracking time-varying signals in general non-linear and non-Gaussian dynamic systems.

In this chapter, parameters which characterize both the prior and likelihood function in the DSS representation, *e.g.* the *p.d.f.* mean and variance, have been assumed perfectly known. However, a such assumption is not realistic in practice. As a consequence, the next chapter is devoted to the parameter estimation problem in general state space model containing also hidden dynamic states. More precisely, we will focus on SMC-based algorithms.

3 PARAMETER ESTIMATION USING SMC METHODS

STOCHASTIC filtering theory is concerned with the estimation of the posterior distribution of a stochastic process given some observations. In this case, the filter is completely determined by the observations and the DSS model as shown in Chapter 2. However, in many cases of interest, both the state transition density and the conditional likelihood function are assumed to depend not only upon the dynamic state x_t but also on unknown static parameter vector λ, i.e. :

$$\begin{cases} x_t \sim p(x_t|x_{t-1}, \lambda) \\ y_t \sim p(y_t|x_t, \lambda) \end{cases} \tag{3.1}$$

Depending on the requirements of a given application, model parameters estimation can be carried out in two different modes. On the one hand, if the data is available in a batch beforehand, the estimation of the parameters will generally be done prior to the state inference, which is referred to as batch or off-line estimation. However in this off-line context, due to the presence of the hidden state, the log-likelihood function of the observations $l_T(\lambda) \triangleq \log p(y_{0:T}|\lambda)$ cannot be easily maximized, except in some particular cases. A generic approach for latent variable models is provided by the Expectation-Maximization (EM) algorithm [Dempster *et al.* 77]. However, in many complex problems such as nonlinear and/or non-Gaussian models, parameter estimate using EM algorithm is analytically intractable. In this case, stochastic versions of EM-type algorithms, like the Monte-Carlo EM (MCEM) [Wei *et al.* 90], the Stochastic Approximation of the EM (SAEM) [Delyon *et al.* 99], can be employed to obtain an approximation of the expectation step of the EM algorithm. It has been observed by several authors that this expectation can be approximated also using SMC methods. However, when processing large data sets or data streams, off-line algorithms for parameter estimation become impractical due to the requirement that the whole data be available at each iteration of the algorithm.

On the other hand, in some cases, the parameters have to be estimated sequentially without the data being stored, which is referred to as on-line estimation. Due to the recursive nature of SMC methods, we can resort to theses methods for on-line parameter estimation. Different approaches have been proposed in the literature to solve this

problem using SMC methods. These several schemes can be divided into two strategies : the point estimation and the filtering estimation methods. The first strategy aims to give point estimate of the parameter λ from $\{y_{0:t}\}$ without a preliminary approximation of the posterior distribution $p(\lambda|y_{0:t})$. This approach have generally good convergence property but at the expense of a slow convergence rate. The second strategy is filtering method which consists in approximating the joint posterior distribution of both the dynamic and static states $p(x_{0:t}, \lambda|y_{0:t})$, and thus $p(\lambda|y_{0:t})$ in order to deduce an estimation of λ. As a consequence, with filtering methods, particle method is required in the parameter space. In comparison to point estimation methods, a diversity in the search space of the parameters is therefore added, thus increasing the convergence speed of the parameter estimation. However, existing filtering schemes suffer from high complexity and/or path degeneracy problems inherent to SMC methods. As a consequence, we propose an efficient parameter estimation using SMC methods which benefits from advantages of both point and filtering estimation methods.

The remainder of the chapter is organized as follows. Firstly in Section 3.1, off-line methods based on SMC methods are briefly reviewed. Secondly, we describe the existing approaches for on-line parameter estimation using sequential Monte-Carlo methods in Section 3.2. Then, we describe the proposed scheme for on-line parameter estimation. Finally, simulation results are presented in two different models and performance comparisons between the proposed method and the on-line Expectation maximization are assessed.

3.1 Off-line Estimation

3.1.1 MC-based algorithms

3.1.1.1 Monte Carlo Expectation Maximization

Since the paper of Dempster et al. [Dempster *et al.* 77], the Expectation-Maximization (EM) algorithm has become a highly appreciated tool for maximum likelihood 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_{i-1} of the parameter, the estimate at the next iteration (i) is given by :

$$\lambda_i = \arg\max_{\lambda} Q(\lambda, \lambda_{i-1}) \tag{3.2}$$

where

$$Q(\lambda, \lambda_{i-1}) = \int \log p(x_{0:T}, y_{0:T}|\lambda) p(x_{0:T}|y_{0:T}, \lambda_{i-1}) dx_{0:T} \tag{3.3}$$

with T the length of observations available for 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 [Wei *et al.* 90], leading to the principle of the Monte Carlo EM (MCEM).

Using MC methods and importance sampling for the estimation of $p(x_{0:T}|y_{0:T}, \lambda_{i-1})$, Eq. (3.3) is approximated by :

$$\widehat{Q}(\lambda, \lambda_{i-1}) = \sum_{j=1}^{M(i)} \widetilde{w}^{(j)} \log p(x_{0:T}^{(j),i}, y_{0:T}|\lambda) \tag{3.4}$$

where $\left\{x_{0:T}^{(j),i}\right\}_{j=1}^{M(i)}$ are the $M(i)$ samples of missing data obtained at the i-th iteration.

This function can be easily maximized when $p(x_{0:T}^{(j),i}, y_{0:T}|\lambda)$ belongs to the exponential family. In this case, $p(x_{0:T}, y_{0:T}|\lambda)$ depends on $(x_{0:T}, y_{0:T})$ only through a set of sufficient statistics. On the one hand, if $\left\{x_{0:T}^{(j),i}\right\}_{j=1}^{M(i)}$ is sampled directly from $p(x_{0:T}|y_{0:T}, \lambda_{i-1})$ then $\left\{\widetilde{w}^{(j)}\right\}_{j=1}^{M(i)} = 1/M(i)$. If direct sampling from the posterior distribution is not possible, samples from $p(x_{0:T}|y_{0:T}, \lambda_{i-1})$ can also be obtained using MCMC methods [Andrieu *et al.* 03b]. On the other hand, if $M(i) = 1$, MCEM reduces to the Stochastic Expectation Maximization (SEM) [Celeux *et al.* 85] :

$$\widehat{Q}(\lambda, \lambda_{i-1}) = \log p(x_{0:T}^{i}, y_{0:T}|\lambda) \tag{3.5}$$

with $x_{0:T}^{i} \sim p(x_{0:T}|y_{0:T}, \lambda_{i-1})$.

However, when approximating Q by $\widehat{Q}$, the Monte Carlo sample size $M(i)$ has to be increased as the algorithm moves along. In fact, Booth et al. argue that MCEM will never converge if $M(i)$ is held fixed across iterations because of a persevering Monte Carlo error [Booth *et al.* 01].

3.1.1.2 *Stochastic Approximation of Expectation Maximization*

A variant of the basic approach is to use a stochastic approximation version of EM, the SAEM [Delyon *et al.* 99]. 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 (3.3) is replaced by the following stochastic averaging procedure :

$$\widehat{Q}_i(\lambda) = (1 - \gamma_i)\widehat{Q}_{i-1}(\lambda) + \gamma_i \sum_{j=1}^{M} \widetilde{w}^{(j)} \log p(x_{0:T}^{(j),i}, y_{0:T}|\lambda) \tag{3.6}$$

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

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 $\widehat{Q}_i(\lambda)$ with a forgetting factor inversely proportional to the step-size.

3.1.2 SMC-based algorithm

Maximum likelihood estimation approaches based on EM techniques share the common feature of requiring the approximation of the expectation of a sum functional of hidden states, conditionally on all the available observations, Eq. (3.3). Such quantities can also be reliably approximated using sequential Monte-Carlo methods as shown in [Cappé *et al.* 05, Doucet *et al.* 03, Andrieu *et al.* 05, Andrieu *et al.* 03a, Cérou *et al.* 01]. Compared to Monte-Carlo approximations of the Q function in the previous section, the resampling step integrated in SMC methods allows to prevent importance sampling dimension from blowing up with time index T.

In the context of the EM algorithm, $Q(\lambda, \lambda_{i-1})$ in (3.3) needs to be estimated. Quite naturally, this approximation based on SMC methods then consists in propagating a system of particle trajectories and associated weights under the current value of the parameters λ_{i-1}, which is denoted by $\left\{x_{0:T}^{(j),i}, \widetilde{w}_T^{(j)}\right\}_{j=1}^{M}$:

$$\widehat{Q}(\lambda, \lambda_{i-1}) = \sum_{j=1}^{M} \widetilde{w}_T^{(j)} \log p(x_{0:T}^{(j),i}, y_{0:T}|\lambda) \tag{3.7}$$

However, all SMC algorithms using a fixed and finite number of particles suffer from a common problem, namely *path degeneracy* which is the result of the resampling stage. As a consequence, it is not possible to approximate properly the smoothed posterior distribution $p(x_{0:T}|y_{0:T}, \lambda)$ as illustrated in Fig. 3.1. To cope with this problem, a modification, relying on forgetting properties of filtering dynamics of the standard method described above is proposed in [Olsson *et al.* 06].

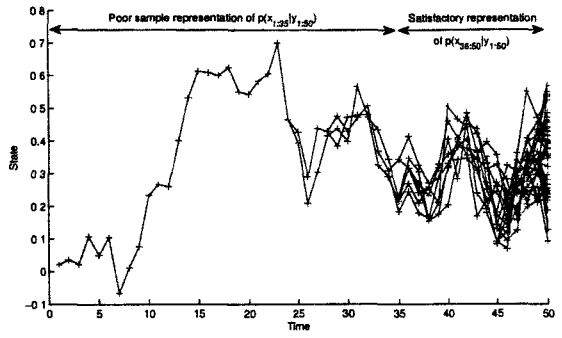

Figure 3.1 – *Path degeneracy illustration of SMC methods (50 particles) with a noisily Gaussian AR(1) ($\phi = 0.95$, $\sigma_v^2 = 0.01$ and $\sigma_b^2 = 0.01$, see Section 3.3.1.1 for details of model).*

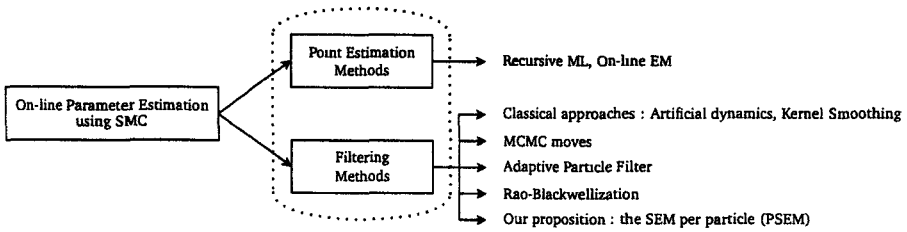

Figure 3.2 – *On-line Parameter Strategies using SMC methods.*

3.2 On-line Estimation

In this section, we focus on the on-line estimation of the static parameter λ. As shown in Fig. 3.2 two different methodologies can be employed in the on-line context: the point estimation and the filtering method in which several schemes have been proposed. This section is devoted to the brief review of particle based methods that have been proposed to solve the problem of the on-line parameter estimation. The proposed scheme based on a stochastic EM per particle is also described.

3.2.1 Point-Estimation Methods

Point-estimation can be done with Recursive ML [Doucet *et al.* 03], gradient-free Stochastic approximation (SA) methods [Poyiadjis *et al.* 06] or on-line EM [Andrieu *et al.* 03a] using SMC methods. In the following of this section, we focus on the on-line EM algorithm.

3.2.1.1 On-line EM

Since the EM algorithm is a numerically well-behaved gradient method, an on-line version of the EM algorithm can be used to obtain sequentially a new parameter estimate. However, similarly to the off-line case, although a direct implementation of the on-line EM algorithm using particle methods is feasible in the nonlinear non-Gaussian case, it would fail in practice because of an accumulation of errors over time.

In [Andrieu *et al.* 03a, Andrieu *et al.* 05], authors propose a general methodology which allows to compute asymptotically consistent point estimates of λ for a large class of dynamic models using an on-line EM algorithm. To prevent the degeneracy inherent to SMC methods, the contrast function to minimize is modified. Indeed, instead of considering the maximization of the average log-likelihood function, authors consider the so-called split-data likelihood (SDL) also called quasi-likelihood as proposed in [Ryden 94] for finite state-space HMM. In this approach, the data is divided in blocks of L data and we maximize the average of the resulting log-SDL. This leads to an alternative Kullback-Leibler contrast function. It can be shown under regularity assumptions that the set of parameters optimizing this contrast function includes the true parameter.

For a given L and any $k \geq 1$, we denote $\mathbf{y}_k = y_{(k-1)L:kL-1}$ and $\mathbf{x}_k = x_{(k-1)L:kL-1}$. Assume $\{x_n\}_{n\geq 0}$ defined by (3.1) is a stationary Markov process with invariant density

$\pi(x_t|\lambda)$. This implies that once the process $\{x_t\}_{t\geq 0}$ has reached its stationary regime, then for any k, $\{\mathbf{x}_k, \mathbf{y}_k\}$ are *identically distributed* according to

$$\bar{p}(\mathbf{x}_k, \mathbf{y}_k|\lambda) = \pi(x_{(k-1)L}|\lambda)p(y_{(k-1)L}|x_{(k-1)L}, \lambda) \prod_{n=(k-1)L+1}^{kL-1} p(x_n|\bar{x}_{n-1}, \lambda)p(y_n|x_n, \lambda) \tag{3.8}$$

we can naturally introduce the likelihood in the stationary regime of the block of data $\mathbf{y}_k$

$$\bar{p}(\mathbf{y}_k|\lambda) = \int \bar{p}(\mathbf{x}_k, \mathbf{y}_k|\lambda)d\mathbf{x}_k \tag{3.9}$$

and denote $\bar{l}_\lambda(\mathbf{y}_k) \triangleq \log \bar{p}(\mathbf{y}_k|\lambda)$ the associated log-likelihood. Now, we introduce the following so-called split-data (marginal) likelihood of p blocks of L ($L \geq 1$) consecutive observations

$$\bar{p}(\mathbf{y}_1, \ldots, \mathbf{y}_p|\lambda) \triangleq \prod_{k=1}^{p} \bar{p}(\mathbf{y}_k|\lambda) \tag{3.10}$$

This split-data likelihood ignores the dependency between adjacent data blocks. The underlying motivation for the introduction of this quantity relies on the following property, which parallels the classical scenario where the observations $\mathbf{y}_k$ are independent, and the true likelihood is, therefore, a simple product of densities. As a consequence, the log-likelihood function of $\bar{p}(\mathbf{y}_1, \ldots, \mathbf{y}_p|\lambda)$ using the SDL is simply given by :

$$\bar{l}_\lambda(\mathbf{y}_1, ..., \mathbf{y}_p) = \sum_{k=1}^{p} \log \bar{p}(\mathbf{y}_k|\lambda) \tag{3.11}$$

The use of the SDL provides therefore an interesting alternative contrast function which has the clear advantage of involving integral over fixed spaces (of length L) only. As a consequence, Monte Carlo algorithms can be efficiently used in this case.

Now in order to find an estimate of the true value of parameters λ^*, we can use an on-line EM algorithm. At iteration k, the on-line version computes the average Q function by a Monte-Carlo method such as SMC methods. More precisely, once the block of observations $\mathbf{y}_k$ is available, and given our current estimate $\tilde{Q}_{k-1}$ of $\tilde{Q}$, we compute

$$\begin{aligned} \tilde{Q}_k(\lambda_{1:k-1}, \lambda) &= (1-\gamma_k)\tilde{Q}_{k-1}(\lambda_{1:k-2}, \lambda) + \gamma_k \int \log \bar{p}(\mathbf{x}_k, \mathbf{y}_k|\lambda)\bar{p}(\mathbf{x}_k|\mathbf{y}_k, \lambda_{k-1})d\mathbf{x}_k \\ &= (1-\gamma_k)\tilde{Q}_{k-1}(\lambda_{1:k-2}, \lambda) + \gamma_k \sum_{j=1}^{M} \widetilde{w}_{kL-1}^{(j)} \log \bar{p}(\mathbf{x}_k^{(j)}, \mathbf{y}_k|\lambda) \end{aligned} \tag{3.12}$$

and update the value of parameter according to

$$\lambda_k = \arg\max_{\lambda} \tilde{Q}_k(\lambda_{1:k-1}, \lambda) \tag{3.13}$$

As we work on data blocks of fixed dimension, the crucial point is that there is no more degeneracy problem. In practice, when the joint density $p(x_{0:t}, y_{0:t}|\lambda)$ belongs to the exponential family, we only need to propagate a set of sufficient statistics. More precisely, in the on-line EM context, we recursively approximate the set of sufficient statistics with the following update :

$$\begin{aligned} \Phi_k &= (1-\gamma_k)\Phi_{k-1} + \gamma_k \mathbb{E}_{\lambda_{k-1}}\left[\Psi\left(\mathbf{x}_k, \mathbf{y}_k\right)|\mathbf{y}_k\right] \\ &= (1-\gamma_k)\Phi_{k-1} + \gamma_k \sum_{j=1}^{M} \widetilde{w}_{kL-1}^{(j)} \Psi\left(\mathbf{x}_k^{(j)}, \mathbf{y}_k\right) \end{aligned} \tag{3.14}$$

where $\mathbb{E}_{\lambda_{k-1}}\left[\Psi\left(\mathbf{x}_k, \mathbf{y}_k\right)|\mathbf{y}_k\right]$ corresponds to the sufficient statistics associated to the data block $\mathbf{y}_k$. The parameter is therefore updated according to :

$$\lambda_k = \Lambda(\Phi_k) \tag{3.15}$$

where Λ is the mapping between the set of sufficient statistics and the parameter space. Let us remark the similarity between the on-line EM and the SAEM described in Section 3.1.1.2. However in the on-line EM, the parameter update is made over a new block of information which consequently allows the on-line parameter estimation unlike in the SAEM where algorithm iterations are always performed using the same block of observations (batch processing).

3.2.2 Filtering Methods

3.2.2.1 Classical Approaches

For Bayesian dynamic models, the most natural approach consists in treating the unknown parameter λ, using the state-space representation, as a component of the state which has no dynamic evolution, also referred to as a *static parameter*. Hence, we can reformulate our initial objectives as trying to simulate from the joint posterior distribution of the unobservable states and parameters. Unfortunately, this scheme is known to be inefficient since the absence of evolution for λ implies that exploration of the parameter space is limited to the first time index. Moreover, only one value of the parameter will survive after several resampling steps.

A pragmatic solution consists in running the sequential Monte Carlo filter using an artificial, hopefully negligible, dynamic equation on the parameter λ. Typically, a random-walk dynamic with a small variance is chosen [Gordon *et al.* 93, Liu *et al.* 01]. This approach can also be related to kernel estimate ideas where the target filtering and smoothing distributions are smoothed using a kernel with a small bandwidth [Musso *et al.* 01, Liu *et al.* 01]. However, such solutions do not solve the fixed-parameter estimation problem.

3.2.2.2 *MCMC moves*

One other solution to parameter estimation is to integrate Markov Chain Monte Carlo (MCMC) moves into the SMC filter [Fearnhead 01, Gilks *et al.* 01, Chopin 02] in order to maintain diversity of the samples in the parameter space. The stationary distribution for the MCMC will be the full joint posterior distribution of both states and parameters, $p(x_{0:t}, \lambda|y_{0:t})$. A natural choice of algorithm structure might be to apply Metropolis-within-Gibbs sampling steps separately to $p(\lambda|x_{0:t}, y_{0:t})$ and $p(x_{0:t}|\lambda, y_{0:t})$. However, note that in general models this will not be feasible for large datasets, since sampling from $p(\lambda|x_{0:t}, y_{0:t})$ may involve recomputing statistics based on the complete trajectory $x_{0:t}$ and $y_{0:t}$. Implementing MCMC within particle filter algorithm in this way suffers from the problem that the trajectories need to be stored. This leads to memory storage problems if long-time series are analyzed. Furthermore, the computational cost of implementing MCMC move will increase with n. Nevertheless, in many models of interest, this will not be necessary, since the influence of the path $x_{0:t}$ and $y_{0:t}$ may be summarized by low-dimensional sufficient statistics. However, another problem with this approach is that SMC estimates of sufficient statistics degrade as t increases since they are based on the approximation of the smoothing distribution $p(x_{0:t}|\lambda, y_{0:t})$.

3.2.2.3 *Adaptive Particle Filter*

In [Papavasiliou 05], 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.

3.2.2.4 *Rao-Blackwellization*

If the general model contains a linear Gaussian sub-structure, it can be exploited to derive an accurate estimator with a lower computational demand as described in details in Section 2.3.6. Using both the Bayes' theorem and the HMM, the linear gaussian variables are marginalized out and estimated by a Kalman filter, which is the optimal filter in this case. This technique is mostly used for time-varying estimation but has also been achieved for fixed parameter estimation [Liu *et al.* 98, Kong *et al.* 94].

By reviewing existing schemes for on-line parameter estimation using SMC methods, we have seen that each technique suffers from its own drawbacks. Indeed, point estimation methods provide good convergence stability but at the expense of a slow convergence rate well known in EM type algorithms. Concerning filtering methods, existing schemes suffer from high complexity (Adaptive Particle filter, MCMC) and/or degeneracy problems inherent to SMC methods (MCMC). Indeed, with the path degeneracy problem, the parameter estimates can drift away as the sufficient statistics used are not properly estimated. Finally, Rao-blackwelization is a powerful variance reduction technique for SMC methods but is restricted to the presence of linear and

Gaussian structure on the parameter of interest. As a consequence, we propose below a new approach to perform an efficient on-line parameter estimation with good convergence rate.

3.2.3 Stochastic EM per particle

The main idea of our parameter estimation scheme is to benefit from both the search diversity of filtering methods which allows to increase the convergence rate and the convergence stability of point estimation methods. In order to avoid the path degeneracy problem inherent to the SMC methods, we consider the split-data likelihood as described in the on-line EM framework, so the same notation is used. We propose to use a Stochastic Expectation Maximization (SEM) type algorithm to update the parameter value associated to each particle. Contrary to the on-line EM where the Q function is computed with a Monte Carlo approximation obtained by all particles of the SMC filter, the proposed approach consists in approximating the Q function by only one sample corresponding to the trajectory of one particle. As a consequence, each particle have its own approximation of the Q function, denoted below by $\left\{\tilde{Q}_k^{(j)}(\lambda_{1:k-1},\lambda)\right\}_{j=1}^{M}$ and defined as follows :

$$\tilde{Q}_k^{(j)}(\lambda_{1:k-1}^{(j)},\lambda) = (1-\gamma_k)\tilde{Q}_{k-1}^{(j)}(\lambda_{1:k-2},\lambda) + \gamma_k \log \bar{p}(\mathbf{x}_k^{(j)},\mathbf{y}_k|\lambda) \tag{3.16}$$

The value of each particle's parameter value is then updated according to :

$$\lambda_k^{(j)} = \arg\max_{\lambda} \tilde{Q}_k^{(j)}(\lambda_{1:k-1}^{(j)},\lambda) \tag{3.17}$$

In practice, when the joint density $p(x_{0:t},y_{0:t}|\lambda)$ is in the exponential family, we only need to propagate a set of sufficient statistics. Contrary to the on-line EM context, in the particle SEM, a set of sufficient statistics is required for each particle. As a consequence, we recursively approximate the set of sufficient statistics associated to the i-th particle with the following update :

$$\Phi_k^{(j)} = (1-\gamma_k)\Phi_{k-1}^{(j)} + \gamma_k \Psi\left(\mathbf{x}_k^{(j)},\mathbf{y}_k\right) \tag{3.18}$$

where $\Psi(\mathbf{x}_k,\mathbf{y}_k)$ corresponds to the sufficient statistics associated to the i-th particle trajectory and to the data block $\mathbf{y}_k$. The parameter associated to the j-th particle is therefore updated according to :

$$\lambda_k^{(j)} = \Lambda(\Phi_k^{(j)}) \tag{3.19}$$

where Λ is the mapping between the set of sufficient statistics and the parameter space.

Since the joint posterior distribution of both the state and the parameters is ap-

proximated as follows :

$$p(\mathbf{x}_{1:k}, \lambda|\mathbf{y}_{1:k}) \approx \sum_{j=1}^{M} \widetilde{w}_{kL-1}^{(j)} \delta(\mathbf{x}_{1:k} - \mathbf{x}_{1:k}^{(j)}; \lambda - \lambda_k^{(j)}) \tag{3.20}$$

with $\delta(.;.)$ the two-dimensional Dirac delta function, the estimate of the parameter is then given by the importance weighted sum of each particle parameter value, *i.e.*

$$\lambda_k = \sum_{j=1}^{M} \widetilde{w}_{(k-1)L}^{(j)} \lambda_k^{(j)} \tag{3.21}$$

3.3 APPLICATION

In this section, performance of the proposed parameter estimation scheme is compared to the on-line EM through numerical simulations. We have applied these algorithms to two different dynamic state-space model described below : a linear and Gaussian model and a stochastic volatility model.

3.3.1 Dynamic State-Space Models

3.3.1.1 Model 1 : Linear and Gaussian model

Let us consider a noisily observed Gaussian autoregressive model of order 1, *i.e.*

$$\begin{cases} x_t = \phi x_{t-1} + v_t \\ y_t = x_t + b_t \end{cases} \tag{3.22}$$

where $v_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_v^2)$, $b_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_b^2)$ are two mutually independent sequences, independent of the initial state x_0 and $\lambda = (\phi, \sigma_v^2, \sigma_b^2)$ denotes the static parameter vector. It can easily be checked that $\pi(x_t|\lambda) = \mathcal{N}(0, \frac{\sigma_v^2}{1-\phi^2})$, $p(x_t|x_{t-1}, \lambda) = \mathcal{N}(x_t; x_{t-1}, \sigma_v^2)$ and $p(y_t|x_t, \lambda) = \mathcal{N}(y_t; x_t, \sigma_b^2)$.

3.3.1.2 Model 2 : Stochastic volatility model

Stochastic volatility models are used in econometrics to model the changing volatility of asset returns. A simple univariate stochastic model is

$$\begin{cases} x_t = \phi x_{t-1} + v_t \\ y_t = \exp(x_t/2)\, b_t \end{cases} \tag{3.23}$$

where $v_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_v^2)$, $b_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_b^2)$ are two mutually independent sequences, independent of the initial state x_0 and $\lambda = (\phi, \sigma_v^2, \sigma_b^2)$ denotes the static parameter vector. The transition probability $p(x_t|x_{t-1}, \lambda)$ and the stationary distribution $\pi(x_t|\lambda)$ are identical to those of the previous model but here $p(y_t|x_t, \lambda) = \mathcal{N}(y_t; 0, \sigma_b^2 \exp(x_t))$.

3.3.2 Derivation of the algorithms

In this section, we are interested in estimating the unknown static parameter vector λ. In the two described models, the stationary distribution of the hidden process is $\mathcal{N}(0, \frac{\sigma_v^2}{1-\phi^2})$. The complete log-likelihood function of the k-th data block associated to each particle, *i.e.* $\bar{p}(\mathbf{x}_k^{(j)}, \mathbf{y}_k|\lambda)$, required in (3.12) and (3.16) respectively for the on-line EM and the Particle SEM can therefore be written as :

$$\begin{aligned}\log \bar{p}(\mathbf{x}_k^{(j)}, \mathbf{y}_k|\lambda) &= cste + \frac{1}{2}\log(1-\phi^2) - \frac{L}{2}\log\sigma_v^2 - \frac{L}{2}\log\sigma_b^2 - \frac{1-\phi^2}{2\sigma_v^2}x_{(k-1)L}^{(j)2} \\ &\quad - \frac{1}{2\sigma_v^2}\sum_{n=1}^{L-1}(x_{(k-1)L+n}^{(j)} - \phi x_{(k-1)L+n-1}^{(j)})^2 - \chi_k^{(j)} \end{aligned} \tag{3.24}$$

with

$$\chi_k^{(j)} = \begin{cases} \frac{1}{2\sigma_b^2}\sum_{n=0}^{L-1}(y_{(k-1)L+n} - x_{(k-1)L+n}^{(j)})^2 & \text{Model 1} \\ \frac{1}{2\sigma_b^2}\sum_{n=0}^{L-1} y_{(k-1)L+n}^2 \exp(-x_{(k-1)L+n}^{(j)}) & \text{Model 2} \end{cases} \tag{3.25}$$

Since the complete likelihood function is in the exponential family, we only need to propagate the set of sufficient statistics for the derivation of both the on-line EM and the particle SEM. The quantity $\Psi\left(\mathbf{x}_k^{(j)}, \mathbf{y}_k\right) = \left[\Psi_1^{(j)}, \Psi_2^{(j)}, \Psi_3^{(j)}, \Psi_4^{(j)}, \Psi_5^{(j)}\right]$ involved in evaluating the set of sufficient statistics as described in Eqs. (3.14) and (3.18) respectively for the on-line EM and the particle SEM can be written using (3.25) as :

$$\Psi_1^{(j)} = x_{(k-1)L}^{(j)2} \tag{3.26}$$

$$\Psi_2^{(j)} = \sum_{n=1}^{L-1} x_{(k-1)L+n}^{(j)2} \tag{3.27}$$

$$\Psi_3^{(j)} = \sum_{n=1}^{L-1} x_{(k-1)L+n}^{(i)} x_{(k-1)L+n-1}^{(j)} \tag{3.28}$$

$$\Psi_4^{(j)} = \sum_{n=1}^{L-1} x_{(k-1)L+n-1}^{(j)2} \tag{3.29}$$

$$\Psi_5^{(j)} = \begin{cases} \sum_{n=0}^{L-1}(y_{(k-1)L+n} - x_{(k-1)L+n}^{(j)})^2 & \text{Model 1} \\ \sum_{n=0}^{L-1} y_{(k-1)L+i}^2 \exp(-x_{(k-1)L+n}^{(j)}) & \text{Model 2} \end{cases} \tag{3.30}$$

3.3.2.1 The on-line EM - SMC

Let us recall that in the on-line EM approach the set of sufficient statistics $\Phi_k = [\Phi_{k,1}\Phi_{k,2}, \Phi_{k,3}, \Phi_{k,4}, \Phi_{k,5}]$ is updated according to Eq. (3.14) using (3.26)-(3.30). The mapping between the set of sufficient statistics and the parameter space introduced in (3.15) is defined for both the models 1 and 2 as follows :

$$\sigma_{b,k}^2 = \frac{1}{L}\Phi_{k,5} \tag{3.31}$$

$$\sigma_{v,k}^2 = \frac{1}{L}\left(\left(1-\phi_k^2\right)\Phi_{k,1} + \phi_k^2\Phi_{k,4} - 2\phi_k\Phi_{k,3} + \Phi_{k,1}\right) \tag{3.32}$$

$$-\frac{\phi_k}{1-\phi_k^2} + \frac{\phi_k(\Phi_{k,1}-\Phi_{k,4})}{\sigma_{v,k}^2} + \frac{\Phi_{k,3}}{\sigma_{v,k}^2} = 0 \tag{3.33}$$

Plugging the expression of $\sigma_{v,k}^2$ in the equation (3.33) gives an equation in ϕ that can be solved by simply computing the eigenvalues of the companion matrix associated to the polynomial. The on-line EM using SMC method is summed up in Table 3.1.

Initialization
$\left\{x_0^{(j)}; \lambda_0^{(j)}; \widetilde{w}_0^{(j)}\right\}_{j=1}^{M} = \{0; \lambda_{Init}; 1/M\}$
For $k = 1...$
  For $n = 0...L-1$
    For $j = 1...M$
      Sample $x_{(k-1)L+n}^{(i)}$ from the prior distribution $p(x_{(k-1)L+n}|x_{(k-1)L+n-1}^{(j)}, \lambda_{k-1})$
      Evaluate the corresponding importance : $w_{(k-1)L+n}^{(j)} = w_{(k-1)L+n-1}^{(j)} p(y_{(k-1)L+n}|x_{(k-1)L+n}^{(j)}, \lambda_{k-1})$
    end for
    Normalize importance weights : $\widetilde{w}_{(k-1)L+n}^{(j)} = w_{(k-1)L+n}^{(j)} / \sum_{m=1}^{M} w_{(k-1)L+n}^{(m)}$
    If $N_{eff} < N_{seuil}$ **then**
      Resample particle trajectories using systematic resampling
    end if
    Evaluate the state estimate : $\widehat{x}_{(k-1)L+n} = \sum_{j=1}^{M} \widetilde{w}_{(k-1)L+n}^{(j)} x_{(k-1)L+n}^{(j)}$
  end for
  Compute the set of sufficient statistics update Φ_k using (3.14) and (3.26)-(3.30)
  Compute the parameter update λ_k using (3.31)-(3.33)
end for

Table 3.1 – *On-line EM using Sequential Monte Carlo algorithm (OEM-SMC)*

3.3.2.2 *The Particle SEM - SMC*

Contrary to the on-line EM framework, a set of sufficient statistics for each particle is required in the proposed approach. As a consequence, let us denote $\Phi_k^{(j)} = \left[\Phi_{k,1}^{(j)}\Phi_{k,2}^{(j)}, \Phi_{k,3}^{(j)}, \Phi_{k,4}^{(j)}, \Phi_{k,5}^{(j)}\right]$ the sufficient statistics associated to the j-th particle which is updated according to Eq. (3.18) using (3.26)-(3.30). The mapping between the set of sufficient statistics and the parameter space introduced in (3.15) is defined for both the models 1 and 2 as follows :

$$\sigma_{b,k}^{(j)2} = \frac{1}{L}\Phi_{k,5}^{(j)} \tag{3.34}$$

$$\sigma_{v,k}^{(j)2} = \frac{1}{L}\left(\left(1-\phi_k^{(j)2}\right)\Phi_{k,1}^{(j)} + \phi_k^{(ij)2}\Phi_{k,4}^{(j)} - 2\phi_k^{(j)}\Phi_{k,3}^{(j)} + \Phi_{k,1}^{(j)}\right) \tag{3.35}$$

$$-\frac{\phi_k^{(j)}}{1-\phi_k^{(j)2}} + \frac{\phi_k^{(j)}(\Phi_{k,1}^{(j)}-\Phi_{k,4}^{(j)})}{\sigma_{v,k}^{(j)2}} + \frac{\Phi_{k,3}^{(j)}}{\sigma_{v,k}^{(j)2}} = 0 \tag{3.36}$$

Plugging the expression of $\sigma_{v,k}^{(j)2}$ in the equation (3.36) gives an equation in ϕ that can be solved like in the on-line EM context. The particle SEM using SMC method is summed up in Table 3.2.

Initialization

$\left\{x_0^{(j)}; \lambda_0^{(j)}; \tilde{w}_0^{(j)}\right\}_{j=1}^{M} = \{0; \lambda_{Init}; 1/M\}$

For $k = 1...$

  For $n = 0...L-1$

    For $j = 1...M$

      Sample $x_{(k-1)L+n}^{(j)}$ from the prior distribution $p(x_{(k-1)L+n}|x_{(k-1)L+n-1}^{(j)}, \lambda_{k-1}^{(j)})$

      Evaluate the corresponding importance : $w_{(k-1)L+n}^{(j)} = w_{(k-1)L+n-1}^{(j)} p(y_{(k-1)L+n}|x_{(k-1)L+n}^{(j)}, \lambda_{k-1}^{(j)})$

    end for

    Normalize importance weights : $\tilde{w}_{(k-1)L+n}^{(j)} = w_{(k-1)L+n}^{(j)} / \sum_{m=1}^{M} w_{(k-1)L+n}^{(m)}$

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

      Resample particle trajectories using systematic resampling

    end if

    Evaluate the state estimate : $\hat{x}_{(k-1)L+n} = \sum_{j=1}^{M} \tilde{w}_{(k-1)L+n}^{(j)} x_{(k-1)L+n}^{(j)}$

    Evaluate the parameter estimate : $\hat{\lambda}_{(k-1)L+n} = \sum_{j=1}^{M} \tilde{w}_{(k-1)L+n}^{(j)} \lambda_k^{(j)}$

  end for

  Compute the set of sufficient statistics update $\left\{\Phi_k^{(j)}\right\}_{j=1}^{M}$ using (3.18) and (3.26)-(3.30)

  Compute the parameter update $\left\{\lambda_k^{(j)}\right\}_{j=1}^{M}$ using (3.34)-(3.36)

end for

Table 3.2 – *Particle SEM using Sequential Monte Carlo algorithm (PSEM-SMC)*

3.3.3 Numerical Simulations

Throughout this section, performance of the proposed on-line parameter strategy is assessed and compared to the on-line EM. In particular, the bias, the variance and the mean square error (MSE) of the estimate are analyzed. Firstly, influence of both the forgetting factor and the block length is studied for both algorithms. Then, we illustrate the sensitivity of the algorithms to initial parameter values. Finally, the case where the hidden Markov process is non-stationary is studied. This last case must be considered since the phase noise process is a non-stationary process (Eq. (1.22)).

3.3.3.1 Influence of both the forgetting factor and the block length

In this section, let us consider $\phi_{Init} = 0.2$, $\sigma_{v,Init}^2 = 0.75$ and $\sigma_{b,Init}^2 = 1.5$ the initial values of parameters in the algorithms.

The evolution over block number of the forgetting factor $\gamma_k = k^{-\alpha}$ involved in both algorithms is depicted in Fig. 3.3. For $\alpha = 1$, the pseudo complete likelihood function $\tilde{p}(\mathbf{x}_k, \mathbf{y}_k|\lambda)$ corresponding to the k-th data block is less taking into account rapidly over time for the computation of the average Q function (in Eq. (3.12) or (3.16) depending on the algorithm) than for $\alpha = 0.5$. Indeed for example, γ_{2500} is equal to 2.10^{-2} and 3.10^{-4} respectively for $\alpha = 0.5$ and $\alpha = 1$. The influence of α in algorithm performances in model 1 can be seen in Figs. 3.4 and 3.5 respectively for the OEM-SMC and the PSEM-SMC with $L = 10$ and $N = 100$ particles. In these two figures, the median and the 2.5th, 97.5th, 25th and 75th percentiles of the parameter estimates are plotted for α equal to 0.5 and 0.75. It can be logically noted that for $\alpha = 0.75$ the parameter estimates of both algorithms converge more slowly than for $\alpha = 0.5$ since the complete likelihood functions for $\alpha = 0.75$ associated to the blocks up to $k-1$ are more significant in the average of the Q function than for the case of $\alpha = 0.5$.

By comparing Figs. 3.4 and 3.5 where percentiles of parameter estimates are depicted for $L = 10$, $\alpha = 0.5$ and $N = 100$, PSEM-SMC converges more rapidly, whatever α, to a value $\widehat{\lambda}$ in the neighborhood of the true parameter value than the OEM-SMC. This remark is also valid for the model 2 as shown in Fig. 3.6. The diversity in the search space of parameter in the PSEM-SMC greatly improves the convergence rate but at the expense of a slight increase of the estimate variance.

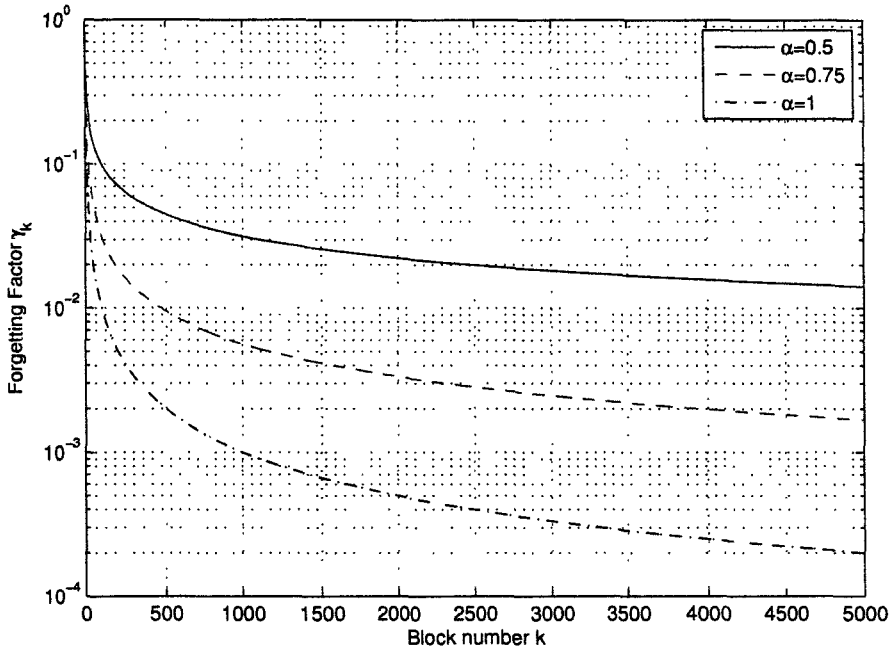

Figure 3.3 – *Evolution of the forgetting factor $\gamma_k = k^{-\alpha}$ vs block number k for 3 different values of α.*

Now, we study the influence of the block length L. Algorithm performances with $N = 100$, $\alpha = 0.5$ and L equal to 5 and 10 are depicted in Figs. 3.7 in term of mean square error (MSE) for the OEM-SMC and PSEM-SMC in model 1 where the true model parameters are $\phi = 0.8$, $\sigma_v^2 = 0.1$ and $\sigma_b^2 = 0.1$. It can be seen that the convergence speed for both algorithms increases with the block length L since the parameter update is based on a longer observation data block. The PSEM-SMC converges more rapidly than the OEM-SMC. Moreover, it can be also remarked that for $L = 5$ and $k = 2000$, mean square errors of the parameter estimates converge to the same values for both algorithms, *i.e.* 2×10^{-3}, 8×10^{-4} and 7×10^{-4} respectively for ϕ, σ_b^2 and σ_v^2. The MSE of parameter estimates of both algorithm in the same context but with $M = 200$ particules have been depicted in Fig 3.8 and 3.9. From Fig. 3.9, it can be seen that the PSEM-SMC converges more rapidly in the first blocks that the OEM-SMC. However from Fig. 3.8, contrary to the previous case with 100 particules, MSE of parameter estimates using the OEM-SMC are better than those of the PSEM-SMC when the block number k is high. In Table 3.3, the bias and the variance of the parameter estimate errors obtained in $k = 4000$ have been presented. It can be denoted that the bias

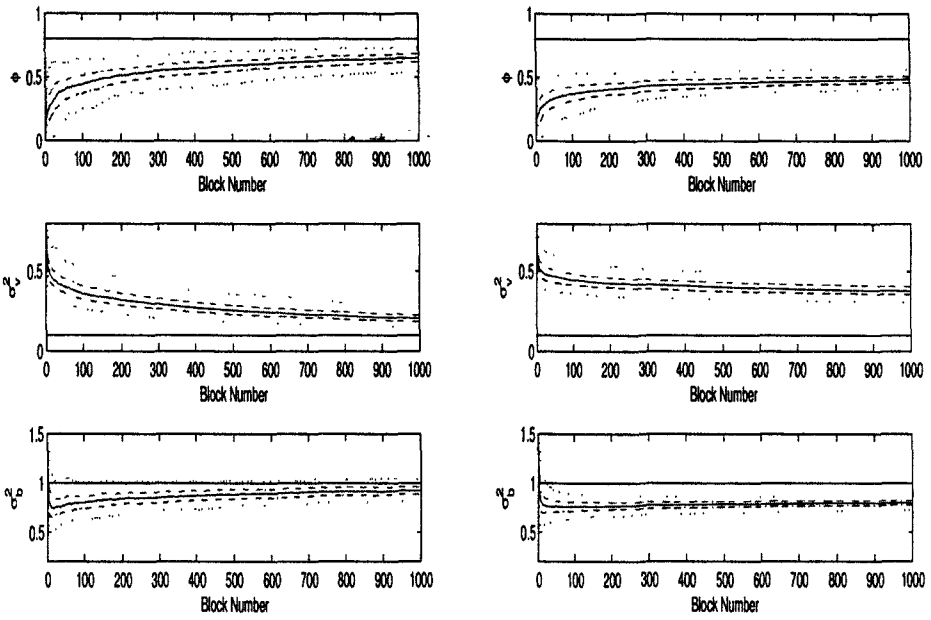

Figure 3.4 – *Median (solid lines), 2.5th and 97.5th percentiles (dotted lines), 25th and 75th percentiles (dash-dotted lines) of the parameter estimates in Model 1 using the OEM-SMC with* $M = 100$ *particles,* $L = 10$ *and* $\alpha = 0.5$ *(left) and* $\alpha = 0.75$ *(right) (*$\phi = 0.8$, $\sigma_v^2 = 0.1$, $\sigma_b^2 = 1$*).*

obtained with the PSEM-SMC is slightly better than one obtained with the OEM-SMC. However, as also remarked previously, a smaller variance is achieved by the OEM-SMC explaining the fact that the MSE curve for the OEM-SMC is slightly better than for the PSEM-SMC.

	Bias		Variance	
	OEM-SMC	PSEM-SMC	OEM-SMC	PSEM-SMC
ϕ	-0.0221	-0.0196	4.67×10^{-4}	9.67×10^{-4}
σ_v^2	0.0159	0.0156	3.34×10^{-5}	1.98×10^{-4}
σ_b^2	-0.0139	-0.01327	5.75×10^{-5}	1.7×10^{-4}

Table 3.3 – *Parameter estimate error bias and variance for* $L = 5$, $N = 200$ *particles and* $\alpha = 0.5$ *(*$\phi = 0.8$, $\sigma_v^2 = 0.1$, $\sigma_b^2 = 0.1$*)*

3.3.3.2 *Influence of the initial values*

In this section, the influence of the initial parameter values is studied where the true model parameters are $\phi = 0.8$, $\sigma_v^2 = 0.2$ and $\sigma_b^2 = 1$. Fig. 3.10 illustrates algorithm performance in model 2 with two different sets of initial values. It can be seen from these two figures that the PSEM-SMC still converges more rapidly to a value in the neighborhood of the true parameter value, especially for the state transition ϕ and the state driving noise σ_v^2.

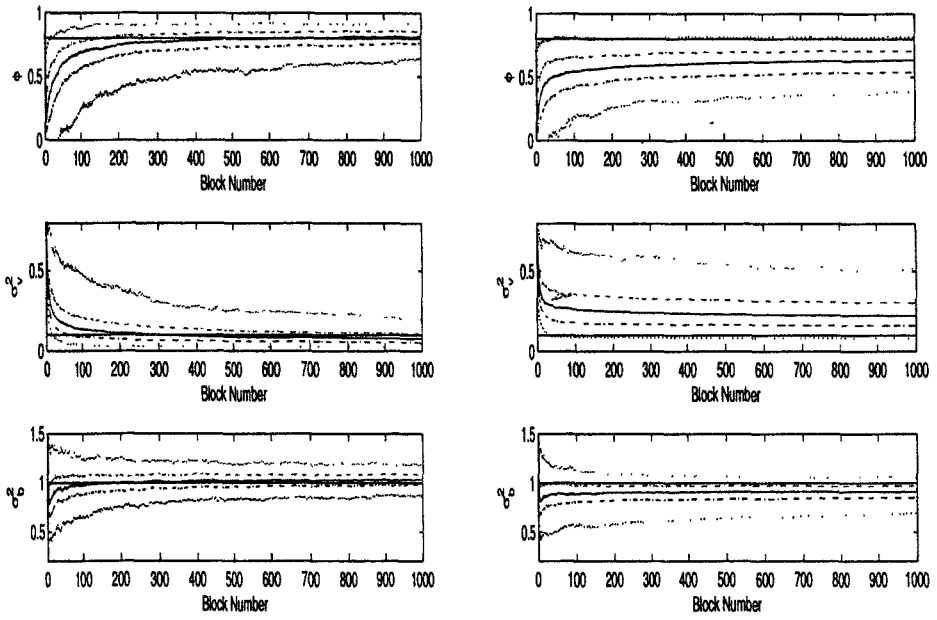

Figure 3.5 – *Median (solid lines), 2.5th and 97.5th percentiles (dotted lines), 25th and 75th percentiles (dash-dotted lines) of the parameter estimates in Model 1 using the PSEM-SMC with* $M = 100$ *particles,* $L = 10$ *and* $\alpha = 0.5$ *(left) and* $\alpha = 0.75$ *(right)* ($\phi = 0.8$, $\sigma_v^2 = 0.1$, $\sigma_b^2 = 1$).

3.3.3.3 *Non-stationary hidden Markov process*

Let us recall that since the phase noise process is a non-stationary hidden Markov (Eq. (1.22)), this case must be considered. However, both the OEM-SMC and the PSEM-SMC are based on the stationary property of the hidden Markov process in Eq. (3.8). In order to allow the derivation of these algorithms in this case, we propose to use the prior transition density $p(x_{(k-1)L}|x_{(k-1)L-1}, \lambda)$ instead of the stationary distribution $\pi(x_{(k-1)L}|\lambda)$. As a consequence, Eq. (3.8) becomes :

$$\tilde{p}(\mathbf{x}_k, \mathbf{y}_k|\lambda) \approx \prod_{n=(k-1)L}^{kL-1} p(x_n|x_{n-1}, \lambda) p(y_n|x_n, \lambda) \tag{3.37}$$

The quantity $\Psi\left(\mathbf{x}_k^{(j)}, \mathbf{y}_k\right) = \left[\Psi_1^{(j)}, \Psi_2^{(j)}, \Psi_3^{(j)}, \Psi_4^{(j)}\right]$ is therefore given by :

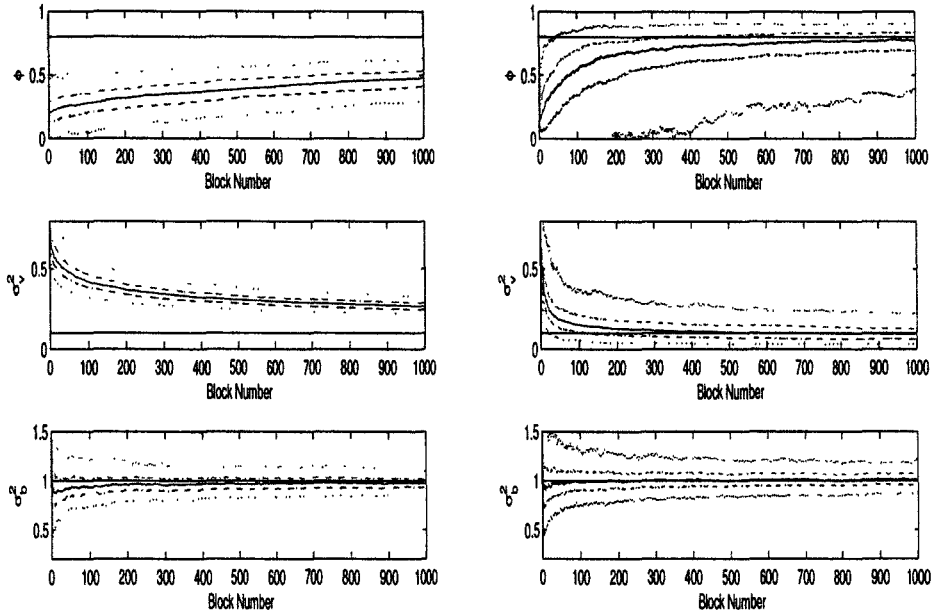

Figure 3.6 – *Median (solid lines), 2.5th and 97.5th percentiles (dotted lines), 25th and 75th percentiles (dash-dotted lines) of the parameter estimates in Model 2 using OEM-SMC (left) and PSEM-SMC (right) with $M = 100$ particles, $L = 10$ and $\alpha = 0.5$ ($\phi = 0.8$, $\sigma_v^2 = 0.1$, $\sigma_b^2 = 1$).*

$$\Psi_1^{(j)} = \sum_{n=0}^{L-1} x_{(k-1)L+n}^{(j)2} \tag{3.38}$$

$$\Psi_2^{(j)} = \sum_{n=0}^{L-1} x_{(k-1)L+n}^{(i)} x_{(k-1)L+n-1}^{(j)} \tag{3.39}$$

$$\Psi_3^{(j)} = \sum_{n=0}^{L-1} x_{(k-1)L+n-1}^{(j)2} \tag{3.40}$$

$$\Psi_4^{(j)} = \begin{cases} \sum_{n=0}^{L-1} (y_{(k-1)L+n} - x_{(k-1)L+n}^{(j)})^2 & \text{Model 1} \\ \sum_{n=0}^{L-1} y_{(k-1)L+i}^2 \exp(-x_{(k-1)L+n}^{(j)}) & \text{Model 2} \end{cases} \tag{3.41}$$

The sufficient statistics associated to the j-th particle in the case of the PSEM-SMC, $\Phi_k^{(j)}$, is always updated according to Eq. (3.18). The mapping between the set of sufficient statistics and the parameter space introduced in (3.15) is defined for both the models 1 and 2 as follows :

$$\sigma_{b,k}^{(j)2} = \frac{1}{L} \Phi_{k,5}^{(j)} \tag{3.42}$$

$$\phi_k^{(j)} = \frac{\Phi_{k,2}^{(j)}}{\Phi_{k,3}^{(j)}} \tag{3.43}$$

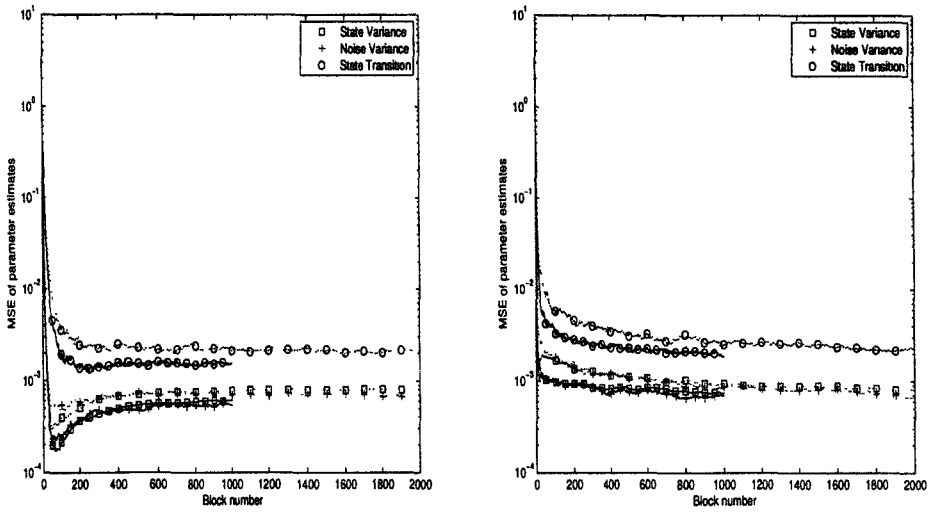

Figure 3.7 – *MSE of parameter estimates in Model 1 using OEM-SMC (left) and PSEM-SMC (right) for two different block lengths, $L = 5$ (dotted lines) and $L = 10$ (solid lines) and with $M = 100$ particles and $\alpha = 0.5$ ($\phi = 0.8$, $\sigma_v^2 = 0.1$, $\sigma_b^2 = 0.1$).*

$$\sigma_{v,k}^{(j)2} = \frac{1}{L}\left(\Phi_{k,1}^{(j)} - 2\phi_k^{(j)}\Phi_{k,2}^{(j)} + \phi_k^{(j)2}\Phi_{k,3}^{(j)}\right) \tag{3.44}$$

The MSE of parameter estimates using OEM-SMC and PSEM-SMC in model 2 is depicted in Fig. 3.11. As previously remarked in the stationary case, the PSEM-SMC converges still more rapidly than the OEM-SMC. It can be denoted that the MSE of σ_b^2 for both algorithms is high. This is due to the non-stationary nature of the hidden Markov process with variance grows with time and leads to a rapid fluctuation of the observation as illustrated in Fig. 3.12. In Fig. 3.13, we focus on the median and the 2.5th, 97.5th, 25th and 75th percentiles of the parameter estimates. This figure shows that the median of the estimate of σ_b^2 is close to the true value but the estimates presents a high dispersion which explains the MSE results. Finally, the estimates of ϕ and σ_v^2 are still more accurate, in the first blocks, for PSEM-SMC than for OEM-SMC in this non-stationary context.

CONCLUSION

In this chapter, we have briefly reviewed the existing scheme for both off-line and on-line parameter estimation using sequential Monte-Carlo methods. We have focused more specifically on the on-line mode where the SMC methods are particularly well adapted due to their sequential nature. In this on-line context, two different strategies, *i.e.* point estimation and filtering methods, can be adopted. On the one hand, point estimation methods like on-line EM using SMC have a good convergence stability but generally suffer from a slow convergence rate. On the other hand, the filtering methods induce diversity in the search space of parameters leading typically to an increase

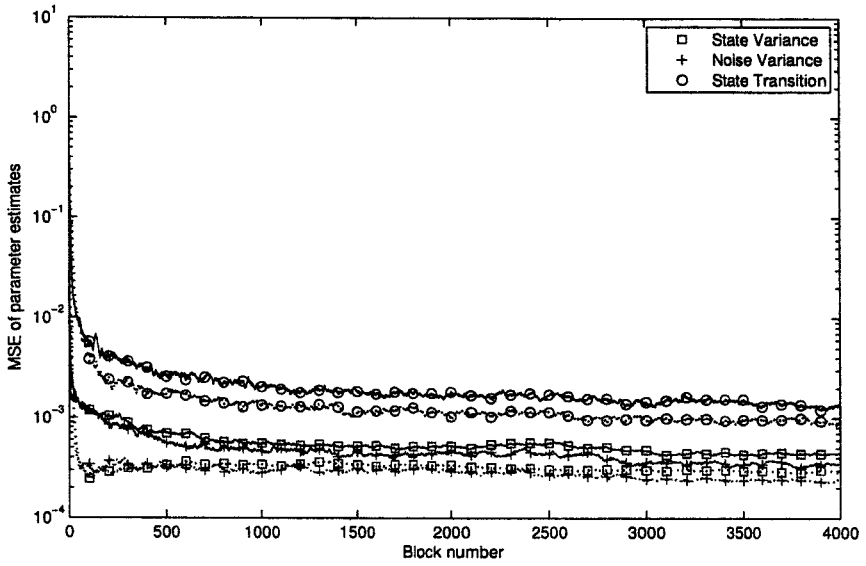

Figure 3.8 – *MSE of parameter estimates in Model 1 using OEM-SMC (dotted lines) and PSEM-SMC (solid lines) for* $L = 5$, $M = 200$ *particles,* $\alpha = 0.5$ *(*$\phi = 0.8$, $\sigma_v^2 = 0.1$, $\sigma_b^2 = 0.1$*).*

of the convergence rate in comparison to point estimation methods. However, principal existing schemes in filtering methods suffer from the path degeneracy problem and the parameter estimates can drift away as the sufficient statistics used are not properly estimated. In order to avoid this disadvantage, we have proposed to use a Stochastic EM per particle (PSEM-SMC).

Numerical simulations show the efficiency of the proposed scheme in comparison to the on-line EM (OEM-SMC). We have seen that the PSEM-SMC outperforms the OEM-SMC in term of convergence rate due to the diversity in the search space of parameters. Moreover, this proposed scheme offer good stability in time due to the use of a forgetting factor. The case of a non-stationary hidden Markov process has also been treated. Indeed, since the phase noise is a non-stationary process, this case must be studied. In this context, we have proposed an adaptation of both OEM-SMC and PSEM-SMC. Numerical results illustrates that the PSEM-SMC still converges more rapidly than the OEM-SMC.

As a consequence, the proposed scheme could be apparently used for an efficient joint parameter and state estimation involved in the phase distortions compensation problem on multicarrier 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 SMC filters.

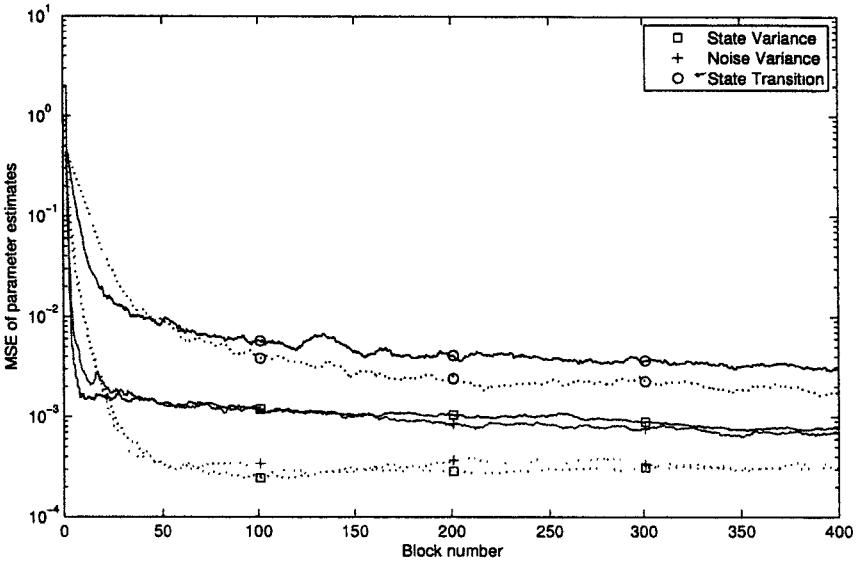

Figure 3.9 – *MSE of parameter estimates in Model 1 using OEM-SMC (dotted lines) and PSEM-SMC (solid lines) for* $L = 5$, $M = 200$ *particles,* $\alpha = 0.5$ *(*$\phi = 0.8$, $\sigma_v^2 = 0.1$, $\sigma_b^2 = 0.1$*)*

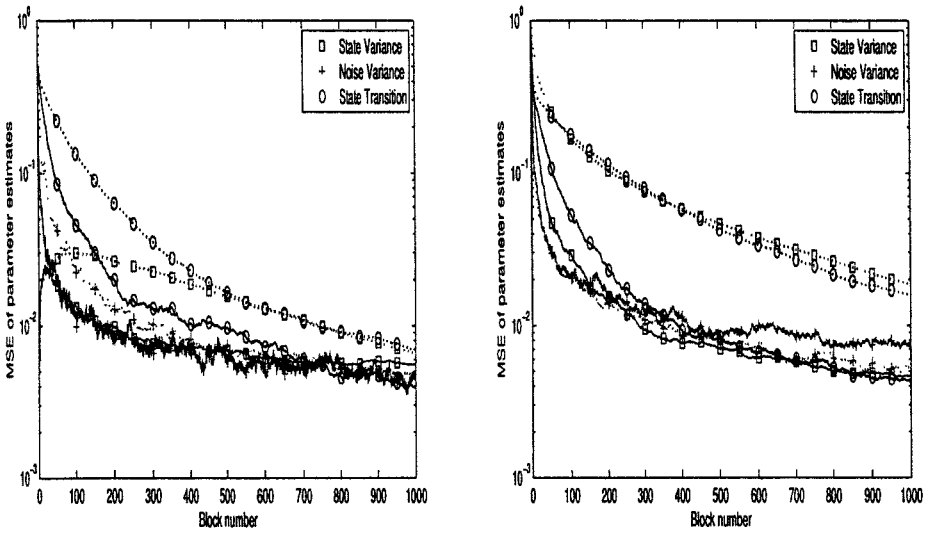

Figure 3.10 – *MSE of parameter estimates in Model 2 using OEM-SMC (dotted lines) and PSEM-SMC (solid lines) with* $\left\{\phi_{Init}, \sigma^2_{v,Init}, \sigma^2_{b,Init}\right\} = \{0.1, 0.3, 1.5\}$ *(left) and* $\left\{\phi_{Init}, \sigma^2_{v,Init}, \sigma^2_{b,Init}\right\} = \{-0.2, 0.6, 0.1\}$ *(right) and for* $L = 10$, $M = 100$ *particles and* $\alpha = 0.5$ *(*$\phi = 0.8$, $\sigma_v^2 = 0.2$, $\sigma_b^2 = 1$*).*

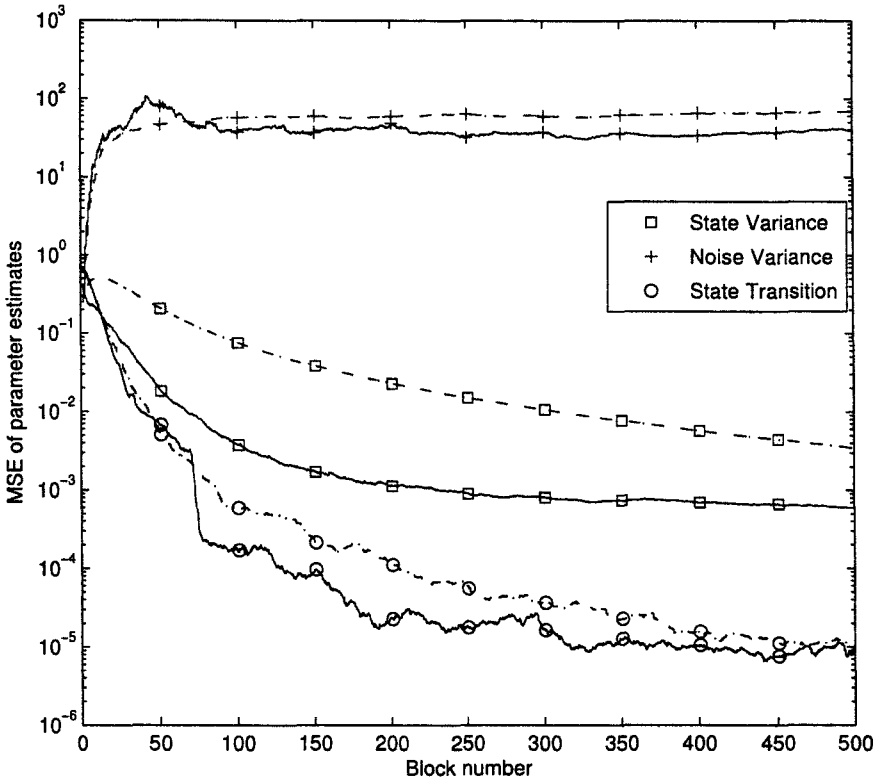

Figure 3.11 – *MSE of parameter estimates in Model 2 using OEM-SMC (dashed lines) and PSEM-SMC (solid lines) for $L = 10$, $M = 100$ particles and $\alpha = 0.5$ ($\phi = 1$, $\sigma_v^2 = 0.1$, $\sigma_b^2 = 1$).*

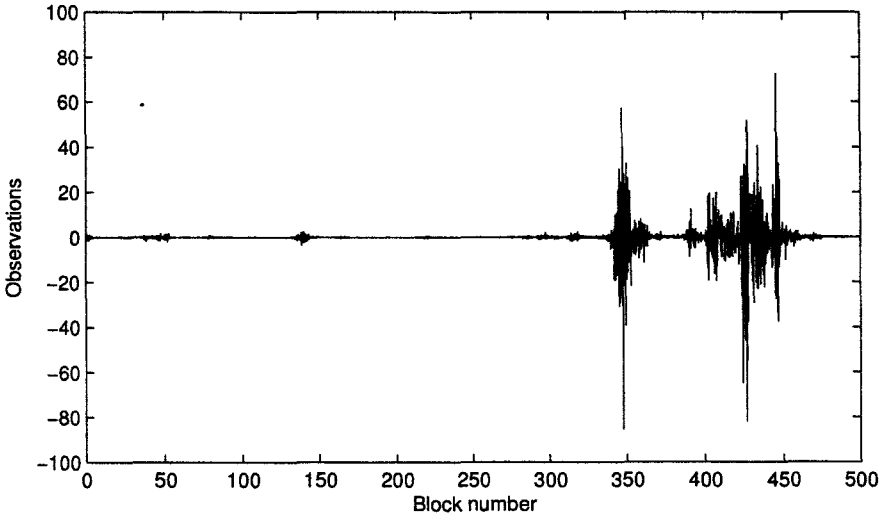

Figure 3.12 – *Example of an observation sequence using model 2 ($\phi = 1$, $\sigma_v^2 = 0.1$, $\sigma_b^2 = 1$).*

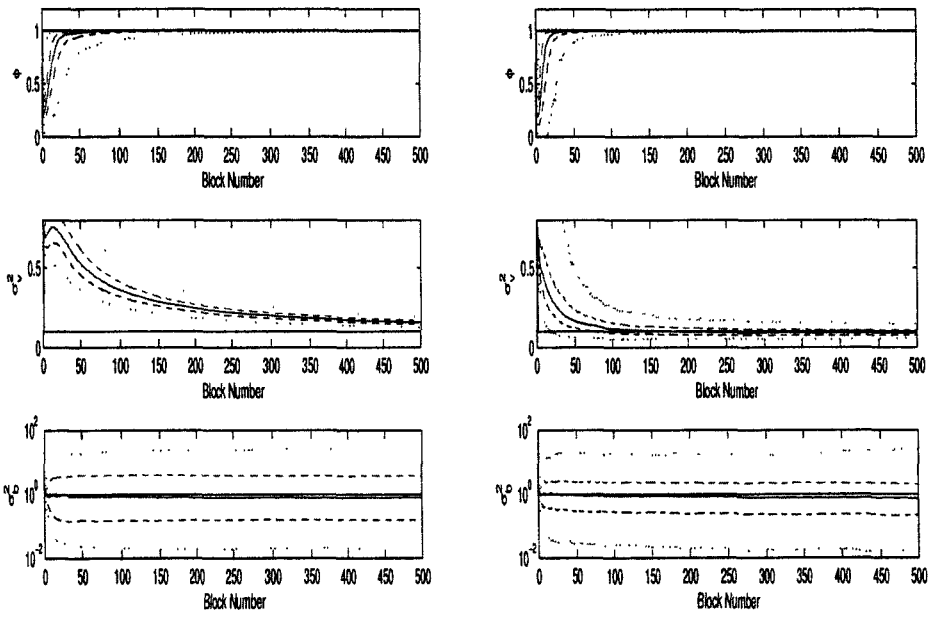

Figure 3.13 – *Median (solid lines), 2.5th and 97.5th percentiles (dotted lines), 25th and 75th percentiles (dash-dotted lines) of the parameter estimates in Model 2 using OEM-SMC (left) and PSEM-SMC (right) with* $M = 100$ *particles,* $L = 10$ *and* $\alpha = 0.5$ ($\phi = 1$, $\sigma_v^2 = 0.1$, $\sigma_b^2 = 1$).

4 CHANNEL ESTIMATION IN THE PRESENCE OF PHN AND CFO

IN this chapter, we focus on the first step of the proposed multicarrier receiver in the presence of both PHN and CFO which consists in estimating the channel impulse response from a training sequence. As described in Section 1.3, only two papers [Wu *et al.* 03b] and [Lin *et al.* 06] deal with channel estimation in the presence of both PHN and CFO. In [Wu *et al.* 03b], the authors mainly focus on channel and CFO estimation and in [Lin *et al.* 06] an approximate 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 here 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 following multicarrier symbols. To estimate these many unknowns from a single training multicarrier symbol, we propose an off-line and an on-line estimator based on Monte Carlo (MC) methods. Firstly, we derive 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 approximate maximum a posteriori estimator proposed in [Lin *et al.* 06] 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 as shown in Chapter 3, 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 on point estimation and filtering method, namely the on-line EM (OEM-SMC) and the stochastic EM per particle (PSEM-SMC) respectively, are derived for the considered problem.

This chapter is organized as follows. In Section 4.1, models of both the received multicarrier signal and the phase distortion are introduced, leading to the dynamic state space (DSS) representation. The estimation strategies based on Monte Carlo

methods is then presented. In Sections 4.2 and 4.3, the different approaches for the estimation problem are detailed. Section 4.5 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 4.6.

4.1 System Model

4.1.1 Signal Model

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

$$r_t = e^{j(\theta_t + 2\pi t\epsilon/N)} \sum_{l=0}^{L_p-1} h_l s_{t-l} + b_t \tag{4.1}$$

where t denotes the t-th sample of the multicarrier symbol, $\{h_l\}_{l=0}^{L_p-1}$, $\{s_t\}_{t=0}^{N-1}$, $\{\theta_t\}_{t=0}^{N-1}$ and $\{b_t\}_{t=0}^{N-1}$ 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 (4.1) can be written as :

$$r_t = e^{j(\theta_t + 2\pi t\epsilon/N)} \mathbf{s}_t^T \mathbf{h} + b_t \tag{4.2}$$

where $\mathbf{s}_t = \begin{bmatrix} s_t & \cdots & s_{t-L_p+1} \end{bmatrix}^T$ is the transmitted multicarrier signal vector and $\mathbf{h} = \begin{bmatrix} h_0 & \cdots & h_{L_p-1} \end{bmatrix}^T$ is the CIR vector. Let us remark that for simplicity the indice n used in Chapter (1) in order to denote the current multicarrier symbol has been omitted since in this chapter only one multicarrier symbol is required.

4.1.2 Phase Distortion Model

In Section 1.2.1, the phase distortions, *i.e.* the CFO and the PHN, have been introduced. Using (1.22), (1.23) and (4.2), a discrete recursive relation for phase distortions including both PHN and CFO can be summarized as :

$$\phi_t = \begin{cases} v_0 & \text{if } t = 0 \\ \phi_{t-1} + \frac{2\pi\epsilon}{N} + v_t & \text{otherwise} \end{cases} \tag{4.3}$$

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

4.1.3 Dynamic State-Space Model

This chapter focuses on the accurate estimation of CIR, CFO and PHN from a single multicarrier symbol. At the receiver, the transmitted multicarrier signal $\{s_t\}_{t=0}^{N-1}$ 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 (4.2) and (4.3), we obtain the following dynamic state-space model :

$$\begin{cases} \phi_t = \begin{cases} v_0 & \text{if } t = 0 \\ \phi_{t-1} + \frac{2\pi\epsilon}{N} + v_t & \text{otherwise} \end{cases} \\ r_t = e^{j\phi_t}\mathbf{s}_t^T\mathbf{h} + b_t \end{cases} \tag{4.4}$$

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 *p.d.f.* $p(\mathbf{h}, \phi_{0:N-1}, \epsilon, \sigma^2 | r_{0:N-1})$. Unfortunately, this distribution is analytically intractable, so we propose its numerical approximation via Monte Carlo methods.

Let us remark 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 :

$$\widehat{\phi}_{0:t} \rightarrow \phi_{0:t} + \Phi \tag{4.5}$$

$$\widehat{\mathbf{h}} \rightarrow e^{-j\Phi}\mathbf{h} \tag{4.6}$$

Indeed, from the observation equation of the DSS model (4.4), introducing two opposite phase rotations to PHN and CIR results in the same overall distortion $e^{j\phi_t}\mathbf{h}$, 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 [Lin *et al.* 06] 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 [Lin *et al.* 06] and does not affect the final symbol detection.

4.1.4 Estimation approaches based on Monte Carlo Methods

In this chapter, 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 multicarrier symbol, on-line estimation is recursive, *i.e.* the parameter estimates are updated at each time instant.

In [Lin *et al.* 06], 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 $\widehat{\epsilon}$. In this chapter, 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 multicarrier 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) introduced in Section 3.1.1.2. 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 multicarrier training symbol is only treated one time in sequential manner while in the batch mode, the convergence of the algorithm requires several iterations over the block of N observations.

4.2 CIR, CFO, PHN and Noise Variance Estimation Based on SAEM and JCPCE

4.2.1 Introduction

The SAEM algorithm aims to iteratively estimate noise variances and CFO from the received multicarrier symbol $r_{0:N-1}$ (cf Fig. 4.2.1) before estimating the channel, the PHN and the CFO by the JCPCE proposed in [Lin *et al.* 06]. The estimate value of the CFO $\widetilde{\epsilon}$ is used as initial value of the search method described in [Lin *et al.* 06] for finding $\widehat{\epsilon}$ the optimal solution of ϵ using the maximum a posteriori criterion.

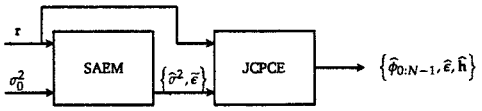

Figure 4.1 – *Structure of the proposed SAEM-JCPCE*

4.2.2 The proposed SAEM algorithm

As described in Section 3.1.1.2, given an estimate of the noise variance obtained at the $(i-1)$-th iteration and denoted by λ_{i-1}, the SAEM requires the approximation of the

posterior distribution $p(\boldsymbol{\phi}, \mathbf{h}|\mathbf{r}, \lambda_{i-1})$ using MC methods with $\boldsymbol{\phi} = \begin{bmatrix} \phi_{N-1} & \cdots & \phi_0 \end{bmatrix}^T$ and $\mathbf{r} = \begin{bmatrix} r_{N-1} & \cdots & r_0 \end{bmatrix}^T$. As detailed in Appendix A.1, the estimate of the posterior distribution of interest obtained by the Monte Carlo filter, using the importance sampling principle, at its i-th iteration is given by :

$$\widehat{p}(\boldsymbol{\phi}, \mathbf{h}|\mathbf{r}, \lambda_{i-1}) = \sum_{j=1}^{M} \widetilde{w}_i^{(j)} \delta(\boldsymbol{\phi} - \boldsymbol{\phi}_i^{(j)}) p(\mathbf{h}|\boldsymbol{\phi}_i^{(j)}, \mathbf{r}, \lambda_{i-1}) \tag{4.7}$$

where the posterior $p.d.f.$ $p(\mathbf{h}|\boldsymbol{\phi}_i^{(j)}, \mathbf{r}, \lambda_{i-1})$ is defined as :

$$p(\mathbf{h}|\boldsymbol{\phi}_i^{(j)}, \mathbf{r}, \lambda_{i-1}) = \mathcal{N}_c\left(\mathbf{h}; \mathbf{h}_i^{(j)}, \boldsymbol{\Sigma}_i^{(j)}\right) \tag{4.8}$$

with :

$$\mathbf{h}_i^{(j)} = \frac{1}{L}\left(\mathbf{C}_i^{(j)}\right)^H \left(\frac{1}{L}\mathbf{C}_i^{(j)}\left(\mathbf{C}_i^{(j)}\right)^H + \sigma_{b,i-1}^2 \mathbf{I}_N\right)^{-1} \mathbf{r} \tag{4.9}$$

$$\boldsymbol{\Sigma}_i^{(j)} = \frac{1}{L}\mathbf{I}_L - \frac{1}{L^2}\left(\mathbf{C}_i^{(j)}\right)^H \left(\frac{1}{L}\mathbf{C}_i^{(j)}\left(\mathbf{C}_i^{(j)}\right)^H + \sigma_{b,i-1}^2 \mathbf{I}_N\right)^{-1} \mathbf{C}_i^{(j)} \tag{4.10}$$

and,

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

Once the posterior distribution of interest has been approximated and given an estimate λ_{i-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.*

$$\lambda_i = \arg\max_{\lambda} \left\{ \widehat{Q}_i(\lambda) \right\} \tag{4.11}$$

where using (3.6) and (4.7),

$$\widehat{Q}_i(\lambda) = (1-\gamma_i)\widehat{Q}_{i-1}(\lambda) + \gamma_i \sum_{j=1}^{M} \widetilde{w}_i^{(j)} \int \log p(\boldsymbol{\phi}_i^{(j)}, \mathbf{h}, \mathbf{r}|\lambda) p(\mathbf{h}|\boldsymbol{\phi}_i^{(j)}, \mathbf{r}, \lambda_{i-1}) d\mathbf{h} \tag{4.12}$$

which can also be written as :

$$\widehat{Q}_i(\lambda) = (1-\gamma_i)\widehat{Q}_{i-1}(\lambda) + \gamma_i \left(A + \widehat{Q}'(\lambda, \lambda_{i-1})\right) \tag{4.13}$$

where A is a constant and $\{\gamma_i\}_{i\geq 1}$ are the forgetting factors of the SAEM which are typically $\gamma_i = i^{-\alpha}$ with $\alpha \in (1/2; 1]$. Since the $p.d.f$ $p(\boldsymbol{\phi}_i^{(j)}, \mathbf{h}, \mathbf{r}|\lambda)$ involved in (4.12) can be decomposed as :

$$p(\boldsymbol{\phi}_i^{(j)}, \mathbf{h}, \mathbf{r}|\lambda) = p(\mathbf{r}|\boldsymbol{\phi}_i^{(j)}, \mathbf{h}, \sigma_b^2) p(\boldsymbol{\phi}_i^{(j)}|\epsilon, \sigma_v^2) p(\mathbf{h}) \tag{4.14}$$

with

$$p(\boldsymbol{\phi}|\epsilon,\sigma_v^2) = \mathcal{N}(\boldsymbol{\phi};\epsilon_{i-1}\boldsymbol{\mu},\sigma_{v,i-1}^2\boldsymbol{\Theta}) \tag{4.15}$$

where the mean vector is $\boldsymbol{\mu} = \begin{bmatrix}2\pi(N-1)/N & \cdots & 2\pi/N & 0\end{bmatrix}^T$ and the covariance matrix :

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

$\widehat{Q}'(\lambda,\lambda_{i-1})$ becomes :

$$\widehat{Q}'(\lambda,\lambda_{i-1}) = \sum_{j=1}^{M} \widetilde{w}_i^{(j)} \left[\log p(\boldsymbol{\phi}_i^{(j)}|\epsilon,\sigma_v^2) + \log\left(\frac{1}{(2\pi\sigma_b^2)^{N/2}}\right) - \frac{1}{2\sigma_b^2}\left\{(\mathbf{r}-\mathbf{C}_i^{(j)}\mathbf{h}_i^{(j)})^H(\mathbf{r}-\mathbf{C}_i^{(j)}\mathbf{h}_i^{(j)}) + \text{trace}\left[\mathbf{C}_i^{(j)}\boldsymbol{\Sigma}_i^{(j)}(\mathbf{C}_i^{(j)})^H\right]\right\}\right] \tag{4.16}$$

Since distributions involved in (4.16) belong to the exponential family, we only need to propagate a set of sufficient statistics as already remarked in Chapter 3. More precisely from (4.13), we recursively approximate the set of sufficient statistics with the following update

$$\Phi_i = (1-\gamma_i)\Phi_{i-1} + \gamma_i \sum_{j=1}^{M} \widetilde{w}_i^{(j)} \Psi\left(\left\{\mathbf{h}_i^{(j)},\boldsymbol{\phi}_i^{(j)}\right\},\mathbf{r}\right) \tag{4.17}$$

The quantity $\Psi\left(\left\{\mathbf{h}_i^{(j)},\boldsymbol{\phi}_i^{(j)}\right\},\mathbf{r}\right) = \left[\Psi_1^{(j)},\Psi_2^{(j)},\Psi_3^{(j)},\Psi_4^{(j)}\right]$ involved in evaluating the set of sufficient statistics Φ_i can be written using (4.16) as :

$$\Psi_1^{(j)} = \boldsymbol{\mu}^T\boldsymbol{\Theta}^{-1}\boldsymbol{\mu} \tag{4.18}$$

$$\Psi_2^{(j)} = \boldsymbol{\mu}^T\boldsymbol{\Theta}^{-1}\boldsymbol{\phi}_i^{(j)} + (\boldsymbol{\phi}_i^{(j)})^T\boldsymbol{\Theta}^{-1}\boldsymbol{\mu} \tag{4.19}$$

$$\Psi_3^{(j)} = (\boldsymbol{\phi}_i^{(j)})^T\boldsymbol{\Theta}^{-1}\boldsymbol{\phi}_i^{(j)} \tag{4.20}$$

$$\Psi_4^{(j)} = (\mathbf{r}-\mathbf{C}_i^{(j)}\mathbf{h}_i^{(j)})^H(\mathbf{r}-\mathbf{C}_i^{(j)}\mathbf{h}_i^{(j)}) + \text{tr}\left[\mathbf{C}_i^{(j)}\boldsymbol{\Sigma}_i^{(j)}(\mathbf{C}_i^{(j)})^H\right] \tag{4.21}$$

Finally, the maximization of $\widehat{Q}_i(\lambda)$ gives the following mapping between the set of sufficient statistics and the parameter space :

$$\epsilon_i = \frac{\Phi_{i,2}}{2\Phi_{i,1}} \tag{4.22}$$

$$\sigma_{v,i}^2 = \frac{\Phi_{i,3} - \epsilon_i\Phi_{i,2} + \epsilon_i^2\Phi_{i,1}}{N} \tag{4.23}$$

$$\sigma_{b,i}^2 = \frac{\Phi_{i,4}}{N} \tag{4.24}$$

The proposed SAEM algorithm is summed up in Table 4.1.

Table 4.1 – *The proposed SAEM algorithm*

Initialization :
Set $\left\{\epsilon_0; \sigma_{v,0}^2; \sigma_{b,0}^2\right\} = \{0; 0.1; 1\}$
For $i = 1...I$
  For $m = 1...M$
    Sample $\boldsymbol{\phi}_i^{(m)}$ from the prior distribution using (4.15)
    Evaluate the *a posteriori p.d.f.* of the CIR $p(\mathbf{h}|\boldsymbol{\phi}_i^{(m)}, \sigma_{b,i-1}^2, \mathbf{r})$ using (4.9) and (4.10)
    Evaluate the corresponding importance weights using (A.6)
  end for
  Normalize importance weights : $\widetilde{w}_i^{(j),n} = w^{(j)} / \sum_{m=1}^{M} w_i^{(m)}$
  Estimation of ϵ_i, $\sigma_{v,i}^2$ and $\sigma_{b,i}^2$ using respectively (4.22), (4.23) and (4.24)
end for
Set $\left\{\widetilde{\epsilon}; \widetilde{\sigma}^2\right\} = \left\{\epsilon_I; \sigma_I^2\right\}$

4.3 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, as in Chapter 3, by the PSEM-SMC and the OEM-SMC algorithms[1].

4.3.1 Rao-Blackwellization for the CIR estimation

The *posterior distribution* $p(\mathbf{h}|\boldsymbol{\phi}_{0:t}^{(j)}, \lambda_{\lfloor t/L \rfloor}^{(j)}, r_{0:t})$ is obtained in a same way as in Appendix A.1.2 by introducing the time domain index $0 \leq t \leq N-1$ (unlike the iteration index i in the SAEM) since a sequential method is considered in this section. This

[1]Let 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 PSEM-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 $\left\{\lambda_{\lfloor t/L \rfloor}^{(j)}\right\}_{j=1}^{M} = \lambda_{\lfloor t/L \rfloor}$.

posterior *p.d.f.* is $\mathcal{N}_c\left(\mathbf{h};\mathbf{h}_t^{(j)},\mathbf{\Sigma}_t^{(j)}\right)$ with :

$$\mathbf{h}_t^{(j)} = \frac{1}{L}\left(\mathbf{C}_t^{(j)}\right)^H \left(\frac{1}{L}\mathbf{C}_t^{(j)}\left(\mathbf{C}_t^{(j)}\right)^H + \sigma^{2(j)}_{b,\lfloor t/L\rfloor}\mathbf{I}_{t+1}\right)^{-1}\mathbf{r}_t \tag{4.25}$$

and,

$$\mathbf{\Sigma}_t^{(j)} = \frac{1}{L}\mathbf{I}_L - \frac{1}{L^2}\left(\mathbf{C}_t^{(j)}\right)^H \left(\frac{1}{L}\mathbf{C}_t^{(j)}\left(\mathbf{C}_t^{(j)}\right)^H + \sigma^{2(j)}_{b,\lfloor t/L\rfloor}\mathbf{I}_{t+1}\right)^{-1}\mathbf{C}_t^{(j)} \tag{4.26}$$

where the operator $\lfloor a\rfloor$ represents the largest integer lower than or equal to a, $\mathbf{r}_t = \begin{bmatrix} r_t & \cdots & r_0 \end{bmatrix}^T$ and

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

4.3.2 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 chapter, we consider the optimal importance function for phase distortions which minimizes the variance of the importance weights conditional upon the particle trajectories and the observations [Zaritskii *et al.* 75]. In our context, it is expressed as :

$$\pi(\phi_t|\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},r_{0:t}) \;=\; p(\phi_t|\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},r_{0:t}) \tag{4.27}$$

The sampling of ϕ_t from (4.27) requires the analytical expression of the optimal importance function. However, this *p.d.f.* 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 $e^{jv_t} = 1 + jv_t$ where v_t is defined in (4.3). Generally, the usual linearization of the PHN term $e^{j\theta_t} = 1 + j\theta_t$ is employed but is less accurate than the proposed one. As detailed in Appendix A.2, this *p.d.f.* can thus be approximated by :

$$p(\phi_t|\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},r_{0:t}) \;\approx\; \mathcal{N}(\phi_t;\vartheta_t^{(j)},\Lambda_t^{(j)}) \tag{4.28}$$

where $\Lambda_t^{(j)} = \frac{\chi_t^{(j)}\sigma^{2(j)}_{v,\lfloor t/L\rfloor}}{|\Gamma_t^{(j)}|^2\sigma^{2(j)}_{v,\lfloor t/L\rfloor}+\chi_t^{(j)}}$ and

$$\vartheta_t^{(j)} = \begin{cases} \Upsilon_0^{(j)} & \text{if } t = 0 \\ \Upsilon_t^{(j)} + \phi_{t-1}^{(j)} + 2\pi\epsilon_{\lfloor t/L\rfloor}^{(j)}/N & \text{otherwise} \end{cases} \tag{4.29}$$

with $\chi_t^{(j)} = \mathbf{s}_t^T \boldsymbol{\Sigma}_{t-1}^{(j)} \mathbf{s}_t^* + \sigma_{b,\lfloor t/L \rfloor}^{2(j)}$, $\Upsilon_t^{(j)} = \frac{\Im(\Gamma_t^{(j)*} r_t)\sigma_{v\lfloor t/L \rfloor}^{2(j)}}{|\Gamma_t^{(j)}|^2 \sigma_{v,\lfloor t/L \rfloor}^{2(j)} + \chi_t^{(j)}}$ (with $\Im(.)$ the imaginary part) and

$$\Gamma_t^{(j)} = \begin{cases} \mathbf{s}_0^T \mathbf{h}_{-1}^{(j)} & \text{if } t = \bar{0} \\ e^{j(\phi_{t-1}^{(j)} + 2\pi\epsilon_{\lfloor t/L \rfloor}^{(j)}/N)} \mathbf{s}_t^T \mathbf{h}_{t-1}^{(j)} & \text{otherwise} \end{cases} \tag{4.30}$$

A Gaussian approximation of the optimal importance function is proposed in this chapter 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 [Vaswani 07]. As shown in Appendix A.3, the Gaussian approximation is accurate in our context since the prior distribution of the phase distortions is narrow enough to ensure that only one mode is significant.

The sampling distribution of ϕ_t is now identified by (4.28). The corresponding unnormalized weights are then computed by :

$$w_t^{(j)} \propto w_{t-1}^{(j)} \frac{p(r_t|\phi_{0:t}^{(j)}, \lambda_{\lfloor t/L \rfloor}^{(j)}, r_{0:t-1}) p(\phi_t^{(j)}|\phi_{0:t-1}^{(j)}, \lambda_{\lfloor t/L \rfloor}^{(j)})}{\widehat{p}(\phi_t^{(j)}|\phi_{0:t-1}^{(j)}, \lambda_{\lfloor t/L \rfloor}^{(j)}, r_{0:t})} \tag{4.31}$$

with $\widehat{p}(\phi_t^{(j)}|\phi_{0:t-1}^{(j)}, \lambda_{\lfloor t/L \rfloor}^{(j)}, r_{0:t})$ the approximate optimal importance function given by (4.28),

$$p(\phi_t|\phi_{0:t-1}^{(j)}, \lambda_{\lfloor t/L \rfloor}^{(j)}) = \begin{cases} \mathcal{N}(\phi_0; 0, \sigma_{v,\lfloor t/L \rfloor}^2) & \text{if } t = 0 \\ \mathcal{N}(\phi_t; \phi_{t-1}^{(j)} + 2\pi\epsilon_{\lfloor t/L \rfloor}^{(j)}/N, \sigma_{v,\lfloor t/L \rfloor}^2) & \text{otherwise} \end{cases} \tag{4.32}$$

$$p(r_t|\phi_{0:t}^{(j)}, \lambda_{\lfloor t/L \rfloor}^{(j)}, r_{0:t-1}) = \mathcal{N}_c(r_t; \rho_t^{(j)}, \chi_t^{(j)}) \tag{4.33}$$

and $\rho_t^{(j)} = e^{j\phi_t^{(j)}} \mathbf{s}_t^T \mathbf{h}_{t-1}^{(j)}$.

4.3.3 Estimation of CFO and noise variances

In this section, the on-line parameter schemes described in Chapter 3, *i.e.* the on-line EM (OEM-SMC) and the stochastic EM per particle (PSEM-SMC), are applied on our estimation problem. Since the PHN is a non-stationary Markov process, the approach introduced in Section 3.3.3.3 is used. In the following, for a given L and any $1 \leq k \leq N/L$, we denote $\mathbf{r}_k = r_{(k-1)L:kL-1}$ and $\boldsymbol{\phi}_k = \phi_{(k-1)L:kL-1}$.

4.3.3.1 The on-line EM estimator

As described in Chapter 3, the parameter update rule is :

$$\lambda_k = \arg\max_{\lambda} \tilde{Q}_k(\lambda_{1:k-1}, \lambda) \tag{4.34}$$

where $\tilde{Q}_k(\lambda_{1:k-1}, \lambda)$ applied on our problem is given by :

$$\tilde{Q}_k(\lambda_{1:k-1}, \lambda) = (1-\gamma_k)\tilde{Q}_{k-1}(\lambda_{1:k-2}, \lambda) + \gamma_k \int\int \log p(\boldsymbol{\phi}_k, \mathbf{h}, \mathbf{r}_k|\lambda) p(\boldsymbol{\phi}_k, \mathbf{h}|\mathbf{r}_k, \lambda_{k-1}) d\mathbf{h} d\boldsymbol{\phi}_k \quad (4.35)$$

Moreover, the SMC filter developped previously approximates the joint posterior distribution as follows :

$$p(\boldsymbol{\phi}_k, \mathbf{h}|\mathbf{r}_k, \lambda_{k-1}) \approx \sum_{j=1}^{M} \tilde{w}_{kL-1}^{(j)} \delta(\boldsymbol{\phi}_k - \boldsymbol{\phi}_k^{(j)}) \mathcal{N}_c\left(\mathbf{h}; \mathbf{h}_{kL-1}^{(j)}, \boldsymbol{\Sigma}_{kL-1}^{(j)}\right) \quad (4.36)$$

As a consequence, we have the following approximation of the Q-function :

$$\tilde{Q}_k(\lambda_{1:k-1}, \lambda) \approx (1-\gamma_k)\tilde{Q}_{k-1}(\lambda_{1:k-2}, \lambda) + \gamma_k \sum_{j=1}^{M} \tilde{w}_{kL-1}^{(j)} \int \log p(\boldsymbol{\phi}_k^{(j)}, \mathbf{h}, \mathbf{r}_k|\lambda) \mathcal{N}_c\left(\mathbf{h}; \mathbf{h}_{kL-1}^{(j)}, \boldsymbol{\Sigma}_{kL-1}^{(j)}\right) d\mathbf{h} \quad (4.37)$$

Since $p(\boldsymbol{\phi}_k^{(j)}, \mathbf{h}, \mathbf{r}_k|\lambda)$ is in the exponential family, we only need to propagate the set of sufficient statistics, Φ_k. More precisely, since the parameters to estimate are $\{\epsilon, \sigma_v^2, \sigma_b^2\}$, we recursively approximate the set of sufficient statistics with the following update rule :

$$\Phi_k = (1-\gamma_k)\Phi_{k-1} + \gamma_k \sum_{j=1}^{M} \tilde{w}_{kL-1}^{(j)} \Psi\left(\left\{\mathbf{h}_{kL-1}^{(j)}, \boldsymbol{\phi}_k^{(j)}\right\}, \mathbf{r}_k\right) \quad (4.38)$$

where the quantity $\Psi\left(\left\{\mathbf{h}_{kL-1}^{(j)}, \boldsymbol{\phi}_k^{(j)}\right\}, \mathbf{r}_k\right) = \left[\Psi_1^{(j)}, \Psi_2^{(j)}, \Psi_3^{(j)}\right]$ involved in evaluating the set of sufficient statistics Φ_k can be written as :

$$\Psi_1^{(j)} = \begin{cases} \frac{1}{L-1}\sum_{n=1}^{L-1}\left(\phi_n^{(j)} - \phi_{n-1}^{(j)}\right) & \text{if } k=1 \\ \frac{1}{L}\sum_{n=(k-1)L}^{kL-1}\left(\phi_n^{(j)} - \phi_{n-1}^{(j)}\right) & \text{otherwise} \end{cases} \quad (4.39)$$

$$\Psi_2^{(j)} = \frac{1}{L}\sum_{n=(k-1)L}^{kL-1}\left(\phi_n^{(j)} - \phi_{n-1}^{(j)}\right)^2 \quad (4.40)$$

$$\Psi_3^{(j)} = \frac{1}{L}\left(\sum_{n=(k-1)L}^{kL-1}\left|r_n - e^{j\phi_n^{(j)}}\mathbf{s}_n^T\mathbf{h}_{kL-1}^{(j)}\right|^2 + \mathrm{tr}\left[\mathbf{C}_{kL-1}^{(j)}\boldsymbol{\Sigma}_{kL-1}^{(j)}(\mathbf{C}_{kL-1}^{(j)})^H\right]\right) \quad (4.41)$$

Finally, the maximization of $\tilde{Q}_k(\lambda_{1:k-1}, \lambda)$ gives the following mapping between the set of sufficient statistics and the parameter space :

$$\epsilon_k = \frac{N}{2\pi}\Phi_{k,1} \quad (4.42)$$

$$\sigma_{v,k}^2 = \left(\Phi_{k,2} - \frac{4\pi\epsilon_k}{N}\Phi_{k,1} + \frac{4\pi^2\epsilon_k^2}{N^2} \right) \tag{4.43}$$

$$\sigma_{b,k}^2 = \Phi_{k,3} \tag{4.44}$$

The OEM-SMC is summed up in Table 4.2.

Table 4.2 – *On-line EM using Sequential Monte Carlo algorithm (OEM-SMC)*

Initialization :
$\left\{\phi_{-1}^{(j)}; \mathbf{h}_{-1}^{(j)}; \mathbf{\Sigma}_{-1}^{(j)}; \tilde{w}_{-1}^{(j)}\right\}_{j=1}^{M} = \left\{0; \mathbf{0}_{L\times 1}; \frac{1}{L}\mathbf{I}_L; 1/M\right\}$
$\{\lambda_0\} = \left\{\widehat{\sigma}_{v,0}^2; \widehat{\sigma}_{b,0}^2; \epsilon_{-1}\right\} = \{0.01; 1; 0\}$
For $k = 1...N/L$
  For $n = 0...L-1$
    For $j = 0...M$
      Sample $\phi_{(k-1)L+n}^{(j)}$ from the optimal importance function using (4.28)
      Evaluate the *a posteriori p.d.f.* of the CIR $p(\mathbf{h}|\phi_{0:(k-1)L+n}^{(j)}, \lambda_{k-1}, r_{0:(k-1)L+n})$ using (4.25) and (4.26)
      Evaluate the corresponding importance weights using (4.31)
    end for
    Normalize importance weights : $\tilde{w}_{(k-1)L+n}^{(j)} = w_{(k-1)L+n}^{(j)} / \sum_{m=1}^{M} w_{(k-1)L+n}^{(m)}$
    If $N_{eff} < N_{seuil}$ **then**
      Resample particle trajectories using systematic resampling
    end if
  end for
  Compute the set of sufficient statistics Φ_k using (4.38) and (4.39)-(4.41)
  Compute the parameter update λ_k using (4.42)-(4.44)
end for
Evaluate the final smoothed estimates:
$\star$ $\widehat{x}_{0:N-1} = \sum_{j=1}^{M} \tilde{w}_{N-1}^{(j)} x_{0:N-1}^{(j)}$ with $x_{0:N-1}^{(j)} = \left\{\phi_{0:N-1}^{(j)}, \mathbf{h}^{(j)}\right\}$
$\star$ $\widehat{\lambda} = \lambda_{N/L}$

4.3.3.2 *The stochastic EM per particle*

The parameter update rule for the j-th particle is :

$$\lambda_k^{(j)} = \arg\max_{\lambda} \tilde{Q}_k^{(j)}(\lambda_{1:k-1}^{(j)}, \lambda) \tag{4.45}$$

where $\tilde{Q}_k^{(j)}(\lambda_{1:k-1}^{(j)}, \lambda)$ applied on our problem is given by :

$$\begin{aligned}\tilde{Q}_k^{(j)}(\lambda_{1:k-1}^{(j)}, \lambda) \approx (1-\gamma_k)\tilde{Q}_{k-1}^{(j)}(\lambda_{1:k-2}^{(j)}, \lambda) \\ + \gamma_k \int \log p(\boldsymbol{\phi}_k^{(j)}, \mathbf{h}, \mathbf{r}_k|\lambda) \mathcal{N}_c\left(\mathbf{h}; \mathbf{h}_{kL-1}^{(j)}, \mathbf{\Sigma}_{kL-1}^{(j)}\right) d\mathbf{h}\end{aligned} \tag{4.46}$$

The set of sufficient statistics associated to each particle is recursively approximated with the following update rule :

$$\Phi_k^{(j)} = (1-\gamma_k)\Phi_{k-1}^{(j)} + \gamma_k \Psi\left(\left\{\mathbf{h}_{kL-1}^{(j)}, \boldsymbol{\phi}_k^{(j)}\right\}, \mathbf{r}_k\right) \tag{4.47}$$

where the quantity $\Psi\left(\left\{\mathbf{h}_{kL-1}^{(j)}, \boldsymbol{\phi}_k^{(j)}\right\}, \mathbf{r}_k\right) = \left[\Psi_1^{(j)}, \Psi_2^{(j)}, \Psi_3^{(j)}\right]$ defined by Eqs.

(4.39)-(4.41) like for the OEM-SMC. Finally, the mapping between the set of sufficient statistics and the parameter space for the j-th particle is given by :

$$\epsilon_k^{(j)} = \frac{N}{2\pi L}\Phi_{k,1}^{(j)} \tag{4.48}$$

$$\sigma_{v,k}^{2(j)} = \frac{1}{L}\left(\Phi_{k,2}^{(j)} - \frac{4\pi\epsilon_k^{(j)}}{N}\Phi_{k,1}^{(j)} + \frac{4\pi^2\epsilon_k^{2(j)}}{N^2}\right) \tag{4.49}$$

$$\sigma_{b,k}^{2(j)} = \frac{1}{L}\Phi_{k,3}^{(j)} \tag{4.50}$$

The PSEM-SMC is summed up in Table 4.3.

Table 4.3 – *Stochastic EM per particle using Sequential Monte Carlo algorithm (PSEM-SMC)*

Initialization :
$\left\{\phi_{-1}^{(m)}; \mathbf{h}_{-1}^{(j)}; \Sigma_{-1}^{(j)}; \tilde{w}_{-1}^{(j)}\right\}_{j=1}^{M} = \left\{0; \mathbf{0}_{L\times 1}; \frac{1}{L}\mathbf{I}_L; 1/M\right\}$
$\left\{\lambda_0^{(j)}\right\}_{j=1}^{M} = \left\{\tilde{\sigma}_{v,0}^{2(j)}; \tilde{\sigma}_{b,0}^{2(j)}; \epsilon_0^{(j)}\right\} = \{0.01; 1; 0\}$
For $k = 1...N/L$
  For $n = 0...L-1$
    For $j = 0...M$
      Sample $\phi_{(k-1)L+n}^{(j)}$ from the optimal importance function using (4.28)
      Evaluate the *a posteriori p.d.f.* of the CIR $p(\mathbf{h}|\phi_{0:(k-1)L+n}^{(j)}, \lambda_{k-1}^{(j)}, r_{0:(k-1)L+n})$ using (4.25) and (4.26)
      Evaluate the corresponding importance weights using (4.31)
    end for
    Normalize importance weights : $\tilde{w}_{(k-1)L+n}^{(j)} = w_{(k-1)L+n}^{(j)} / \sum_{m=1}^{M} w_{(k-1)L+n}^{(m)}$
    If $N_{eff} < N_{seuil}$ **then**
      Resample particle trajectories using systematic resampling
    end if
  end for
  Compute the set of sufficient statistics $\Phi_k^{(j)}$ using (4.47) and (4.39)-(4.41)
  Compute the parameter update $\lambda_k^{(j)}$ using (4.48)-(4.50)
end for
Evaluate the final smoothed estimates:
$\star$ $\widehat{x}_{0:N-1} = \sum_{j=1}^{M} \tilde{w}_{N-1}^{(j)} x_{0:N-1}^{(j)}$ with $x_{0:N-1}^{(j)} = \left\{\phi_{0:N-1}^{(j)}, \mathbf{h}^{(j)}\right\}$
$\star$ $\widehat{\lambda} = \sum_{j=1}^{M} \tilde{w}_{N-1}^{(j)} \lambda_{N/L}^{(j)}$

4.4 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 (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 [van Trees 68]; this bound is usually referred to as the Van Trees version of the CRB, or posterior CRB (PCRB) [Tichavsky 95].

In this topic, the PCRB for the joint estimation of $\{\boldsymbol{\phi}, \mathbf{h}, \epsilon, \sigma^2\}$ is analytically intractable. Since the CIR is the main quantity of interest, we derive in this chapter 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 :

$$p(\mathbf{r}|\mathbf{h}) = \mathcal{N}_c\left(\mathbf{r}; \mathbf{S}\mathbf{h}, \sigma_b^2 \mathbf{I}_N\right) \tag{4.51}$$

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 multicarrier symbol is equal to the covariance matrix $\mathbf{\Sigma}$ of the posterior *p.d.f.* $p(\mathbf{h}|\mathbf{r})$ and is defined using (4.51) and (A.8) by :

$$\mathbf{\Sigma} = \frac{1}{L}\mathbf{I}_L - \frac{1}{L^2}\mathbf{S}^H \left(\frac{1}{L}\mathbf{S}\mathbf{S}^H + \sigma_b^2 \mathbf{I}_N\right)^{-1} \mathbf{S} \tag{4.52}$$

where the elements of $\mathbf{S}$ correspond to the multicarrier training signal. The PCRB is thus obtained through numerical simulations by the expectation of (4.52) over the multicarrier symbol $\mathbf{S}$:

$$\text{PCRB} = \mathbb{E}\left[\text{tr}\left[\mathbf{\Sigma}\right]\right] \tag{4.53}$$

4.5 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 $L_p = 10$ taps and a uniform power delay profile, a multicarrier 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 multicarrier 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 4.1.3 is not the subject of this chapter, we assume that it can be perfectly corrected to facilitate the analysis of channel estimation as in [Lin *et al.* 06].

Through this section, we firstly study the influence of the block length L and the forgetting factor α on parameter estimation. Secondly, comparisons between the different on-line parameter estimation techniques are assessed. Finally, performance in term of MSE of both channel and phase distortions estimate is depicted and compared to the PCRB.

The MSE of variance estimates versus the block number is depicted for different values of L and SNR in Figures 4.2 and 4.3, respectively for the OEM-SMC and the PSEM-SMC. From these two figures and as remarked in Chapter 3, it can be seen that the convergence speed of both algorithms increases with L whatever

the SNR is since the parameter update is based on a longer observation data block. However in order to have an accurate parameter estimate at the end of the training multicarrier symbol, *i.e.* after processing the 64 available observations, $L = 1$ appears as a good choice for both OEM-SMC and PSEM-SMC, especially for high SNR. Indeed, for a SNR equal to 40 dB, the MSE of $\{\sigma_b^2, \sigma_v^2\}$ estimates obtained at the end of the multicarrier symbol for respectively the OEM-SMC and the PSEM-SMC is $\{1.7 \times 10^{-7}, 5.4 \times 10^{-7}\}$ and $\{7 \times 10^{-9}, 3.7 \times 10^{-9}\}$ with $L = 1$ against $\{2.5 \times 10^{-5}, 2.3 \times 10^{-5}\}$ and $\{3.2 \times 10^{-6}, 1.2 \times 10^{-7}\}$ with $L = 5$.

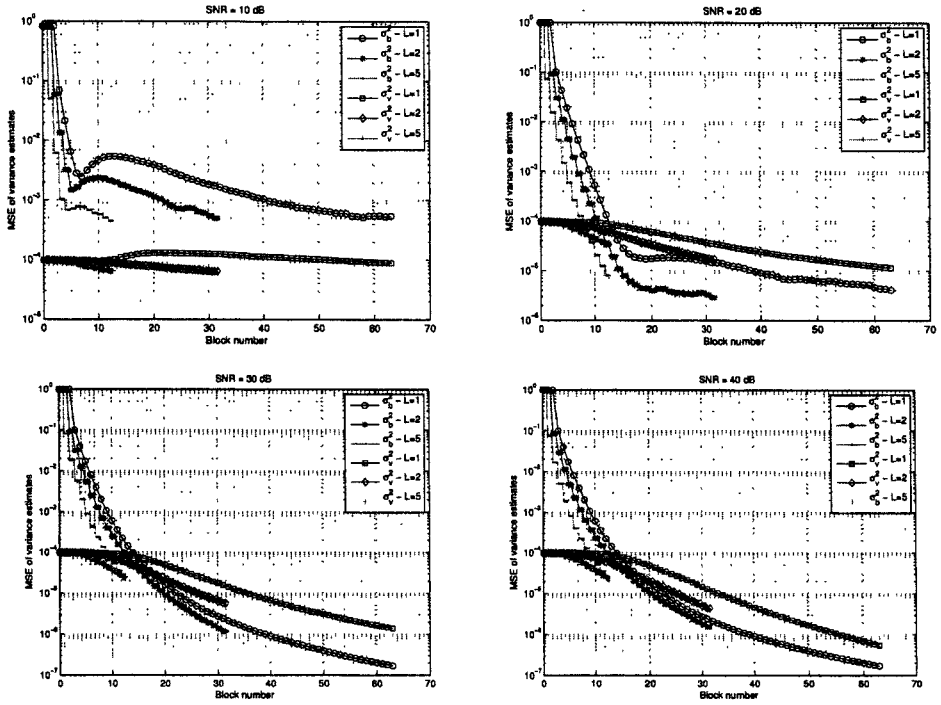

Figure 4.2 – *Influence of block length on MSE of variance estimates using OEM-SMC for different SNR with* $\alpha = 0.5$, $\beta T = 10^{-3}$ *and* $L_p = 10$ *paths.*

Now, the influence of the value α in the forgetting factor is studied for both OEM-SMC and PSEM-SMC in Figures 4.4 and 4.5 where the length of block is set to $L = 1$. As denoted in Chapter 3, both algorithms converges more slowly when α increases, especially for high SNR. By comparing these two figures, it can be remarked, that the PSEM-SMC outperforms the OEM-SMC in parameter estimation. Indeed, the PSEM-SMC converges more rapidly than the OEM-SMC. From these remarks concerning the influence on accuracy of on-line algorithm parameter, both algorithms are derived with $L = 1$ and $\alpha = 0.5$ in the following of this section.

For comparison purposes between parameter estimators, we study now 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 PSEM-SMC are drawn in Fig. 4.6 with $\beta T = 10^{-3}$ and in Fig. 4.7 with $\beta T = 10^{-2}$. A Box-and-whiskers

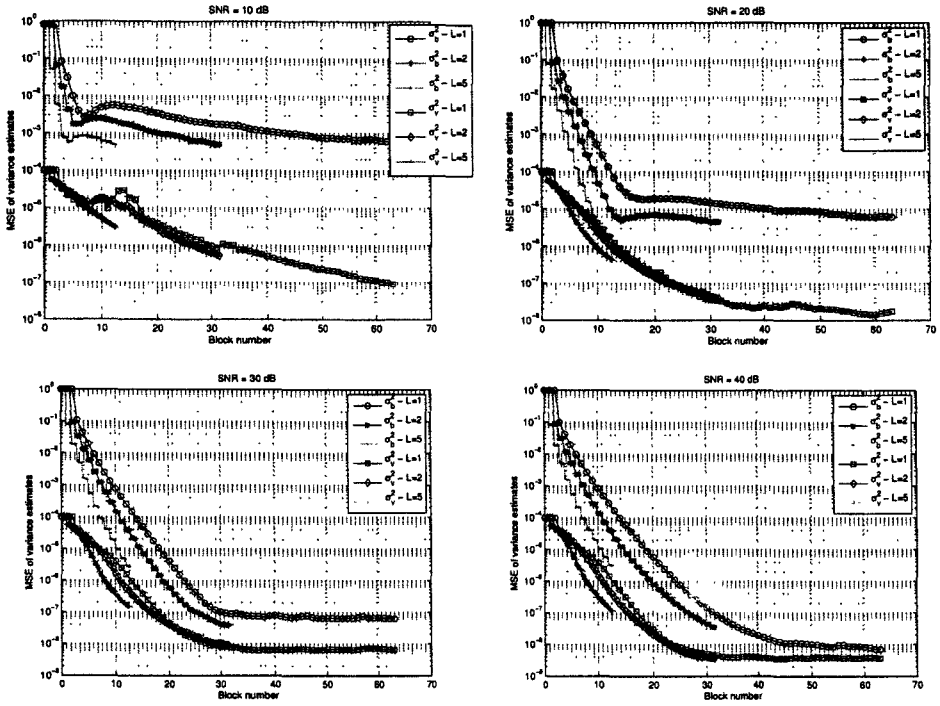

Figure 4.3 – *Influence of block length on MSE of variance estimates using PSEM-SMC for different SNR with* $\alpha = 0.5$, $\beta T = 10^{-3}$ *and* $L_p = 10$ *paths.*

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 PSEM-SMC than with the OEM-SMC. Indeed, the bias obtained by the PSEM-SMC is smaller than the one obtained by the OEM-SMC. Moreover, the PSEM-SMC offers a smaller estimation variance. These remarks about bias and variance become more pronounced for PHN variance and when the SNR decreases. Therefore, it is obvious that the diversity introduced in the search space of λ in the PSEM-SMC significantly improves the precision of the final estimates for both PHN and AWGN powers. Finally, the PSEM-SMC is more accurate in the estimation of noise variance than the SAEM. It can be explained by the sequential nature of the SMC filter which is more adapted to time-varying state estimation than MC methods. Indeed, the sampling step is improved in the PSEM-SMC since the phase distortions values are simulated sequentially by using an approximate optimal importance function which takes into account the observations. Moreover, the

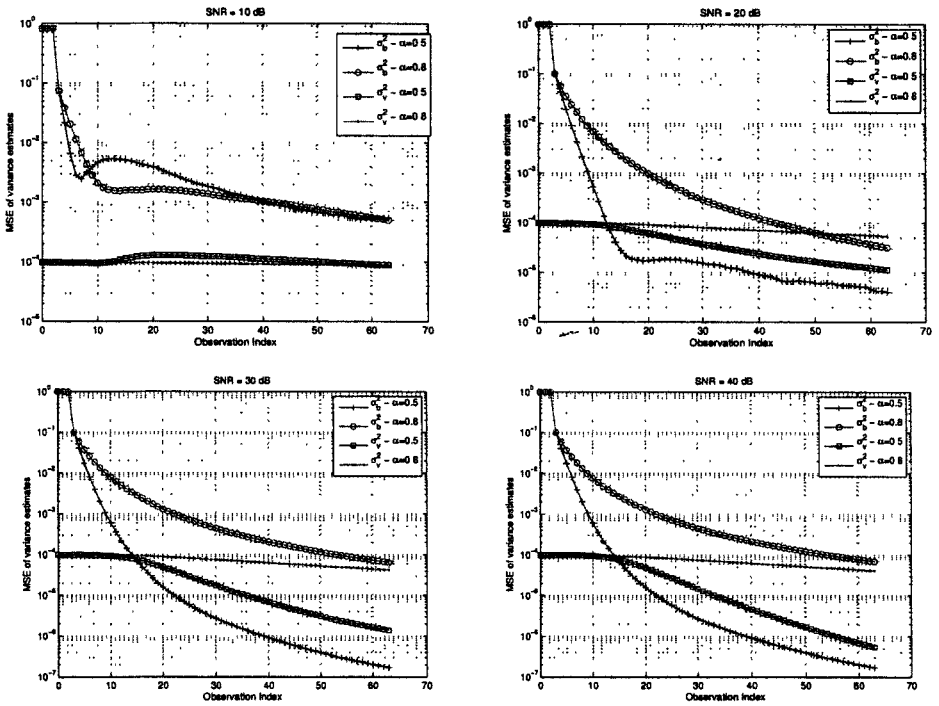

Figure 4.4 – *Influence of α on MSE of variance estimates using OEM-SMC for different SNR with $L = 1$, $\beta T = 10^{-3}$ and $L_p = 10$ paths.*

resampling step integrated in SMC algorithms allows to prevent importance sampling dimension from blowing up with time.

The MSE of phase distortions (CFO + PHN) and channel estimation are respectively depicted in Figures 4.8 and 4.9 for several values of SNR and βT and for $L_p = 10$ 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 $MSE = \frac{1}{Q}\sum_{n=1}^{Q} \text{tr}\left[\left(\mathbf{h}^n - \widehat{\mathbf{h}}^n\right)\left(\mathbf{h}^n - \widehat{\mathbf{h}}^n\right)^H\right]$ where Q is the number of multicarrier symbols used in simulations and n denotes the n-th multicarrier 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 PSEM-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 PSEM-SMC is more accurate in the noise variance estimation, the SAEM-JCPCE slightly outperforms the PSEM-SMC. It can be explained by the fact that at the beginning of the multicarrier symbol, the PSEM-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,

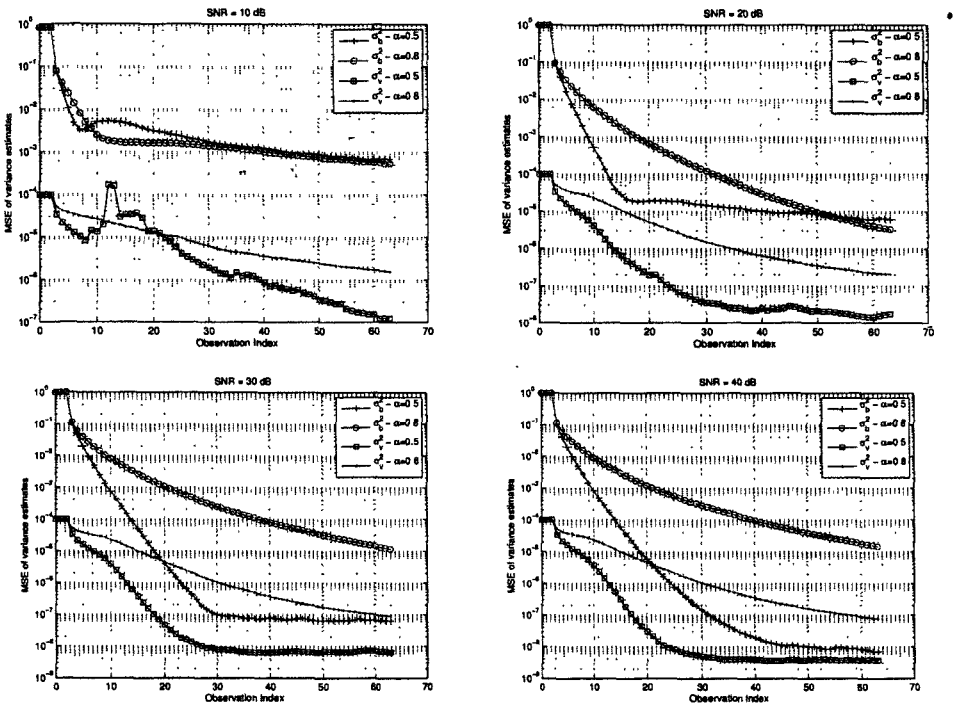

Figure 4.5 – *Influence of α on MSE of variance estimates using PSEM-SMC for different SNR with $L = 1$, $\beta T = 10^{-3}$ and $L_p = 10$ paths.*

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 multicarrier 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.

4.6 Conclusion

This chapter deals with the major problem of multicarrier channel estimation in the presence of CFO and PHN. Unlike [Lin *et al.* 06], 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 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 EM method using SMC methods proposed in [Andrieu *et al.* 03a], called the

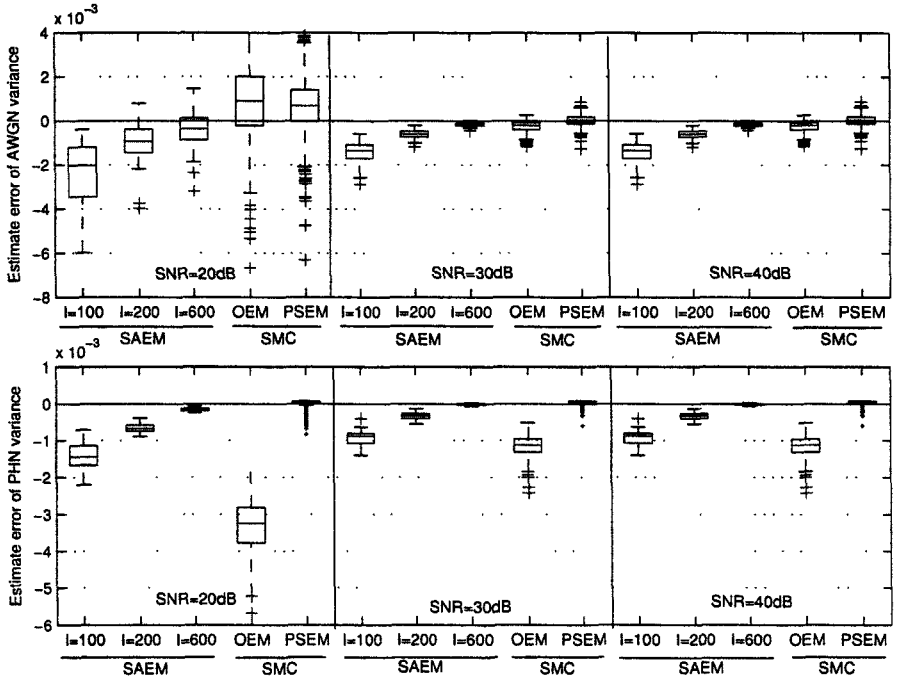

Figure 4.6 – *Box-and-whiskers plots of the estimate error of AWGN (top) and PHN (bottom) power obtained by SAEM (with different numbers of iterations I), the PMAP-SMC and the OEM-SMC algorithms for several SNR values and with* $\beta T = 10^{-3}$ *and* $L_p = 10$ *paths.*

OEM-SMC algorithm, is derived. In the latter case, an on-line SEM per particle for the parameters estimation has been proposed leading to an efficient parameter estimator, the PSEM-SMC.

Numerical simulations demonstrate the efficiency of the proposed algorithms for multicarrier 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 PSEM-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 PSEM-SMC consequently offer a good compromise between performances and complexity.

The aim is now to design an efficient multicarrier receiver for the data detection step in the presence of phase impairments.

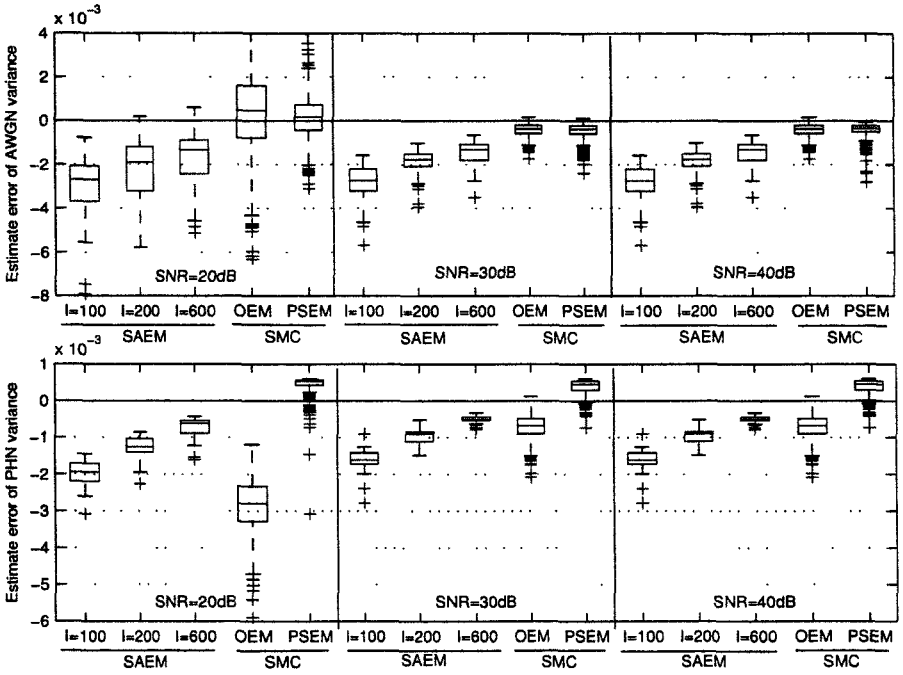

Figure 4.7 – *Box-and-whiskers plots of the estimate error of AWGN (top) and PHN (bottom) power obtained by SAEM (with different numbers of iterations I), the PMAP-SMC and the OEM-SMC algorithms for several SNR values and with* $\beta T = 10^{-2}$ *and* $L_p = 10$ *paths.*

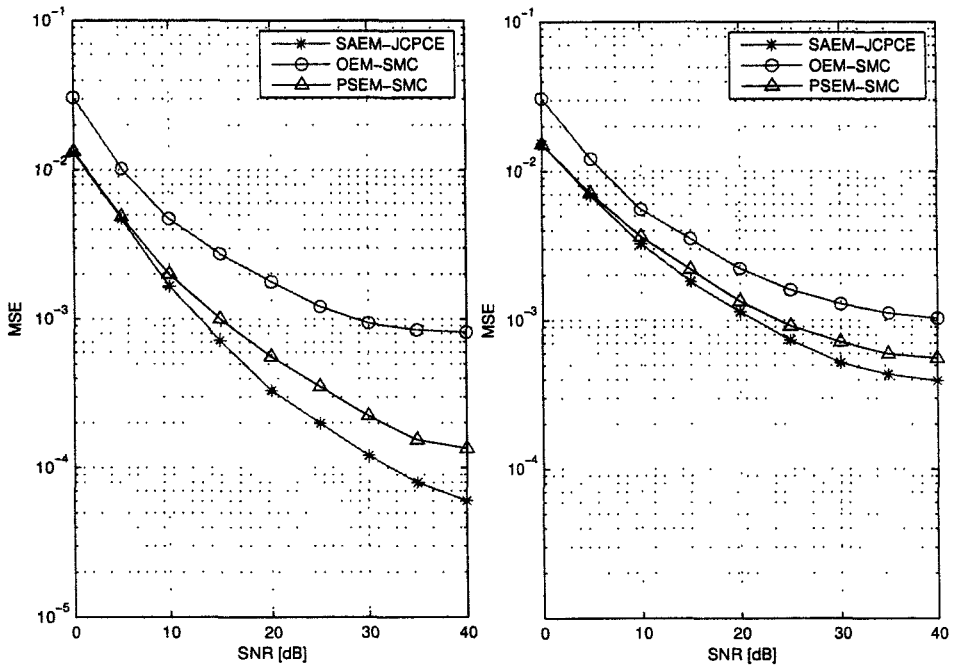

Figure 4.8 – *MSE of phase distortions (CFO + PHN) vs. SNR with* $L_p = 10$ *paths and for phase noise rate* $\beta T = 10^{-3}$ *(left) and* $\beta T = 10^{-2}$ *(right).*

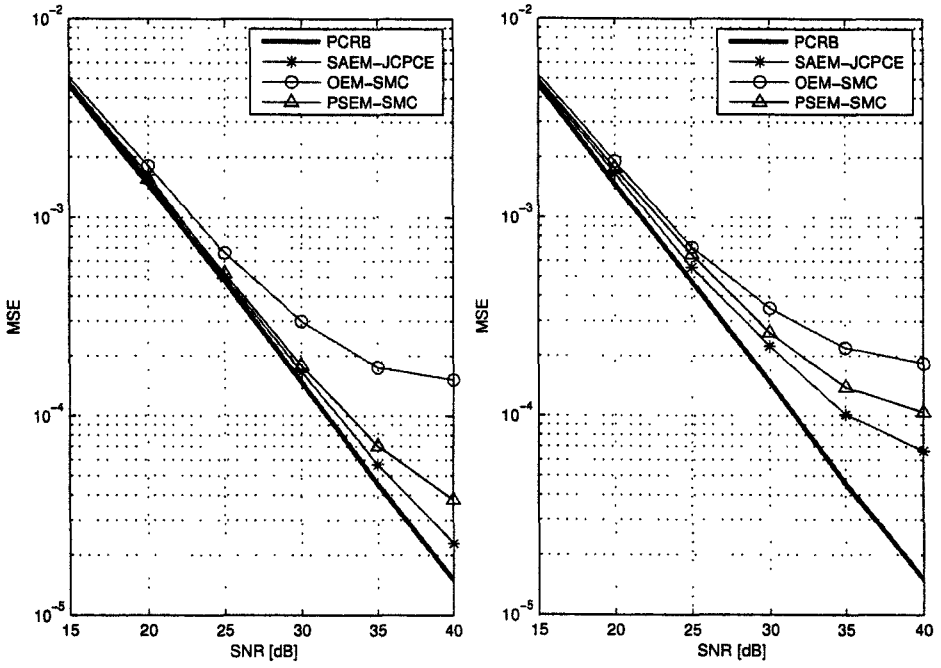

Figure 4.9 – *MSE of channel estimation vs. SNR with* $L_p = 10$ *paths and for phase noise rate* $\beta T = 10^{-3}$ *(left) and* $\beta T = 10^{-2}$ *(right).*

5 JOINT SIGNAL, PHASE NOISE AND FREQUENCY OFFSET ESTIMATION

IN this chapter, we deal with the major problem of data detection in multicarrier systems in the presence of PHN and CFO. Indeed, as shown in Section 1.2.2, multicarrier systems are very sensitive to phase noise (PHN) and carrier frequency offset (CFO) caused by the oscillator instabilities.

In literature, many approaches have been proposed to estimate and compensate phase distortions in OFDM systems. However as discussed in Section 1.3.2, all these existing schemes suffer from several drawbacks that seriously limit their performances in practical conditions. Unlike these previous works, we do not focus directly on the data symbols estimation but rather on the estimation of multicarrier signal in time domain, *i.e* $\{s_{n,t}\}_{t=0}^{N-1}$. This estimation strategy in time domain has several advantages. Firstly, the statistical a priori information about time evolution of phase distortions can be taken into account. Secondly, the redundancy information induced by the cyclic prefix can be exploited to improve the joint estimation. This additional information has been originally exploited in OFDM systems to improve data detection [Tarighat *et al.* 03]. Finally, the proposed scheme could be used without any modifications whatever the multicarrier system, the channel coding and the symbol mapping are.

In order to perform this joint estimation in time domain, we propose a sequential Monte Carlo filter. To improve the proposed filter accuracy, a variance reduction technique described in Section 2.3.6 can be applied for the multicarrier signal. Indeed, by using the central limit theorem, the signal of interest, which corresponds to a sum of independent variables, can be considered as Gaussian and thus the Rao-Blackwellization principle can be used in the SMC filter. In addition, if pilot or null subcarriers are present in the multicarrier symbol, an original autoregressive modeling of the multicarrier signal is proposed leading to efficient and robust joint estimator. Let us note that CIR and both AWGN and PHN power estimates obtained from the multicarrier training sequence using SMC algorithm described in the previous chapter remain valid in the present due to their slow changes in time. For simplicity, we simply assume in this chapter that their values are perfectly known by the receiver.

This chapter is organized as follows. In section 5.1, after a brief description of the observed signal, the unknown multicarrier signal is statistically studied leading to

its original autoregressive modeling. All available information concerning the states of interest is finally summed up in a DSS representation. The proposed SMC filter is described in Section 5.2 and the PCRB is derived in Section 5.2.6. In section 5.3, numerical results are given to demonstrate the validity of our approach. The efficiency and the robustness of the proposed particle filter algorithm are assessed in different system configurations and are compared to those of existing schemes. Finally, section 5.4 presents some concluding remarks

5.1 SYSTEM MODEL

In this chapter, we propose a new robust scheme to jointly estimate in the time domain both the transmitted signal and the phase distortions (Fig. 5.1). Specifically, this original approach can be implemented in many multicarrier system configurations without any decision-directed scheme unlike existing schemes. Since the mathematical foundation of our solution is the Bayesian theory, a dynamic state-space system (DSS) is required. Measurement equation depends on two unknown states : phase distortions $\phi_{n,k}$, which includes the CFO and the PHN, and the multicarrier signal in time domain $s_{n,k}$.

In the first part of this section, the observation equation of the DSS will be given. Secondly in order to take into account all available information about the unknown multicarrier signal, a statistical analysis of $s_{n,k}$ is achieved. An original modeling of this process by an autoregressive model including the redundancy information of the cyclic prefix is then proposed. Finally, all these states and observation equations lead to the description of the dynamic state-space (DSS) model.

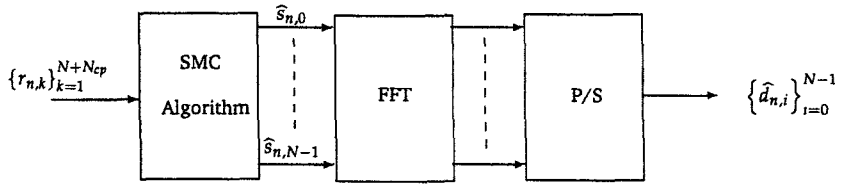

Figure 5.1 – *Block diagram of the proposed multicarrier receiver.*

5.1.1 Observation equation

The received signal $r_{n,t}$ corrupted by both PHN and CFO can be written before removal of the cyclic prefix as follows, for $t = 0 \ldots N + N_{cp} - 1$:

$$r_{n,t} = e^{j\phi_{n,t}} \sum_{l=0}^{L_p-1} h_{n,l} s_{n,t-N_{cp}-l} + w_{n,t} \tag{5.1}$$

where $w_{n,t}$ is a circular zero mean Gaussian noise with variance σ_w^2. This observation equation takes into account both the insertion of the cyclic prefix and the inter-symbol interference (ISI) due to the multipath channel with the use of the following definition

of $s_{n,t}$ for $t < 0$:

$$s_{n,t} = \begin{cases} s_{n,N+t} & \text{if } -N_{cp} \leq t \leq -1 \\ s_{n-1,t+N+N_{cp}} & \text{if } t < -N_{cp} \end{cases} \tag{5.2}$$

Since in a realistic multicarrier systems as introduced in Section 1.1.4, P pilot tones and N_g null subcarriers have been inserted at the transmitter side, we can decompose $s_{n,t}$ into a known term $f_{n,t}$ and an unknown data signal $u_{n,t}$ defined respectively as :

$$f_{n,t} = \frac{1}{\sqrt{N}} \sum_{i \in \Omega} d_{n,i} e^{j2\pi it/N} \tag{5.3}$$

and

$$u_{n,t} = \frac{1}{\sqrt{N}} \sum_{\substack{i=0 \\ i \notin \Omega}}^{N-1} d_{n,i} e^{j2\pi it/N} \tag{5.4}$$

Consequently, (5.1) can be rewritten in the following matrix form :

$$r_{n,t} = e^{j\phi_{n,t}} \mathbf{h}_n^T (\mathbf{f}_{n,t} + \mathbf{u}_{n,t}) + w_{n,t} \tag{5.5}$$

where

$$\mathbf{u}_{n,t} = \begin{bmatrix} u_{n,t-N_{cp}} & \cdots & u_{n,-N_{cp}-L_p+1} & \mathbf{0}_{1\times(N+N_{cp}-t-1)} \end{bmatrix}^T \tag{5.6}$$

$$\mathbf{f}_{n,t} = \begin{bmatrix} f_{n,t-N_{cp}} & \cdots & f_{n,-N_{cp}-L_p+1} & \mathbf{0}_{1\times(N+N_{cp}-t-1)} \end{bmatrix}^T \tag{5.7}$$

$$\mathbf{h}_n = \begin{bmatrix} h_{n,0} & \cdots & h_{n,L_p-1} & \mathbf{0}_{1\times(N+N_{cp}-1)} \end{bmatrix}^T \tag{5.8}$$

The observation equation (5.5) involves two unknown states: the CFO and the PHN included in $\phi_{n,t}$ and the useful part of the transmitted multicarrier signal $\mathbf{u}_{n,t}$. The general objective is to jointly and adaptively estimate these two dynamic states using the set of received signals, $r_{n,t}$, with $t = 0, \cdots, N + N_{cp} - 1$. Since the *a priori* dynamic feature of $\phi_{n,t}$ is already given by (4.3), only the state equation of $\mathbf{u}_{n,t}$ is required for the joint a posteriori estimation.

5.1.2 Modeling of the unknown data signal process

5.1.2.1 *Statiscal properties of the unknown data signal*

In this section, the unknown data signal $u_{n,t}$ is analyzed in order to benefit of all *a priori* statistical information about its evolution. By using the independence property of the information data in Eq. (1.35), the following relation holds :

$$E[u_{n,t}u_{n-1,l}^*] = 0 \qquad \forall \quad t,l \in [-N_{cp}; N-1] \tag{5.9}$$

Moreover, by taking into account that the cyclic prefix in the multicarrier is a copy of the last portion of the symbol appended to the front of the multicarrier symbol, we

have :

$$u_{n,t} = u_{n,t+N} \qquad \forall \quad t \in [-N_{cp}; -1] \tag{5.10}$$

Therefore with this redundancy information, in a same multicarrier symbol, the useful signal may be characterized as a cyclostationary process with period N. Consequently, the signal autocorrelation function is only studied for $(t-l) < N$. In this case, it can be written as :

$$C_{uu}(t-l) = E[u_{n,t}u^*_{n,l}] = \frac{1}{N}\sum_{\substack{i=0 \\ \notin\Omega}}^{N-1}\sum_{\substack{i'=0 \\ \notin\Omega}}^{N-1} E[d_{n,i}d^*_{n,i'}]e^{j2\pi it/N}e^{-j2\pi i'l/N} \tag{5.11}$$

Since $d_{n,i}$ are *i.i.d.* and have unit power, the latter expression can be simplified :

$$\begin{aligned} C_{uu}(t-l) &= \frac{1}{N}\sum_{\substack{i=0 \\ \notin\Omega}}^{N-1} e^{j2\pi i(t-l)/N} \\ &= \begin{cases} \frac{N-P-N_g}{N} & \text{if} \quad t = l \\ -\frac{1}{N}\sum_{i\in\Omega} e^{j2\pi i(t-l)/N} & \text{if} \quad t \neq l \end{cases} \end{aligned} \tag{5.12}$$

This correlation function clearly shows that, for $(t-l) < N$, $u_{n,t}$ is colored. Moreover, in (5.4), we can denote that the term $u_{n,t}$ is the sum of $(N-P-N_g)$ independent random variables. According to the central limit theorem, this signal can be approximated as a circular Gaussian distributed random variable. The accuracy of this Gaussian approximation is illustrated in Appendix A.4. Consequently, we propose to model $u_{n,t}$ by an autoregressive model.

5.1.2.2 *Autoregressive modeling of the unknown data signal process*

In this section, an autoregressive model for the unknown data signal $u_{n,t}$ is proposed in order to take into account, in the dynamic state-space model for the derivation of the proposed joint a posteriori estimator, all statistical information previously obtained.

Due to the independence of the process between two different multicarrier symbols (5.9), the process at time t can be modeled via the time domain recursion by the complex AR model of order $(t-1)$ (*i.e.* AR$(t-1)$) :

$$u_{n,t} = -\sum_{i=1}^{t-1} a_{n,i}u_{n,t-i} + b_{n,t} \tag{5.13}$$

where $b_{n,t}$ is a circular white Gaussian noise. The AR model parameters consist of the filter coefficients $\{a_{n,1}, a_{n,2}, \cdots, a_{n,t-1}\}$ and the driving noise variance $\sigma^2_{b_{n,t}}$.

They are obtained by solving the Yule-Walker equation :

$$\mathbf{C_{n,t}a_{n,t}} = -\mathbf{x_{n,t}} \tag{5.14}$$

where $\mathbf{C_{n,t}}$ is defined as :

$$\mathbf{C_{n,t}} = \begin{bmatrix} C_{uu}(0) & C_{uu}(-1) & \cdots & C_{uu}(-k+2) \\ C_{uu}(1) & C_{uu}(0) & \cdots & C_{uu}(-k+3) \\ \vdots & \vdots & \ddots & \vdots \\ C_{uu}(k-2) & C_{uu}(p-2) & \cdots & C_{uu}(0) \end{bmatrix} \tag{5.15}$$

and,

$$\mathbf{a}_{n,t} = \begin{bmatrix} a_{n,1} & a_{n,2} & \cdots & a_{n,t-1} \end{bmatrix}^T$$

$$\mathbf{x}_{n,t} = \begin{bmatrix} C_{uu}(1) & C_{uu}(2) & \cdots & C_{uu}(k-1) \end{bmatrix}^T$$

with $C_{uu}(.)$ given by (5.12). Finally, the variance of the noise process is obtained using :

$$\sigma^2_{b_{n,t}} = C_{uu}(0) + \mathbf{a}^T_{n,t}\mathbf{x}_{n,t} \tag{5.16}$$

Equation (5.14) can be solved efficiently by the Levinson-Durbin recursion. However, as the process $u_{n,t}$ consists of a sum of $(N - P - N_g)$ sinusoids, a necessary condition is to have the order of the AR process such as $t - 1 \leq (N - P - N_g - 1)$ [Haykin 91]. With this assumption, the inverse $\mathbf{C^{-1}_{n,t}}$ exists and the Yule-Walker equations have an unique solution : $\mathbf{a}_{n,t} = -\mathbf{C^{-1}_{n,t}}\mathbf{x_{n,t}}$. Nevertheless, as shown in [Baddour *et al.* 05], adding a small value τ_0 to its principle diagonal enables the stability and accuracy of larger order AR models. This strategy is equivalent to adding white noise of variance τ_0 to the original process. The addition of this spectral bias removes the bandlimitation of the original spectrum and creates a nondeterministic or regular process that in some sense closely approximates the original process. Consequently with the use of this approach, the order of the AR process can be chosen up to $N - 1$ (the largest possible order in our context). The choice of τ_0 represents thus a good tradeoff between the improvement of the AR modeling of $u_{n,t}$ and the bias introduced in the zeroth autocorrelation lag (*i.e.* in the signal power).

Moreover, since this procedure requires solving set of all orders less than and including the maximum AR order $(N - 1)$, the AR coefficients can be obtained at the intermediate steps of a single Levinson-Durbin execution in $O\left((N-1)^2\right)$.

5.1.2.3 *State equation of the unknown data signal*

Using relations obtained previously, the state equation of the vector $\mathbf{u}_{n,t}$ can finally be written in the matrix form as :

$$\mathbf{u}_{n,t} = \mathbf{A}_{n,t}\mathbf{u}_{n,t-1} + \mathbf{b}_{n,t} \tag{5.17}$$

where the transition matrix $\mathbf{A}_{n,t}$ is defined as :

$$\mathbf{A}_{n,t} = \begin{bmatrix} & \xi^T_{n,t} \\ \mathbf{I}_{(N+N_{cp}+L-2)} & \\ & \mathbf{0}_{(N+N_{cp}+L-2)\times 1} \end{bmatrix} \tag{5.18}$$

with using (5.10) and (5.13),

$$\boldsymbol{\xi}_{n,t} = \begin{cases} \left[\mathbf{a_{n,t}} \quad \mathbf{0}_{1\times(N+N_{cp}+L-t)}\right]^T & \text{if} \quad 0 \leq t \leq N-1 \\ \left[\mathbf{0}_{1\times(N-1)} \quad 1 \quad \mathbf{0}_{1\times(N_{cp}+L-1)}\right]^T & \text{if} \quad N \leq t \leq N+N_{cp}-1 \end{cases} \tag{5.19}$$

Finally, $\mathbf{b}_{n,t}$ is a $(N + N_{cp} + L - 1)$-by-1 zero mean Gaussian noise vector with covariance matrix :

$$E[\mathbf{b}_{n,t}\mathbf{b}_{n,t}^H] = \begin{bmatrix} \sigma^2_{b_{n,t}} & & 0 \\ & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \tag{5.20}$$

where using (5.16),

$$\sigma^2_{b_{n,t}} = \begin{cases} C_{uu}(0) + \sum_{i=1}^{k-1} a_{n,i} C_{uu}(-i) & \text{if} \quad 0 \leq t \leq N-1 \\ 0 & \text{if} \quad N \leq t \leq N+N_{cp}-1 \end{cases} \tag{5.21}$$

Remark : If we consider that no pilot and no nullsubcarrier are present in the multicarrier symbol, i.e. $\Omega = \emptyset$, the useful data signal correspond to the transmitted multicarrier signal and thus can be expressed as :

$$u_{n,t} = s_{n,t} \tag{5.22}$$

From this equation and using (5.12), we can denote that the correlation function of this process is null except in 0 where $E[|u_{n,t}|^2] = 1$. Therefore, the vector $\boldsymbol{\xi}_{n,t}$ defined in general case in (5.19) can be simplified as follows :

$$\boldsymbol{\xi}_{n,t} = \begin{cases} \left[\mathbf{0}_{1\times(N+N_{cp}+L-1)}\right]^T & \text{if} \quad 0 \leq t \leq N-1 \\ \left[\mathbf{0}_{1\times(N-1)} \quad 1 \quad \mathbf{0}_{1\times(N_{cp}+L-1)}\right]^T & \text{if} \quad N \leq t \leq N+N_{cp}-1 \end{cases} \tag{5.23}$$

and the variance $\sigma^2_{b_{n,t}} = 1$ for $\leq t \leq N-1$.

5.1.3 Dynamic state space model

By using the phase distortions equation including both CFO and PHN (4.3) and the proposed AR modeling of the unknown multicarrier signal (5.17), we obtain the following DSS model :

$$\begin{cases} \phi_{n,t} = \begin{cases} v_{n,0} & \text{if } t = 0 \\ \phi_{n,t-1} + 2\pi\epsilon/N + v_{n,t} & \text{if } t = 1, ..., N+N_{cp}-1 \end{cases} \\ \mathbf{u}_{n,t} = \mathbf{A}_{n,t}\mathbf{u}_{n,t-1} + \mathbf{b_{n,t}} \\ r_{n,t} = e^{j\phi_{n,t}}\mathbf{h}_n^T(\mathbf{f}_{n,t} + \mathbf{u}_{n,t}) + w_{n,t} \end{cases} \tag{5.24}$$

In order to jointly estimate $\phi_{n,t}$, ϵ and $\mathbf{u}_{n,t}$, we need the joint posterior probability density function (*p.d.f.*) $p(\phi_{n,t}, \epsilon, \mathbf{u}_{n,t}|r_{n,t})$. Unfortunately, this *p.d.f.* is analytically intractable and so we propose to numerically approximate $p(\phi_{n,0:t}, \epsilon, \mathbf{u}_{n,0:t}|r_{n,0:t})$ via SMC

methodology [Doucet *et al.* 01]. Let us remark that, unlike in the previous chapter, $\{\mathbf{h}_n, \sigma_v^2; \sigma_b^2\}$ are assumed perfectly known since these variables have been previously estimated from the preambule section of the multicarrier frame by algorithm described in Chapter 4.

5.2 SMC METHOD FOR JOINT MULTICARRIER SIGNAL, CFO AND PHN ESTIMATION

In this section, we describe in detail all the steps required in the implementation of the proposed scheme which aims to approximate $p(\phi_{n,0:t}, \epsilon, \mathbf{u}_{n,0:t}|r_{n,0:t})$. Using Bayes' theorem, the posterior density function of interest can be decomposed as :

$$p(\phi_{n,0:t}, \epsilon, \mathbf{u}_{n,t}|r_{n,0:t}) = p(\mathbf{u}_{n,t}|\phi_{n,0:t}, r_{n,0:t})p(\phi_{n,0:t}, \epsilon|r_{n,0:t}) \tag{5.25}$$

In order to provide the best approximation of this distribution, we take advantage of the linear substructure contained in the DSS model. The corresponding variables are marginalized out and estimated using an optimal linear filter. Indeed, conditioned on the nonlinear state variable $\phi_{n,t}$, there is a linear and Gaussian sub-structure in (5.24). As a consequence, $p(\mathbf{u}_{n,t}|\phi_{n,0:t}, r_{n,0:t})$ is, unlike $p(\phi_{n,0:t}, \epsilon|r_{n,0:t})$, analytically tractable and is obtained via a Kalman filter. The marginal posterior distribution $p(\phi_{n,0:t}, \epsilon|r_{n,0:t})$ can be approximated with a SMC filter :

$$\widehat{p}(\phi_{n,0:t}, \epsilon|r_{n,0:t}) = \sum_{j=1}^{N} \delta(\phi_{n,0:t} - \phi_{n,0:t}^{(j)}; \epsilon - \epsilon_{\lfloor t/L \rfloor}^{(j)})\widetilde{w}_{n,t}^{(j)} \tag{5.26}$$

where $\delta(.;.)$ is the two-dimensional Dirac delta function. Thus, substituting (5.26) in (5.25), we obtain an estimate of the joint a posteriori *p.d.f.* :

$$\widehat{p}(\phi_{n,0:t}, \epsilon, \mathbf{u}_{n,t}|r_{n,0:t}) = \sum_{j=1}^{N} p(\mathbf{u}_{n,t}|\phi_{n,0:t}^{(j)}, r_{n,0:t})\delta(\phi_{n,0:t} - \phi_{n,0:t}^{(j)}; \epsilon - \epsilon_{\lfloor t/L \rfloor}^{(j)})\widetilde{w}_{n,t}^{(j)} \tag{5.27}$$

5.2.1 Posterior distribution of the useful signal

The posterior distribution $p(\mathbf{u}_{n,t}|\phi_{n,0:t}^{(j)}, r_{n,0:t})$ is equal to $\mathcal{N}_c(\mathbf{u}_{n,t|t}^{(j)}, \mathbf{\Sigma}_{n,t|t}^{(j)})$ where the mean vector and the covariance matrix are obtained using the following Kalman filtering equations :

$$\text{Time update equations} \begin{cases} \mathbf{u}_{n,t|t-1}^{(j)} = \mathbf{A}_{n,t}\mathbf{u}_{n,t-1|t-1}^{(j)} \\ \mathbf{\Sigma}_{n,t|t-1}^{(j)} = \mathbf{A}_{n,t}\mathbf{\Sigma}_{n,t-1|t-1}^{(j)}\mathbf{A}_{n,t}^H + E[\mathbf{b}_{n,t}\mathbf{b}_{n,t}^H] \end{cases} \tag{5.28}$$

and

$$\text{Measurement update equations} \begin{cases} G_{n,t}^{(j)} = \mathbf{h}_n^T \boldsymbol{\Sigma}_{n,t|t-1}^{(j)} \mathbf{h}_n^* + \sigma_w^2 \\ \mathbf{k}_{n,t}^{(j)} = \boldsymbol{\Sigma}_{n,t|t-1}^{(j)} (e^{j\phi_{n,t}^{(j)}} \mathbf{h}_n^T)^H (G_{n,t}^{(j)})^{-1} \\ \mathbf{u}_{n,t|t}^{(j)} = \mathbf{u}_{n,t|t-1}^{(j)} + \mathbf{k}_{n,t}^{(j)} (r_{n,t} - e^{j\phi_{n,t}^{(j)}} \mathbf{h}_n^T (\mathbf{u}_{n,t|t-1}^{(j)} + \mathbf{f}_{n,t})) \\ \boldsymbol{\Sigma}_{n,t|t}^{(j)} = \boldsymbol{\Sigma}_{n,t|t-1}^{(j)} - \mathbf{k}_{n,t}^{(j)} e^{j\phi_{n,t}^{(j)}} \mathbf{h}_n^T \boldsymbol{\Sigma}_{n,t|t-1}^{(j)} \end{cases} \tag{5.29}$$

In (5.29), we can notice that $G_{n,t}^{(j)}$ and $\boldsymbol{\Sigma}_{n,t|t}^{(j)}$ are independent of the particle coordinates $\phi_{n,t}^{(j)}$ and thus are identical for all the particles. This remark can be used to reduce the complexity of our algorithm.

Now, the posterior distribution of the multicarrier signal $\mathbf{u}_{n,t}$ is identified, so the remaining task is the simulation of the particles in (5.27).

5.2.2 Phase distortion sampling

As in Chapter 4, in order to improve the importance sampling step, the optimal importance function

$$\pi(\phi_{n,t}|\phi_{n,0:t-1}^{(j)}, \epsilon_{\lfloor t/L \rfloor}^{(j)}, r_{0:t}) = p(\phi_{n,t}|\phi_{n,0:t-1}^{(j)}, \epsilon_{\lfloor t/L \rfloor}^{(j)}, r_{0:t}) \tag{5.30}$$

has to be approximated since this *p.d.f.* is analytically intractable. In a similar way as in Chapter 4 and also in EKF principle, a local linearization of the state equation (*i.e.* $e^{jv_{n,t}} = 1 + jv_{n,t}$) is used in order to approximate the optimal importance function.

As detailed in Appendix A.6, this *p.d.f.* can thus be approximated by :

$$p(\phi_{n,t}|\phi_{n,0:t-1}^{(j)}, \epsilon_{\lfloor t/L \rfloor}^{(j)}, r_{0:t}) \approx \mathcal{N}(\phi_{n,t}; \mu_{n,t}^{(j)}, \Lambda_t^{(j)}) \tag{5.31}$$

where

$$\mu_{n,t}^{(j)} = \begin{cases} \gamma_0^{(j)} & \text{if } t = 0 \\ \gamma_t^{(j)} + \phi_{n,t-1}^{(j)} + 2\pi\epsilon_{\lfloor t/L \rfloor}^{(j)}/N & \text{otherwise} \end{cases} \tag{5.32}$$

and,

$$\Lambda_t^{(j)} = \frac{\chi_t^{(j)} \sigma_v^2}{|\Gamma_t^{(j)}|^2 \sigma_v^2 + \chi_t^{(j)}} \tag{5.33}$$

with $\gamma_t^{(j)} = \frac{\Im(\Gamma_t^{(j)*} r_{n,t}) \sigma_v^2}{|\Gamma_t^{(j)}|^2 \sigma_v^2 + \chi_t^{(j)}}$ (where $\Im(.)$ denotes the imaginary part), $\chi_t^{(j)} = \mathbf{h}_n^T \boldsymbol{\Sigma}_{n,t|t-1}^{(j)} \mathbf{h}_n^* + \sigma_b^2$ and,

$$\Gamma_t^{(j)} = \begin{cases} \mathbf{h}_n^T (\mathbf{u}_{n,0|-1}^{(j)} + \mathbf{f}_{n,0}) & \text{if } t = 0 \\ e^{j(\phi_{n,t-1}^{(j)} + 2\pi\epsilon_{\lfloor t/L \rfloor}^{(j)}/N)} \mathbf{h}_n^T (\mathbf{u}_{n,t|t-1}^{(j)} + \mathbf{f}_{n,t}) & \text{otherwise} \end{cases} \tag{5.34}$$

where $\mathbf{u}_{n,0|-1}^{(j)}$ is only composed of the $L-1$ useful signal estimate samples obtained for the previous $n-1$-th multicarrier symbol.

5.2.3 Evaluation of the importance weights

The importance weights in the proposed marginalized particle filter are updated according to the relation :

$$w_{n,t}^{(j)} \propto w_{n,t-1}^{(j)} \frac{p(r_{n,t}|\phi_{n,t}^{(j)}, r_{n,t-1}) p(\phi_{n,t}^{(j)}|\phi_{n,t-1}^{(j)}, \epsilon_{\lfloor t/L \rfloor}^{(j)})}{\widehat{p}(\phi_{n,t}^{(j)}|\phi_{n,0:t-1}^{(j)}, \epsilon_{\lfloor t/L \rfloor}^{(j)}, r_{n,0:t})} \tag{5.35}$$

where $\widehat{p}(\phi_{n,t}^{(j)}|\phi_{n,0:t-1}^{(j)}, \epsilon_{\lfloor t/L \rfloor}^{(j)}, r_{n,0:t})$ is the approximate optimal importance function given by (5.31),

$$p(r_{n,t}|\phi_{n,t}^{(j)}, r_{n,t-1}) = \mathcal{N}_c(r_{n,t}; e^{j\phi_{n,t}^{(j)}} \mathbf{h}_n^T (\mathbf{u}_{n,t|t-1}^{(j)} + \mathbf{f}_{n,t}), \mathbf{G}_{n,t}^{(j)}) \tag{5.36}$$

with $\mathbf{G}_{n,t}^{(j)}$ the innovation covariance of the j-th Kalman filter given by (5.29) and the prior distribution of $\phi_{n,t}$ is given using (5.24) by :

$$p(\phi_{n,t}^{(j)}|\phi_{n,t-1}^{(j)}, \epsilon^{(j)}) = \begin{cases} \mathcal{N}\left(\phi_{n,0}^{(j)}; 0, \sigma_v^2\right) & \text{if } k = 0 \\ \mathcal{N}\left(\phi_{n,i}^{(j)}; \phi_{n,i-1}^{(j)} + 2\pi\epsilon_{\lfloor t/L \rfloor}^{(j)}/N, \sigma_v^2\right) & \text{otherwise} \end{cases} \tag{5.37}$$

5.2.4 CFO estimation

For the CFO estimation, the proposed SEM per particle is applied as in Chapter 4. For a given L and any $1 \leq k \leq N/L$, we denote $\mathbf{r}_{n,k} = r_{n,(k-1)L:kL-1}$ and $\boldsymbol{\phi}_{n,k} = \phi_{n,(k-1)L:kL-1}$.

$$\epsilon_{n,k}^{(j)} = \arg\max_{\epsilon} \tilde{Q}_{n,k}^{(j)}(\epsilon_{n,1:k-1}^{(j)}, \epsilon) \tag{5.38}$$

where $\tilde{Q}_{n,k}^{(j)}(\epsilon_{1:k-1}^{(j)}, \lambda)$ applied on our estimation problem is given by :

$$\begin{aligned} \tilde{Q}_{n,k}^{(j)}(\epsilon_{n,1:k-1}^{(j)}, \epsilon) &\approx (1-\gamma_k)\tilde{Q}_{n,k-1}^{(j)}(\epsilon_{n,1:k-2}^{(j)}, \epsilon) \\ &+ \gamma_k \int \log p(\boldsymbol{\phi}_{n,k}^{(j)}, \mathbf{u}_{n,kL-1}, \mathbf{r}_{n,k}|\epsilon) \mathcal{N}_c\left(\mathbf{u}_{n,kL-1}; \mathbf{u}_{n,kL-1|kL-1}^{(j)}, \boldsymbol{\Sigma}_{n,kL-1|kL-1}^{(j)}\right) d\mathbf{u}_{n,kL-1} \end{aligned} \tag{5.39}$$

The set of sufficient statistics associated to each particle is recursively approximated with the following update rule :

$$\Phi_k^{(j)} = (1-\gamma_k)\Phi_{k-1}^{(j)} + \gamma_k \Psi\left(\left\{\mathbf{u}_{n,kL-1}^{(j)}, \boldsymbol{\phi}_{n,k}^{(j)}\right\}, \mathbf{r}_k\right) \tag{5.40}$$

where the quantity $\Psi\left(\left\{\mathbf{u}_{n,kL-1}^{(j)}, \boldsymbol{\phi}_{n,k}^{(j)}\right\}, \mathbf{r}_k\right) = \left[\Psi_1^{(j)}\right]$ is defined as follows :

$$\Psi_1^{(j)} = \sum_{i=(k-1)L}^{kL-1} \left(\phi_{n,i}^{(j)} - \phi_{n,i-1}^{(j)}\right) \tag{5.41}$$

The mapping between the set of sufficient statistics and the parameter space for the j-th particle is given by :

$$\epsilon_k^{(j)} = \frac{N}{2\pi L}\Phi_{k,1}^{(j)} \tag{5.42}$$

5.2.5 MMSE estimate of multicarrier signal, PHN and CFO

Every element required in the implementation of the marginalized particle filtering algorithm has been identified. The resulting weighted samples $\left\{\mathbf{u}_{n,t|t}^{(j)}, \mathbf{\Sigma}_{n,t|t}^{(j)}, \phi_{n,0:t}^{(j)}, \epsilon^{(j)}, \widetilde{w}_{n,t}^{(j)}\right\}_{j=1}^{M}$ approximate the posterior density function $p(\mathbf{u}_{n,t}, \phi_{n,0:t}, \epsilon | r_{n,0:t})$. Consequently, the minimum mean square error (MMSE) estimates of $\mathbf{u}_{n,t}$, $\phi_{n,t}$ and ϵ are obtained at the end of the n-th multicarrier symbol by the respective expressions :

$$\widehat{\mathbf{u}}_{n,N+N_{cp}-1} = \sum_{j=1}^{N} \widetilde{w}_{n,N+N_{cp}-1}^{(j)} \mathbf{u}_{n,N+N_{cp}-1|N+N_{cp}-1}^{(j)} \tag{5.43}$$

$$\widehat{\phi}_{n,0:N+N_{cp}-1} = \sum_{j=1}^{N} \widetilde{w}_{n,N+N_{cp}-1}^{(j)} \phi_{n,0:N+N_{cp}-1}^{(j)} \tag{5.44}$$

$$\widehat{\epsilon} = \sum_{j=1}^{N} \widetilde{w}_{n,N+N_{cp}-1}^{(j)} \epsilon_{(N+N_{cp})/L}^{(j)} \tag{5.45}$$

The transmitted multicarrier signal is given by :

$$\widehat{\mathbf{s}}_{n,N+N_{cp}-1} = \widehat{\mathbf{u}}_{n,N+N_{cp}-1} + \mathbf{f}_{n,N+N_{cp}-1} \tag{5.46}$$

The proposed sequential Monte Carlo algorithm for Joint Signal, CFO and PHN Estimation, denoted by PSEM-SMC, is summed up in Table 5.1.

5.2.6 The Posterior Cramér-Rao Bound

As in the previous chapter, the posterior Cramér-Rao bound for this estimation problem is derived in order to have the lower bound of the optimal estimator. Unfortunately, the PCRB for the joint estimation of $\{\phi_{n,t}, \epsilon, \mathbf{u}_{n,t}\}$ is analytically intractable. Since the multicarrier signal is the main quantity of interest, we derive in this chapter the conditional PCRB of $\mathbf{u}_{n,t}$ where $\{\phi_{n,t}, \epsilon\}$ are assumed perfectly known. Under this assumption, the DSS model (5.24) becomes linear and Gaussian and the PCRB obtained at the end of each multicarrier symbol is equal to the covariance matrix $\mathbf{\Sigma}_{n,N+N_{cp}-1|N+N_{cp}-1}$ of the posterior $p.d.f.$ $p(\mathbf{u}_{n,N+N_{cp}-1} | \phi_{n,0:N+N_{cp}-1}, \epsilon, r_{n,0:N+N_{cp}-1})$ given by the kalman filter [Tichavsky *et al.* 98] :

Table 5.1 – *Stochastic EM per particle using Sequential Monte Carlo algorithm (PSEM-SMC)*

For $n = 0...EndofSymbols$
  Initialization
  For $k = 1...(N + N_{cp})/L$
    For $i = 0...L-1$
      For $j = 1...M$
        Update the predicted equations of the Kalman filter using (5.28)
        Sample $\phi^{(j)}_{(k-1)L+i}$ from the optimal importance function using (4.28)
        Update the filtered equations of the Kalman filter using (5.29)
        Evaluate the corresponding importance weights using (4.31)
      end for
      Normalize importance weights : $\widetilde{w}^{(j)}_{(k-1)L+i} = w^{(j)}_{(k-1)L+i} / \sum_{m=1}^{M} w^{(m)}_{(k-1)L+i}$
      If $N_{eff} < N_{seuil}$ **then**
        Resample particle trajectories using systematic resampling
      end if
    end for
    Compute the set of sufficient statistics $\Phi^{(j)}_k$ using (5.40) and (5.41)
    Compute the parameter update $\epsilon^{(j)}_k$ using (4.48)-(4.50)
  end for
  Evaluate the final MMSE smoothed estimates $\widehat{\mathbf{s}}_{n,N+N_{cp}-1}$, $\widehat{\phi}_{n,0:N+N_{cp}-1}$ and $\widehat{\epsilon}$ using respectively (5.46), (5.44) and (5.45)
end for

$$\text{PCRB} = \mathbb{E}\left[\frac{1}{N+N_{cp}}\text{tr}\left[\mathbf{\Sigma}_{n,N+N_{cp}-1|N+N_{cp}-1}\right]\right] \tag{5.47}$$

This PCRB is estimated using the Monte-Carlo method by recursively evaluating the predicted and filtered equations (5.28), (5.29) where $\{\phi_{n,t}, \epsilon\}$ are set to their true values.

5.3 Results

In order to show the validity of our approach, extensive simulations have been performed. In a first part, the case of PHN without CFO is considered in order to study the joint PHN and multicarrier signal estimation. Then in the last part, the performances of the PSEM-SMC are assessed when both CFO and PHN are present in multicarrier systems. In these two different cases, non-pilot and pilot aided version of the PSEM-SMC is studied in terms of mean square error (MSE) and bit error rate (BER) and compared to existing methods. Moreover, the BER performance of a multicarrier system in the absence of phase distortions and using the classical frequency domain MMSE equalizer defined in Eq. 1.13 is also depicted and denoted below by MMSE-FEQ.

With regard to the system parameters, 16-QAM modulation is assumed and we have chosen $N = 64$ subcarriers with a cyclic prefix of length $N_{cp} = 8$. A Rayleigh frequency selective channel with $L = 4$ paths and an uniform power delay profile, perfectly known by the receiver, has been generated for each multicarrier symbol. The proposed PSEM-SMC has been implemented with 100 particles, $\alpha = 1$ and $L = 2$.

5.3.1 Performances with PHN only (i.e. $\epsilon = 0$)

We first perform simulations with no CFO in order to study the joint PHN and multicarrier signal estimation. Two different system configurations are studied. The performances of the proposed PSEM-SMC without pilot subcarriers are compared to the non-pilot-aided variational scheme proposed in [Lin *et al.* 07] and to a CPE correction with a perfect knowledge of the CPE value corresponding to the ideal case of [Nikitopoulos *et al.* 05]. Finally, performances of the pilot-aided version of the proposed algorithm are analyzed and compared to the CPE and ICI correction method proposed in [Petrovic *et al.* 04b, Wu *et al.* 03a] and also to a perfect CPE correction scheme.

5.3.1.1 Non-Pilot-aided algorithm

Fig. 5.2 depicts the BER performance of the PSEM-SMC algorithm compared to the variational scheme proposed in [Lin *et al.* 07] and to a perfect CPE correction scheme. Since the multicarrier signal estimation is achieved by a Kalman filter, we can denote that, in a phase distortion-free context and by excluding the cyclic prefix in the received signal, the proposed algorithm leads to a time-domain MMSE equalizer. Time domain and frequency domain MMSE equalization are mathematically equivalent and result in the same performance [Hrycak *et al.* 06]. Consequently, the performance gain between the MMSE-FEQ and the PSEM-SMC without distortions clearly highlights the benefit of considering the additional information induced by the cyclic prefix. As depicted in this figure, the proposed PSEM-SMC outperforms conventional schemes whatever the PHN rate. Moreover, for $\beta T = 10^{-3}$, the PSEM-SMC curve is close to the optimal bound and outperforms the MMSE-FEQ without PHN.

The multicarrier signal estimation performance of the PSEM-SMC is shown in Fig. 5.3. The performance of the proposed estimator is compared to the Posterior Cramér-Rao bound (PCRB) of a multicarrier system without phase distortions derived in section 5.2.6. For a small phase noise rate βT, it can be seen that the proposed PSEM-SMC almost achieves the optimal performance without PHN given by the PCRB. Consequently, the proposed approximate optimal importance function for the PHN sampling leads to an efficient non-pilot-based algorithm for phase noise tracking as also illustrated in Fig 5.4 where the MSE of PHN estimate using the PSEM-SMC is depicted.

5.3.1.2 Pilot-aided algorithm

In this paragraph, we study the impact of the use of pilots subcarriers (*i.e.* $P \neq 0$) on the PSEM-SMC accuracy. Performances are compared to the CPE and ICI correction proposed in [Petrovic *et al.* 04b, Wu *et al.* 03a]. From figure 5.5, it can be seen that the PSEM-SMC outperforms the CPE+ICI correction scheme. By comparing performances in term of BER of the non-pilot and pilot-aided version of the PSEM-SMC (Fig. 5.2 vs Fig. 5.5), we can remark a slight improvement with the presence of pilot subcarriers in the JSPE-MPF performances, especially for $\beta T = 10^{-2}$ and large signal-to-noise ratio.

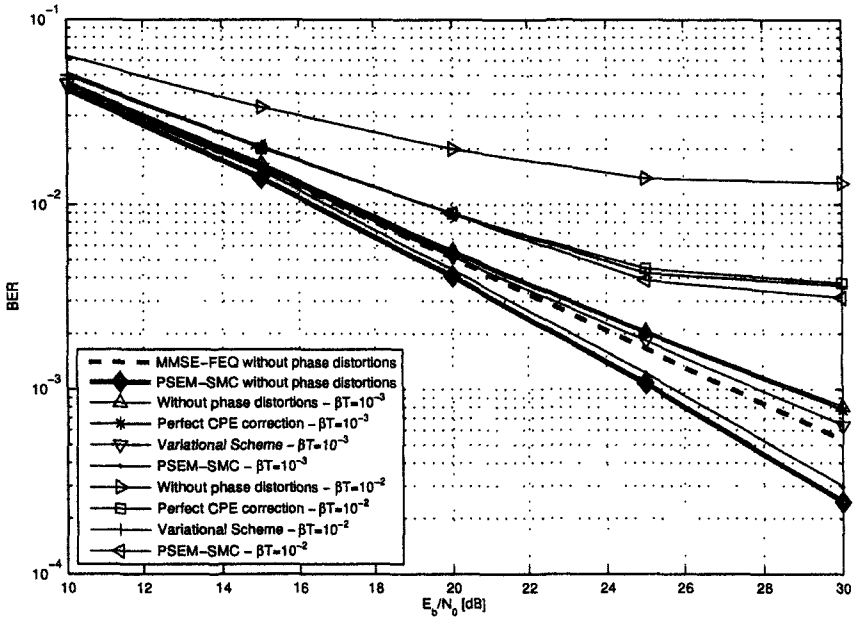

Figure 5.2 – *BER performance of the proposed PSEM-SMC vs E_b/N_0 for different PHN rates βT in an OFDM system ($N_g = 0$, $P = 0$, $\epsilon = 0$).*

Fig. 5.6 shows the impact of the number of null-subcarriers on PSEM-SMC performance in term of BER. From this figure, it can be denoted that its performance increases with the number of null-subcarriers, especially for large PHN rate. In fact, the proposed AR model of the multicarrier signal gives more prior information if the number of pilots or null-subcarriers increases and thus improves the robustness of the proposed estimator. In order to illustrate the gain obtained by the use of the proposed signal model, the MSE of both multicarrier signal and phase distortions obtained using the PSEM-SMC with or without our AR modeling is shown in Figs. 5.7 and 5.8. For the latter configuration (without AR model), $u_{n,t}$ and $u_{n,t-1}$ are assumed independent for $t = 1, ..., N-1$ but, like in the proposed AR model, the redundancy information given by the cyclic prefix is taken into account. Consequently, the DSS model described by (5.24) remains valid by considering both the vector $\boldsymbol{\zeta}_{n,t}$ in the transition matrix $\mathbf{A}_{n,t}$ defined as in (5.23) and the power of the driving noise defined as :

$$\sigma^2_{b_{n,t}} = \begin{cases} \frac{N-P-N_g}{N} & \text{for} \quad t = 0, ..., N-1 \\ 0 & \text{for} \quad t = N, ..., N+N_{cp}-1 \end{cases} \tag{5.48}$$

From these two figures, it can be observed that the MSE of both multicarrier signal and phase distortions is significantly improved by the use of the proposed AR model. Moreover, when the number of null-subcarriers increases, the state equation given by the proposed AR model gives more a priori information since the unknown multicarrier signal $u_{n,t}$ is more correlated with its past values (Eq. 5.12). As a consequence, the MSE of both the multicarrier signal and phase distortion estimate is improved, espe-

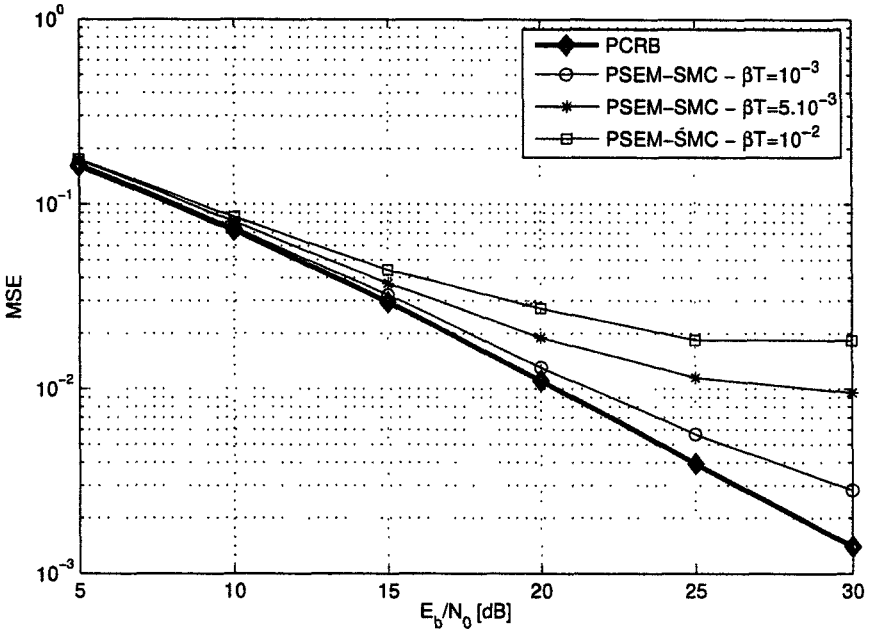

Figure 5.3 – *MSE of the multicarrier signal estimate vs E_b/N_0 for different PHN rates βT ($N_g = 0$, $P = 0$, $\epsilon = 0$).*

cially in severe PHN context. Indeed, for $E_b/N_0 = 30$ dB and $\beta T = 10^{-2}$, the MSE of the multicarrier signal obtained with the PSEM-SMC using the AR model is 1.1×10^{-2} and 7×10^{-3} for respectively $N_g = 0$ and $N_g = 8$.

Most of the works, related in multicarrier receivers, assume that the duration of cyclic prefix is longer than the maximum delay spread of the multipath channel. Under this assumption, the system is free from ISI. Nevertheless, the addition of the cyclic prefix reduces the bandwidth utilization efficiency. It is intuitive that the duration of cyclic prefix should be kept as short as possible. As a result, the duration of cyclic prefix may be shorter than the maximum delay of the multipath channel in some occasional cases. In these cases, ISI and ICI exist and, hence deteriorate the performance of the system. Here, performances of the proposed PSEM-SMC algorithm are assessed with insufficient cyclic prefix and with $P = 4$ pilots inserted in each OFDM symbol ($N_g = 0$).

Fig. 5.9 shows BER performance of the proposed scheme versus E_b/N_0 for two different number of channel paths and PHN rates. Whatever the PHN rate is, the PSEM-SMC significantly outperforms a classical OFDM receiver using MMSE equalizer and without PHN. Since the proposed scheme is implemented in time-domain and has the advantage of taking into account the redundancy of the cyclic prefix, the PSEM-SMC is more robust against the ISI than the classical MMSE equalizer.

5.3.2 Performances with both PHN and CFO

In the following simulations, the CFO term ϵ is generated from a uniform distribution in $[-0.4; 0.4]$, for each multicarrier symbol. Like in the previous case where only the

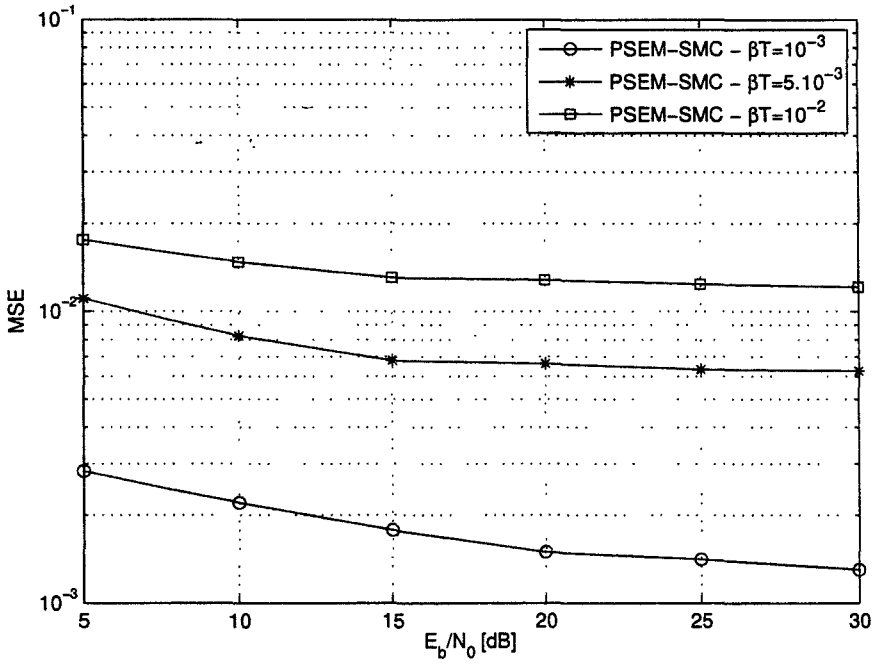

Figure 5.4 – *MSE of phase distortion estimate vs E_b/N_0 for different PHN rates βT ($N_g = 0$, $P = 0$, $\epsilon = 0$).*

PHN has been considered, both the non-pilot and pilot aided version of the PSEM-SMC is studied in this section. Nevertheless, a such CFO severely degrades the received signal (Section 1.2.2) and a perfect CPE correction obtain unsatisfactory BER performances (Fig. 1.18) due to a large ICI (Fig. 1.14). In this context, the existing algorithms based either or non-pilot and pilot subcarriers leads consequently to poor performances since a CPE correction scheme is required for the decision-directed method at the initialization step.

5.3.2.1 *Non-Pilot-aided algorithm*

Fig. 5.10 depicts the BER performance as a function of signal-to-noise ratio for different PHN rates in an OFDM system. As previously remarked, the perfect CPE correction scheme leads to poor system performances due to the large ICI caused by the phase distortions including a severe CFO. However, for $\beta T = 10^{-3}$, the proposed PSEM-SMC still outperforms the MMSE-FEQ without phase distortions. Moreover, even if the PHN rate increases, the PSEM-SMC still achieves accurate estimation.

Figs 5.11 and 5.12 depict the MSE respectively of the multicarrier signal and the phase distortions. Theses figures clearly highlights the robustness of the proposed SMC algorithm even if a severe CFO corrupts the received signal. It can be remarked that the MSE of the phase distortion estimate curves tend towards a minimum MSE threshold depending on PHN rate βT. Concerning the multicarrier signal estimation, the MSE curves shows that the PSEM-SMC still performs close to the PCRB for $\beta T = 10^{-3}$ and

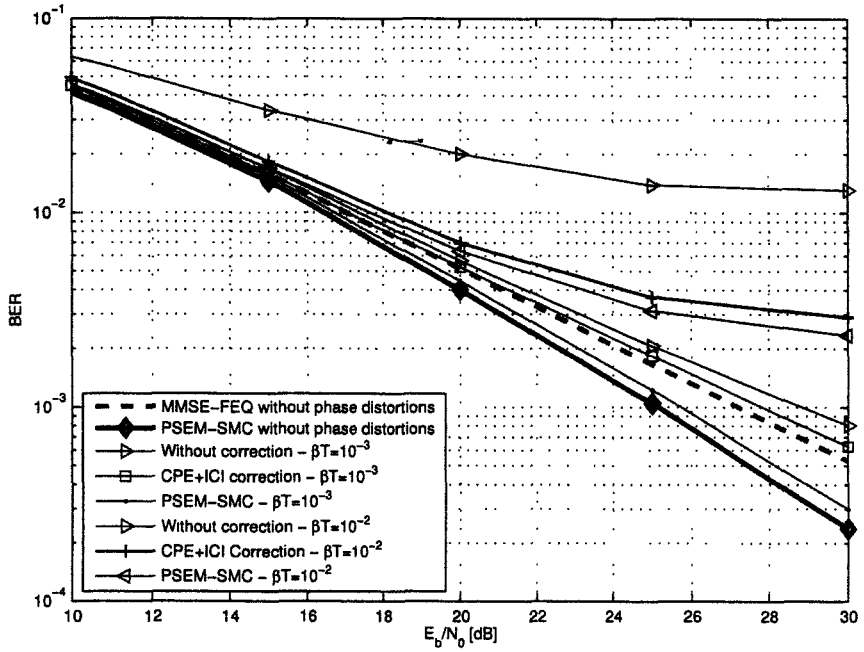

Figure 5.5 – *BER performance of the proposed PSEM-SMC vs E_b/N_0 for different PHN rates βT in an OFDM system ($N_g = 0$, $P = 4$, $\epsilon = 0$).*

the gap between the PCRB and the PSEM-SMC increases with PHN rate since the phase distortions get stronger, especially in the presence of a severe CFO.

By comparing these results with the ones obtained without CFO (Section 5.3.1.1), it can be denoted that the PSEM-SMC performance is only slightly degraded in the presence of CFO. These results illustrate the accuracy of both the CFO estimation using the proposed SEM per particle and the PHN sampling strategy using an approximate optimal importance function.

Finally, the BER performances of the PSEM-SMC for a full and half-loaded MC-CDMA system are respectively shown in Figs. 5.13-5.14. A time-synchronous downlink transmission is considered and Walsh codes are thus selected for their orthogonality property. From these figures, we still observe that, for $\beta T = 10^{-3}$, the PSEM-SMC slightly outperforms the MMSE-FEQ without phase distortions for both a full and a half-loaded system. This is principally due to the cyclic prefix additional information. Since the quantity of interest of the estimator is the time-domain multicarrier signal, the estimation accuracy of the proposed algorithm does not depend on the system load. As a consequence, the performance gap of the PSEM-SMC between a full and a half-loaded system is simply explained by the multiple access interference (MAI) induced by both the frequency selective channel and the residual phase distortion errors at the data detection stage.

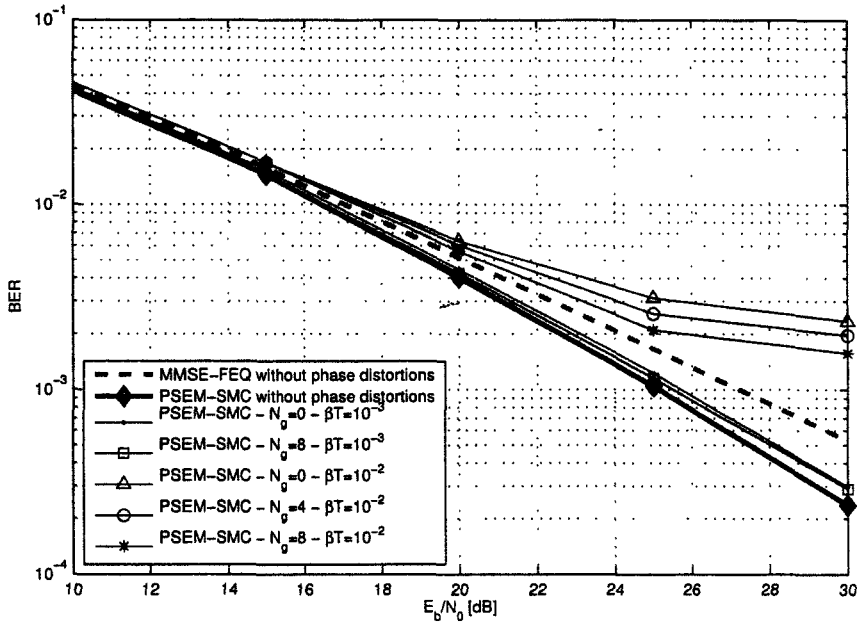

Figure 5.6 – *BER performance of the proposed PSEM-SMC vs E_b/N_0 for different PHN rates βT and number of null subcarriers N_g in an OFDM system ($P = 4, \epsilon = 0$).*

5.3.2.2 *Pilot-aided algorithm*

Now, the PSEM-SMC performance in the presence of pilot subcarriers is studied. Fig. 5.15 shows the BER performance of the proposed algorithm for different PHN rate and also different number of null-subcarriers with and without the AR modeling of the multicarrier signal. Firstly, only a slight performance degradation is observed when a severe CFO is considered in the multicarrier system. However compared to the non-pilot aided version, a significant improvement is made. Indeed, for $\beta T = 10^{-2}$ and $E_b/N_0 = 30$ dB, BER performances of the PSEM-SMC are 5×10^{-3} and 2×10^{-3} respectively for the non-pilot and pilot aided algorithm. Secondly, the BER performance and also the MSE of both the multicarrier signal and the phase distortions depicted in Figs. 5.16 and 5.17 are improved when the number of null subcarriers increases, especially with the proposed AR model. The performance gap with and without the AR model becomes more significant when the number of null subcarrier increases. These different results clearly highlight the benefit of the proposed AR model which makes more robust the proposed SMC estimator to severe phase distortions.

5.4 Conclusion

In this chapter, we address the challenging problem of data detection in multicarrier systems that suffer from the presence of phase noise (PHN) and carrier frequency offset (CFO). The originality of this work consists in using the Sequential Monte-Carlo methods for time domain processing of the nonlinear received signal in both nonpilot

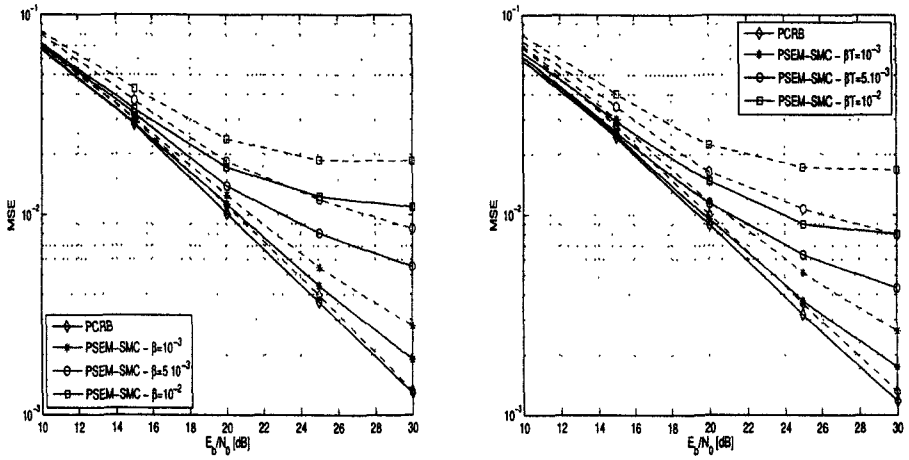

Figure 5.7 – *MSE of the multicarrier signal estimate using the proposed AR model (solid lines) and without AR model (dashed lines) vs E_b/N_0 for different PHN rates βT ($P = 4$, $\epsilon = 0$) with $N_g = 0$ (left) and $N_g = 8$ (right).*

and pilot-aided configurations. Moreover, an autoregressive modeling of a general multicarrier system in the presence of either pilot or null subcarriers is proposed in order to improve both the efficiency and the robustness of the proposed SMC filter.

Numerical simulations show that even with significant PHN rates and severe CFO, the PSEM-SMC achieves good performances both in terms of the phase distortion estimation and BER either in non-pilot or pilot aided context. In particular, it was found that, for a small PHN rate, the PSEM-SMC is more efficient in term of BER than a frequency MMSE equalizer in multicarrier system without phase distortions. Compared to existing schemes, the PSEM-SMC has better performance in the presence of only PHN. Moreover, unlike existing approaches, the PSEM-SMC enables to cope with severe CFO and number of channel paths higher than the cyclic prefix length.

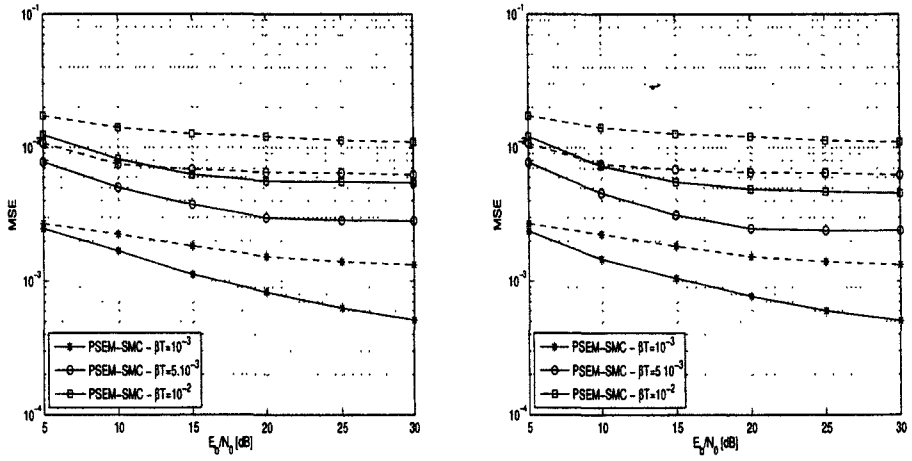

Figure 5.8 – *MSE of phase distortion estimate using the proposed AR model (solid lines) and without AR model (dashed lines) vs E_b/N_0 for different PHN rates βT ($P = 4$, $\epsilon = 0$) with $N_g = 0$ (left) and $N_g = 8$ (right).*

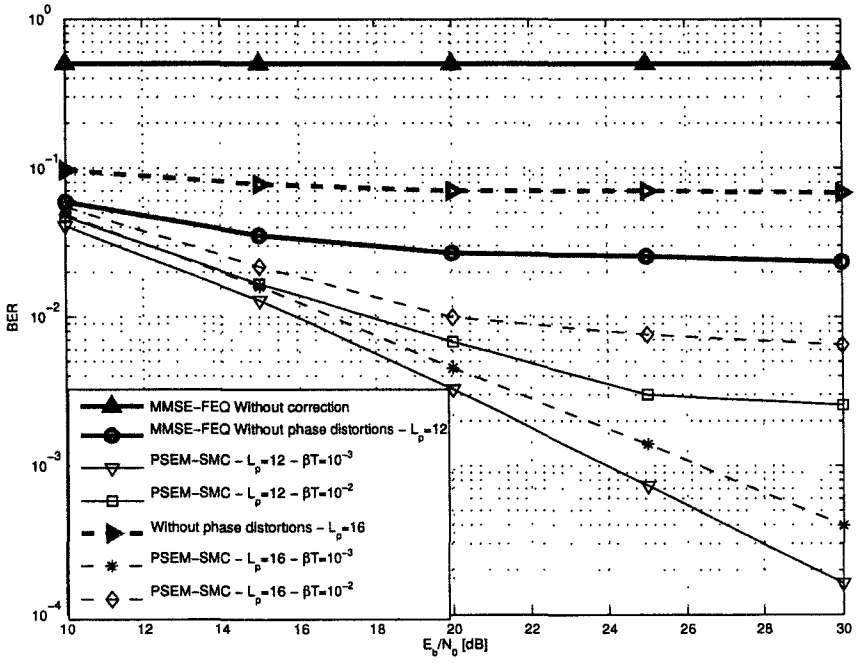

Figure 5.9 – *BER performance of the proposed PSEM-SMC when the number of channel paths is superior to the cyclic prefix length, i.e $L_p > N_{cp} = 8$, for different PHN rates βT in an OFDM system ($N_g = 0$, $P = 4, \epsilon = 0$).*

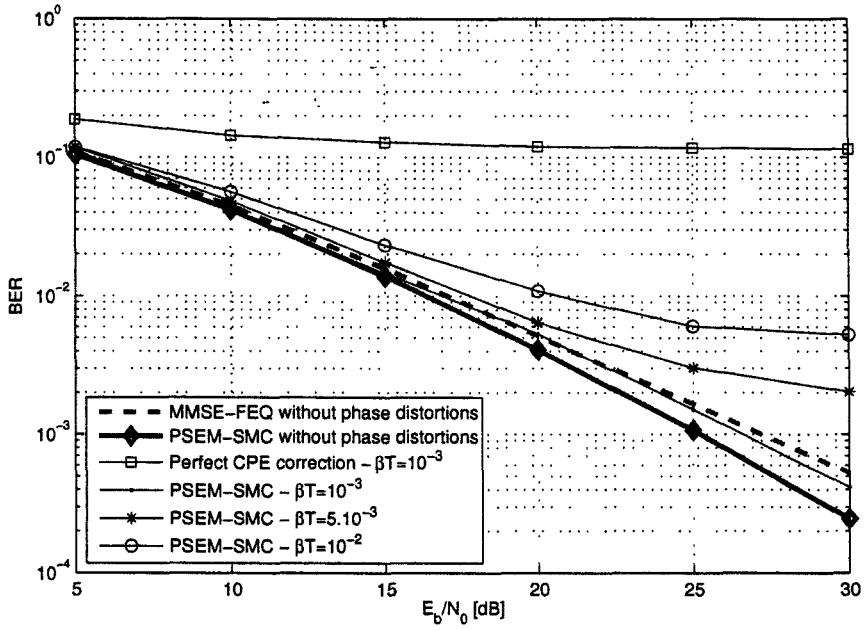

Figure 5.10 – *BER performance of the proposed PSEM-SMC vs E_b/N_0 for different PHN rates βT in an OFDM system ($N_g = 0$, $P = 0, \Delta_{CFO} = 0.4$).*

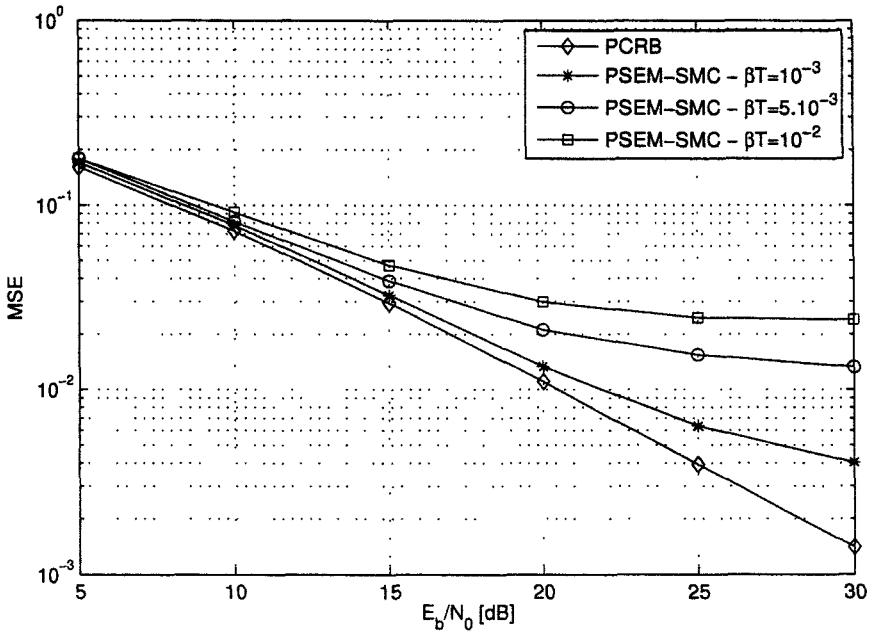

Figure 5.11 – *MSE of the multicarrier signal estimate vs E_b/N_0 for different PHN rates βT ($N_g = 0$, $P = 0$, $\Delta_{CFO} = 0.4$).*

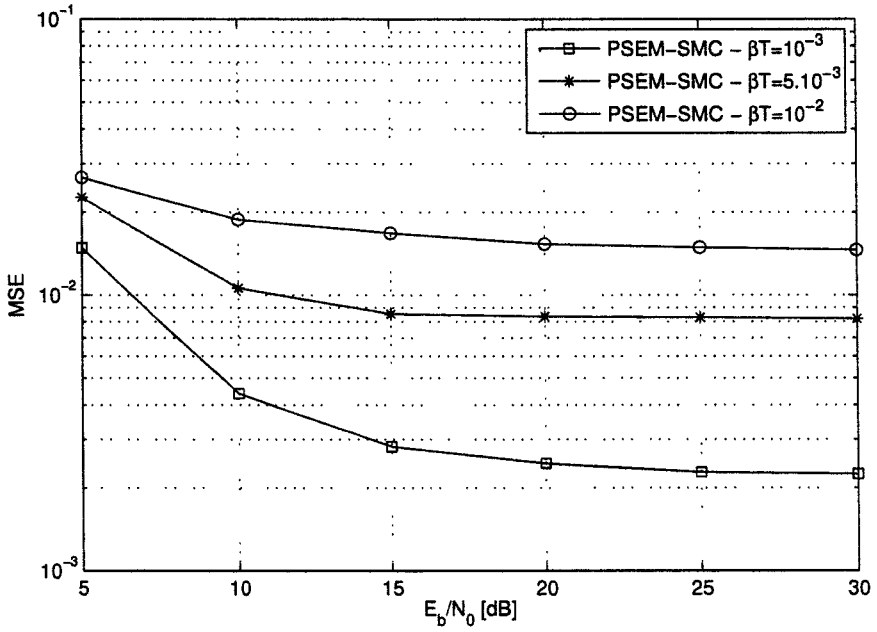

Figure 5.12 – *MSE of phase distortion estimate vs E_b/N_0 for different PHN rates βT ($N_g = 0$, $P = 0$, $\Delta_{CFO} = 0.4$).*

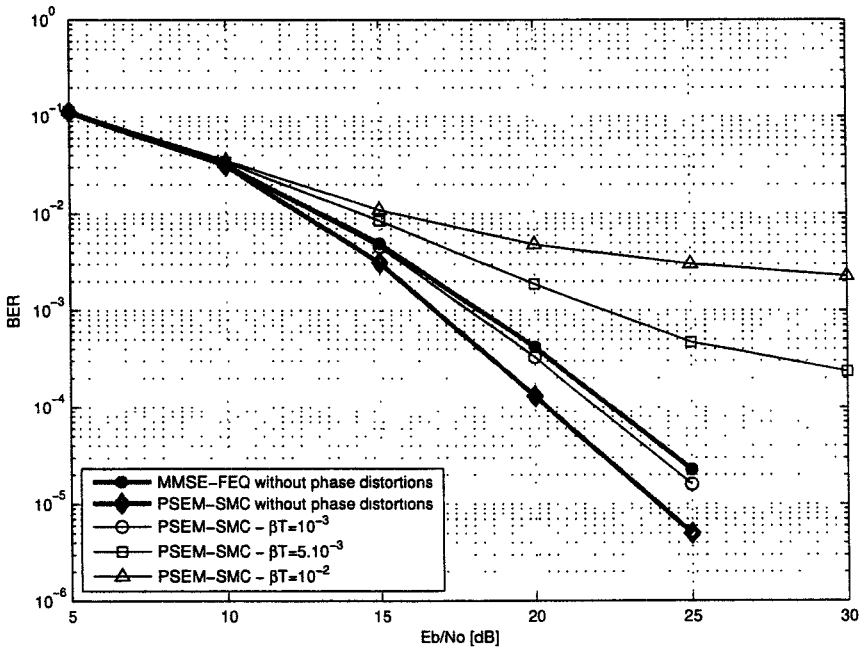

Figure 5.13 – *BER performance of the proposed PSEM-SMC vs E_b/N_0 for different PHN rates βT in a half-loaded MC-CDMA system ($N_g = 0$, $P = 0, \Delta_{CFO} = 0.4$).*

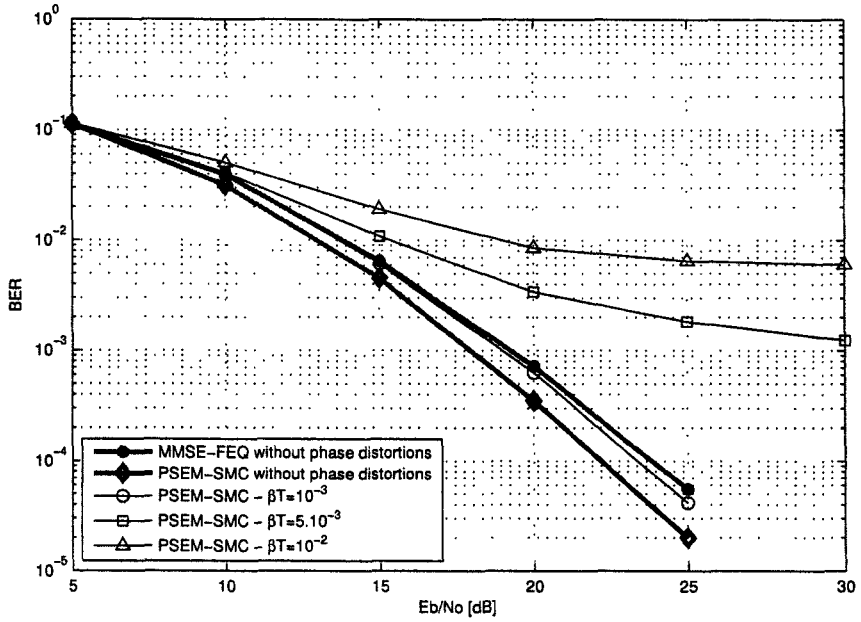

Figure 5.14 – *BER performance of the proposed PSEM-SMC vs E_b/N_0 for different PHN rates βT in a full-loaded MC-CDMA system ($N_g = 0$, $P = 0, \Delta_{CFO} = 0.4$).*

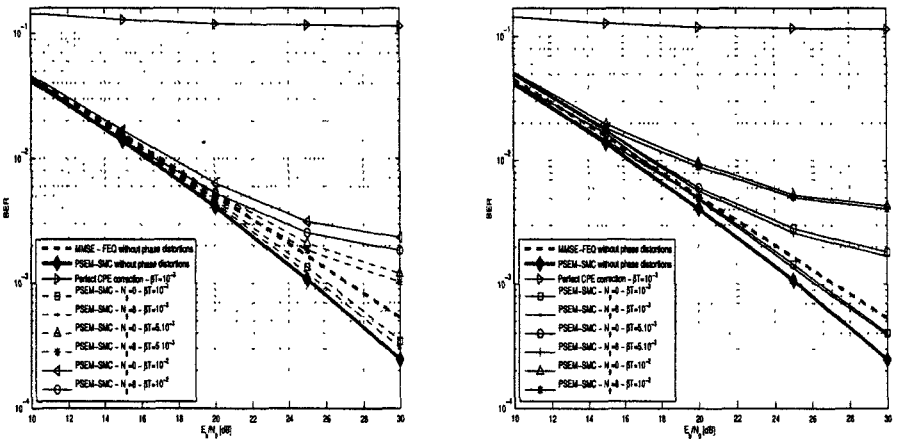

Figure 5.15 – *BER performance of the proposed PSEM-SMC vs E_b/N_0 for different PHN rates βT and number of null-subcarriers ($P = 4$, $\epsilon = 0.4$) using the proposed AR model (left) and without the AR model (right).*

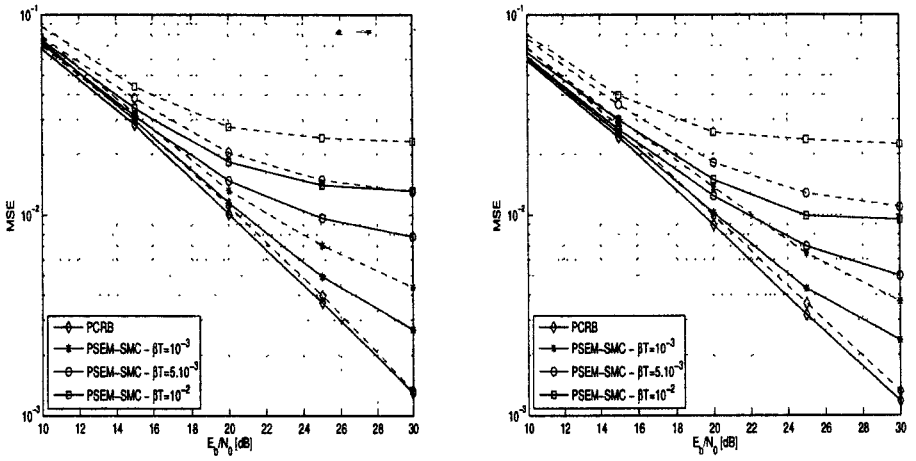

Figure 5.16 – *MSE of the multicarrier signal estimate using the proposed AR model (solid lines) and without AR model (dashed lines) vs E_b/N_0 for different PHN rates βT ($N_g = 0$, $P = 4$, $\epsilon = 0.4$) with $N_g = 0$ (left) and $N_g = 8$ (right).*

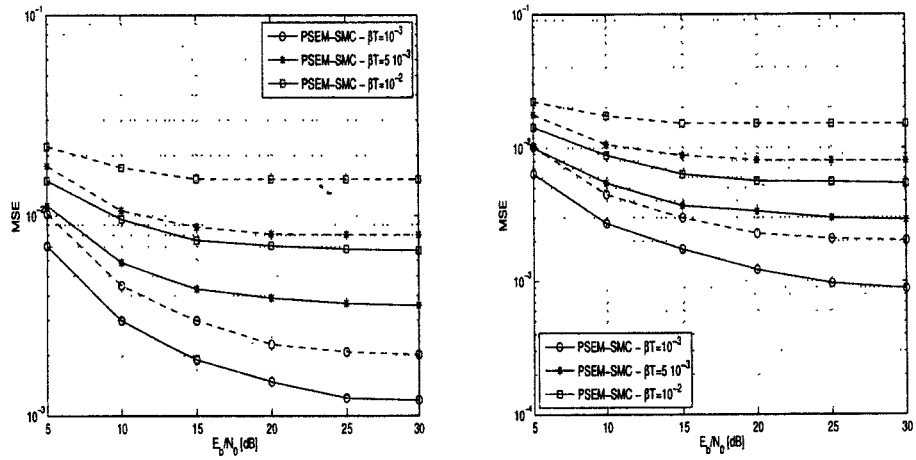

Figure 5.17 – *MSE of phase distortion estimate using the proposed AR model (solid lines) and without AR model (dashed lines) vs E_b/N_0 for different PHN rates βT ($N_g = 0$, $P = 4$, $\epsilon = 0.4$) with $N_g = 0$ (left) and $N_g = 8$ (right).*

GENERAL CONCLUSION

OFDM systems have been standardized in many recent high-data rate applications due to its ability to combat multipath effects and make better use of the system available bandwidth. The benefits and success of OFDM modulation on the one hand, and the flexibility offered by spread spectrum techniques on the other hand, have, since 1993, motivated many researchers to investigate the combination of both techniques, known as multicarrier CDMA techniques.

Unfortunately, multicarrier systems are very sensitive to phase noise (PHN) and carrier frequency offset (CFO) caused by oscillator instabilities. After a brief statistical description of phase distortions, their effects have been analyzed in Chapter 1 with a closed form expression of the SINR, with which, system behavior can be clearly judged for any phase distortions levels. These results clearly highlight the severe performance limitations induced by phase distortions.

In order to have a reliable multicarrier system, a phase distortion compensation scheme is consequently required. Existing schemes dealing with phase impairments in OFDM systems are briefly reviewed in Chapter 1. These algorithms are based on two consecutive steps which consists in estimating respectively the channel impulse response using a training multicarrier symbol and then the transmitted data symbols. However as seen in Section 1.3, these approaches suffer from several drawbacks. On the one hand, for the channel estimation problem, both the AWGN and PHN powers are assumed known to the receiver in existing schemes, which is not a realistic assumption. On the other hand, for data detection, efficient pilot or non-pilot aided algorithms are based on decision-directed scheme and thus lead to poor performance for high phase distortions levels. Moreover, these data estimator are only derived for OFDM systems.

In this thesis, we have focused on the design of an efficient and general multicarrier receiver based on Bayesian inference in order to improve the performance and reliability of multicarrier transmission in the presence of phase distortions. More precisely, in order to solve this task, we propose to use sequential Monte Carlo methods which are a set of powerful simulation-based algorithms to perform optimal state estimation in nonlinear non-Gaussian state space models (Chapter 2).

Nevertheless, although SMC methods are well adapted to problems of estimating a sequence of potentially quickly varying distributions whose dimension is increasing over time, parameter estimation using particle filtering is still a major issue. Since the estimation problem we deal with consists in estimating both dynamic states and static parameters, this task has to be considered with care. In Chapter 3, we have

briefly reviewed the existing scheme for both off-line and on-line parameter estimation using sequential Monte-Carlo methods. Given existing scheme limitations and our application requirements, a new on-line parameter estimation method is proposed and is denoted in this manuscript by PSEM-SMC. The main idea is to benefit from the diversity induced by the different particle trajectories and also from the stability of a deterministic method for parameter recursive update. More precisely, we propose to associate a stochastic expectation maximization (SEM) algorithm to each particle trajectory. Numerical simulations have been provided in two different models. The results show the efficiency of the proposed scheme in comparison to the on-line EM (OEM-SMC). We have seen that the PSEM-SMC outperforms the OEM-SMC in term of convergence rate principally due to an increased diversity in the search space of parameters. Moreover, this proposed scheme offers good stability in time due to the use of a forgetting factor in the deterministic method. The case of a non-stationary hidden Markov process has also been treated. Indeed, since the phase noise is a non-stationary process, this case has been studied. In this context, we have proposed an adaptation of both OEM-SMC and PSEM-SMC. Numerical results illustrate that the PSEM-SMC still converges more rapidly than the OEM-SMC.

By using the packet structure of existing standards like Hiperlan/2 or IEEE802.11a, we propose a multicarrier receiver based on SMC methodology which consists in two separate estimation problems. More precisely, we firstly propose in Chapter 4 an OFDM channel estimator in the presence of phase distortions from a training multicarrier symbol which also deals with the unknowledge of both PHN and AWGN powers. Different strategies for off-line and on-line estimation based on Monte Carlo methods have been introduced. 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 a SMC filter with two different strategies for parameter estimation: the proposed SEM per particle (PSEM-SMC) and the on-line EM (OEM-SMC). The efficiency of the proposed algorithms for multicarrier channel estimation in the presence of CFO and PHN, when both PHN and AWGN powers are assumed unknown, have been illustrated through numerical simulations. Because of both the computational cost and the high memory requirements of the SAEM, the on-line SMC methods and particularly the PSEM-SMC consequently offers a good compromise between performances and overall complexity.

Finally, in Chapter 5, a joint multicarrier signal, CFO and PHN estimator has been proposed. A SMC based algorithm is derived for time domain processing by combining the Rao-Blackwellization technique and an approximate optimal importance function. Moreover, an autoregressive modeling of a general multicarrier system in the presence of either pilot or null subcarriers is proposed in order to improve both the efficiency and the robustness of the proposed SMC filter. The proposed algorithm has also the great advantage of not requiring any modifications for its implementation in different system configurations. Numerical simulations show that even with significant PHN rates and severe CFO, the PSEM-SMC achieves good performances both in terms of

the phase distortion estimation and BER either in non-pilot or pilot aided context. Moreover, unlike existing approaches, the PSEM-SMC enables to cope with severe CFO and number of channel paths higher than the cyclic prefix length.

FUTURE WORK

Some directions for future research are now discussed. In this thesis, a multicarrier receiver based on SMC methods have been proposed to deal with phase impairments. On the one hand, the proposed SMC receiver could be extended to the multiple-input and multiple output (MIMO) case of multicarrier systems which consists in using multiple antennas at both the transmitter and receiver in order to improve communication performance. On the other hand, instead of dealing separately with the channel estimation and then with the data detection, the application of SMC methods to the joint estimation of channel, multicarrier signal and phase distortions could be studied. In this context, a doubly-selective (time and frequency selective) channel with a time coherence smaller than the duration of the multicarrier symbol could be considered in order to tackle the problem of high mobility in multicarrier systems. Moreover, the problem of time synchronization could be taken into account. Even if this estimation problem seems to be extremely difficult to deal with, the time domain processing, as proposed in this thesis with the autoregressive modeling of the multicarrier signal, using sequential Bayesian methods could be studied. However, in this case, the number of unknown states would be higher than the one studied in this thesis. As a consequence, we could resort to combination of several techniques such as particle filtering and Markov chain Monte-Carlo (MCMC) methods which are not sequential but they are more robust than particle filter to address the problem of estimating a state of high dimension.

Finally, concerning the parameter estimation problem in particle filtering, a theorical convergence analysis of the proposed on-line parameter estimation could be studied. On the other hand, a more general parameter estimation scheme including both the on-line EM and the proposed SEM per particle as special cases could also be investigated.

CONCLUSION GÉNÉRALE EN FRANÇAIS

LA technique OFDM a été standardisée dans de nombreuses applications récentes à haut débit grâce notamment à sa capacité à combattre les effets néfastes des multitrajets mais également à utiliser efficacement la bande passante disponible pour le système. Les avantages et le succès de la modulation OFDM d'une part, et la flexibilité offerte par les techniques à étalement de spectre d'autre part, ont motivé depuis 1993 de nombreux chercheurs à étudier la combinaison de ces deux techniques, connue sous le nom de technique CDMA multiporteuses.

Malheureusement, ces systèmes multiporteuses sont extrêmement sensibles aux distorsions de phase, comme le *bruit de phase* et le *décalage fréquentiel de la porteuse*, engendrées par l'instabilité des oscillateurs locaux présent dans la chaîne de transmission. Après une brève description statistique de ces phénomènes perturbateurs, leur impact sur les performances du système ont été analysé dans le Chapitre 1 avec la dérivation exacte du rapport signal à bruit plus interférence (SINR) avec lequel le comportement du système peut être objectivement jugé pour différents niveaux de distorsions de phase. Les résultats obtenus dans ce premier chapitre mettent clairement en évidence les sévères limitations des performances d'un système multiporteuses en présence de distorsions de phase.

Dans le soucis d'avoir un système multiporteuses fiable, une technique de compensation de ces perturbations est en conséquence indispensable. Les méthodes existantes traitant des distorsions de phase dans les systèmes OFDM sont décrites dans le Chapitre 1. Ces algorithmes sont basés sur deux étapes consécutives consistant respectivement à estimer la réponse impulsionnelle du canal grâce à une séquence d'apprentissage et ensuite à estimer les données transmises. Cependant, comme remarqué dans la Section 1.3, ces approches souffrent de plusieurs inconvénients. D'une part, pour le problème de l'estimation du canal de propagation, les puissances du bruit additif Gaussien et du bruit de phase sont supposées parfaitement connues du récepteur, ce qui ne s'avère pas être une hypothèse très réaliste. D'autre part, pour la détection des données, les algorithmes existants, aidés ou non par la présence de porteuse pilotes, et fournissant les meilleures performances sont basés sur un schéma consistant à effectuer une tentative de décision sur le signal toujours corrompu par les imperfections de phase. En conséquence, les performances de ces méthodes baissent drastiquement pour des niveaux sévères de distorsions de phase. De plus, ces estimateurs de données ne sont dérivés que dans le cas d'un système OFDM.

Dans cette thèse, nous nous sommes donc intéressés sur la conception d'un récepteur basé sur l'inférence Bayésienne qui soit général et surtout efficace pour lutter

contre les effets néfastes de ces distorsions de phase. Plus précisément, nous proposons d'utiliser les méthodes séquentielles de Monte-Carlo qui sont de puissants algorithmes de simulation stochastique destinés à l'estimation optimale de processus dans un modèle d'état non linéaire et non Gaussien (Chapitre 2).

Néanmoins, bien que les méthodes séquentielle de Monte-Carlo sont bien adaptées aux problèmes d'estimation de séquence variant rapidement dans le temps dont la dimension augmente également avec le temps, l'estimation de paramètres du modèle par filtrage particulaire est encore un problème majeur. Comme le cas traité dans cette thèse consiste à la fois en l'estimation de variables dynamiques et statiques, ce problème doit être considéré avec attention. Dans le Chapitre 3, nous décrivons les principales solutions permettant l'estimation soit "off-line" ou "on-line" de ces paramètres en utilisant les méthodes séquentielles de Monte-Carlo. Etant donné les limitations de ces solutions et surtout les besoins spécifiques concernant notre application, une nouvelle méthode d'estimation "on-line" des paramètres est proposée et est dénotée dans le manuscrit par PSEM-SMC. L'idée principale de ce nouvel algorithme est de profiter pour la mise à jour des paramètres d'une part de la diversité des différentes trajectoires des particules et d'autre part de la stabilité des méthodes déterministes. Plus précisément, nous proposons d'associer un algorithme d'espérance maximisation stochastique (SEM) à chaque particule. Dans ce Chapitre 3, des simulations numériques ont été fournies dans deux différents modèles. Les résultats montrent l'efficacité du schéma proposé en comparaison avec l'algorithme "on-line" d'espérance maximisation (OEM-SMC). Nous avons vu que l'algorithme PSEM-SMC a de meilleures performances en terme de vitesse de convergence, cela s'expliquant principalement par le fait que dans le PSEM-SMC une plus grande diversité dans l'espace de recherche des paramètres est présente. De plus, la méthode proposée offre une bonne stabilité dans le temps grâce au facteur d'oubli utilisé dans la méthode déterministe par particule. Le cas d'un processus de Markov caché non stationnaire est également étudié car le processus correspondant au bruit de phase est lui-même non stationnaire. Dans ce contexte, nous avons proposé quelques modifications pour à la fois le PSEM-SMC et l'OEM-SMC. Les résultats de simulation amènent les mêmes conclusions que dans le cas stationnaire.

En utilisant la structure d'une trame des standards existants comme l'Hiperlan/2 et l'IEEE 802.11a, nous proposons ainsi un récepteur pour systèmes multiporteuses utilisant le principe des méthodes séquentielles de Monte-Carlo et consistant en deux problèmes distincts d'estimation. Plus précisément, dans le Chapitre 4, nous proposons un estimateur du canal OFDM en présence à la fois de bruit de phase et de décalage fréquentiel de la porteuse en s'appuyant sur une séquence d'apprentissage. De plus, contrairement aux méthodes existantes, nous supposons que les puissances à la fois du bruit additif Gaussien et du bruit de phase sont inconnues du récepteur. Différentes stratégies basées sur les méthodes de Monte-Carlo pour l'estimation ("off-line" et "on-line") des paramètres sont proposées. D'une part, dans le cas "off-line", nous proposons une combinaison du SAEM et du JCPCE pour effectuer l'estimation de tous les états d'intérêt. D'autre part, dans le cas "on-line", on propose un filtre particulaire avec

deux différentes méthodes pour l'estimation des paramètres : notre proposition de SEM par particule (PSEM-SMC) et l'"on-line" d'espérance maximisation (OEM-SMC). L'efficacité de ces différents algorithmes est illustrée à travers de nombreuses simulations numériques. A cause du coût de calcul et des grands besoins de mémoire du SAEM, les méthodes "on-line" utilisant les méthodes séquentielles de Monte-Carlo, et tout particulièrement le PSEM-SMC, offre un bon compromis entre performances et complexité globale.

Finalement, dans le Chapitre 5, un estimateur conjoint du signal multiporteuses, du bruit de phase et du décalage fréquentiel de la porteuse est proposé. Un filtre particulaire est dérivé pour le traitement du signal dans le domaine temporel en combinant la technique de marginalisation ("Rao-Blackwellization") et une approximation de la fonction optimale d'importance grâce à une linéarisation local du model d'état. De plus, dans le but d'améliorer à la fois l'efficacité et la robustesse de notre filtre Bayésien, un modèle autorégressif général concernant l'évolution temporel du signal multiporteuses, dans le cas où des porteuses soit nulles ou pilotes sont présentes, est proposé. L'algorithme proposé a le grand avantage de ne nécessiter aucune modification quelque soit le système multiporteuses considéré. Les performances du filtre ont été étudiées dans de nombreuses configurations différentes. Les résultats montrent, que même en présence d'un niveau significatif de bruit de phase et d'un sévère décalage fréquentiel de la porteuse, le PSEM-SMC permet d'obtenir de très bons résultats en terme de taux d'erreur binaire. De plus, contrairement aux méthodes existantes, le filtre proposé permet de lutter contre un sévère décalage fréquentiel de la porteuse mais aussi contre un canal dont le nombre de trajets est supérieur à la longueur du préfixe cyclique.

PERSPECTIVES

Quelques perspectives concernant de futures recherches sont maintenant discutées. Dans cette thèse, un récepteur complet pour système multiporteuses basé sur les méthodes séquentielles de Monte-Carlo a été proposé pour lutter contre les imperfections de phase. Tout d'abord, les algorithmes proposés dans cette thèse pourraient être étendus au cas "multiple input multiple output" (MIMO) des systèmes multiporteuses qui consiste à utiliser plusieurs antennes à l'émission et à la réception dans le but d'améliorer les performances de transmission. Ensuite, au lieu de traiter indépendamment l'estimation du canal et ensuite l'estimation des données, il pourrait être envisageable d'étudier la conception d'un filtre Bayésien permettant l'estimation conjointe de toutes ces entités. Dans ce contexte, on pourrait alors considérer le problème de la grande mobilité et ainsi considérer un canal doublement sélectif (en temps et en fréquence) avec donc un temps de cohérence inférieur à la durée d'un symbole multiporteuses. De plus, le problème de désynchronisation pourrait également être pris en compte. Même si le problème s'avère être extrêmement difficile aux premiers abords, le traitement dans le domaine temporel du signal, comme proposé dans cette

thèse avec le modèle autorégressif, par des méthodes Bayésienne séquentielles reste à étudier. Cependant, le nombre d'inconnus serait alors beaucoup plus élevé que dans notre étude et il faudrait certainement avoir recours à une combinaison de techniques comme par exemple l'association de méthodes particulaires et de méthodes de Monte-Carlo par chaines de Markov (MCMC). Ces dernières étant plus robustes que le filtrage particulaire face à un problème d'estimation où l'état caché est de grande dimension.

Finalement, concernant le problème d'estimation "on-line" des paramètres par filtrage particulaire, une analyse théorique des propriétés de la méthode proposée qui consiste à associer un SEM par particule pourrait être étudiée. D'autre part, un estimateur de paramètres plus général englobant à la fois notre méthode et l'algorithme "on-line" d'espérance maximisation comme cas spéciaux pourrait être analysé.

LIST OF PUBLICATIONS

Journal papers :

⋆ Septier, F., Delignon, Y., Menhaj-Rivenq, A. & Garnier, C. *Monte Carlo Methods for Channel, Phase Noise and Frequency Offset Estimation with Unknown Noise Variances in OFDM Systems*. Accepted to IEEE Transactions on Signal Processing. To appear in 2008.

⋆ Septier, F., Delignon, Y., Menhaj-Rivenq, A. & Garnier, C. *Nonpilot-Aided Sequential Monte-Carlo Method to Joint Signal, Phase Noise and Frequency Offset Estimation in Multicarrier Systems*. Eurasip Journal on Advances in Signal Processing, Vol. 2008, http://www.hindawi.com/GetArticle.aspx?doi=10.1155/2008/612929.

Conference papers :

⋆ Septier, F., Delignon, Y., Menhaj-Rivenq, A. & Garnier, C. *OFDM channel estimation in the presence of phase noise and frequency offset by particle filtering*. Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, 2007. (IEEE ICASSP' 07), Honolulu, Hawaii, 15-20 April 2007, Vol. 3, pp289-292.

⋆ Septier, F., Delignon, Y., Menhaj-Rivenq, A. & Garnier, C. *Particle filtering with hybrid importance function for joint symbol detection and phase tracking*. Proc. IEEE 7th Workshop onSignal Processing Advances in Wireless Communications, 2006. (IEEE SPAWC' 06), Cannes, France, 2-5 July 2006, pp1-5.

⋆ Septier, F., Delignon, Y., Menhaj-Rivenq, A. & Garnier, C. *Traitement du bruit de phase par filtrage particulaire en communication numerique*. Proc. Traitement et Analyse de l'Information : Methodes et Applications, 2005. (TAIMA), Hammamet, 26 Sep.-1 Oct 2005, pp. 9-18.

APPENDICES

A

CONTENTS

A.1 Monte Carlo approximation of the posterior distribution for the SAEM

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

$$p(\boldsymbol{\phi}, \mathbf{h}|\mathbf{r}, \lambda_{i-1}) \propto p(\mathbf{h}|\boldsymbol{\phi}, \mathbf{r}, \lambda_{i-1}) p(\boldsymbol{\phi}|\mathbf{r}, \lambda_{i-1}) \tag{A.1}$$

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

A.1.1 Importance sampling of the phase distortions

The marginal posterior distribution $p(\boldsymbol{\phi}|\mathbf{r}, \lambda_{i-1})$ is approximated by Monte Carlo algorithm as :

$$\widehat{p}(\boldsymbol{\phi}|\mathbf{r}, \lambda_{i-1}) = \sum_{j=1}^{M} \widetilde{w}_i^{(j)} \delta(\boldsymbol{\phi} - \boldsymbol{\phi}_i^{(j)}) \tag{A.2}$$

The phase distortions values are sampled from its prior distribution,

$$\pi(\boldsymbol{\phi}|\mathbf{r}, \lambda_{i-1}) = p(\boldsymbol{\phi}|\lambda_{i-1}) \tag{A.3}$$

where $p(\boldsymbol{\phi}|\lambda_{i-1})$ is obtained from the state equation (4.3) and is equal to

$$p(\boldsymbol{\phi}|\lambda_{i-1}) = \mathcal{N}(\boldsymbol{\phi}; \epsilon_{i-1}\boldsymbol{\mu}, \sigma_{v,i-1}^2 \boldsymbol{\Theta}) \tag{A.4}$$

with $\boldsymbol{\mu} = \begin{bmatrix} 2\pi(N-1)/N & \cdots & 2\pi/N & 0 \end{bmatrix}^T$ and

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

Using (2.40), the unnormalized importance weights are given by:

$$w_i^{(j)} = \frac{p(\boldsymbol{\phi}_i^{(j)}|\mathbf{r}, \lambda_{i-1})}{\pi(\boldsymbol{\phi}_i^{(j)}|\mathbf{r}, \lambda_{i-1})} \tag{A.5}$$

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

$$w_i^{(j)} \propto p(\mathbf{r}|\boldsymbol{\phi}_i^{(j)}, \lambda_{i-1}) \tag{A.6}$$

The distribution $p(\mathbf{r}|\boldsymbol{\phi}_i^{(j)}, \lambda_{i-1})$ is obtained by marginalizing $p(\mathbf{r}, \mathbf{h}|\boldsymbol{\phi}_i^{(j)}, \lambda_{i-1})$. Be-

cause $\mathbf{h}$ is independent of $\boldsymbol{\phi}$ and λ,

$$p(\mathbf{r},\mathbf{h}|\boldsymbol{\phi}_i^{(j)},\lambda_{i-1}) = p(\mathbf{r}|\mathbf{h},\boldsymbol{\phi}_i^{(j)},\lambda_{i-1})p(\mathbf{h}) \tag{A.7}$$

where

$$p(\mathbf{h}) = \mathcal{N}_c\left(\mathbf{0}_{L\times 1},\frac{1}{L}\mathbf{I}_L\right) \tag{A.8}$$

and the likelihood distribution is given by:

$$p(\mathbf{r}|\mathbf{h},\boldsymbol{\phi}_i^{(j)},\lambda_{n-1}) = \mathcal{N}_c\left(\mathbf{r};\mathbf{C}_i^{(j)}\mathbf{h},\sigma_{b,i-1}^2\mathbf{I}_N\right) \tag{A.9}$$

with

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

As a consequence, we obtain :

$$w_i^{(j)} \propto \mathcal{N}_c(\mathbf{r};\mathbf{0}_{N\times 1},\sigma_{b,i-1}^2\mathbf{I}_N + \frac{\mathbf{C}_i^{(j)}\left(\mathbf{C}_i^{(j)}\right)^H}{L}) \tag{A.10}$$

A.1.2 Posterior distribution of the channel

Since the CIR $\mathbf{h}$ has a circular gaussian *a priori p.d.f.*, the posterior *p.d.f.* $p(\mathbf{h}|\boldsymbol{\phi}_i^{(j)},\mathbf{r},\lambda_{i-1})$ of the j-th particle can be written as :

$$p(\mathbf{h}|\boldsymbol{\phi}_i^{(j)},\mathbf{r},\lambda_{i-1}) \propto p(\mathbf{r}|\mathbf{h},\boldsymbol{\phi}_i^{(j)},\lambda_{i-1})p(\mathbf{h}) \tag{A.11}$$

where $p(\mathbf{h})$ and $p(\mathbf{r}|\mathbf{h},\boldsymbol{\phi}_i^{(j)},\lambda_{i-1})$ are respectively given by (A.8) and (A.9)

Consequently, the *a posteriori p.d.f.* (A.11) of $\mathbf{h}$ is :

$$p(\mathbf{h}|\boldsymbol{\phi}_i^{(j)},\mathbf{r},\lambda_{i-1}) = \mathcal{N}_c\left(\mathbf{h};\mathbf{h}_i^{(j)},\boldsymbol{\Sigma}_i^{(j)}\right) \tag{A.12}$$

with :

$$\mathbf{h}_i^{(j)} = \frac{1}{L}\left(\mathbf{C}_i^{(j)}\right)^H\left(\frac{1}{L}\mathbf{C}_i^{(j)}\left(\mathbf{C}_i^{(j)}\right)^H + \sigma_{b,i-1}^2\mathbf{I}_N\right)^{-1}\mathbf{r} \tag{A.13}$$

and,

$$\boldsymbol{\Sigma}_i^{(j)} = \frac{1}{L}\mathbf{I}_L - \frac{1}{L^2}\left(\mathbf{C}_i^{(j)}\right)^H\left(\frac{1}{L}\mathbf{C}_i^{(j)}\left(\mathbf{C}_i^{(j)}\right)^H + \sigma_{b,i-1}^2\mathbf{I}_N\right)^{-1}\mathbf{C}_i^{(j)} \tag{A.14}$$

A.2 Derivation of the PHN optimal importance function for the channel estimation problem

The optimal importance function for PHN sampling in the proposed particle filter requires the derivation of the *p.d.f.* defined in (5.30). This *p.d.f.* can be rewritten as :

$$p(\phi_t|\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},r_{0:t}) = \frac{p(r_t|\phi_t,\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},r_{0:t-1})p(\phi_t|\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)})}{p(r_t|\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},r_{0:t-1})} \quad \text{(A.15)}$$

with

$$p(\phi_t|\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)}) = \begin{cases} \mathcal{N}(\phi_0;0,\sigma_{v,\lfloor t/L\rfloor}^2) & \text{if } t=0 \\ \mathcal{N}(\phi_t;\phi_{t-1}^{(j)}+2\pi\epsilon_{\lfloor t/L\rfloor}^{(j)}/N,\sigma_{v,\lfloor t/L\rfloor}^2) & \text{otherwise} \end{cases} \quad \text{(A.16)}$$

and,

$$p(r_t|\phi_t,\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},r_{0:t-1}) = \int p(r_t|\mathbf{h},\phi_t,\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},r_{0:t-1})p(\mathbf{h}|\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},r_{0:t-1})d\mathbf{h} \quad \text{(A.17)}$$

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

$$p(r_t|\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},r_{0:t-1}) = \int p(r_t|\phi_t,\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},,r_{0:t-1})p(\phi_t|\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)})d\phi_t \quad \text{(A.18)}$$

The summand in (A.17) is the product of the likelihood *p.d.f.* and the posterior *p.d.f.* of the CIR which are defined respectively by :

$$p(r_t|\mathbf{h},\phi_t,\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},r_{0:t-1}) = \mathcal{N}_c\left(r_t;e^{j\phi_t}\mathbf{s}_t^T\mathbf{h},\sigma_{b,\lfloor t/L\rfloor}^{2(j)}\right) \quad \text{(A.19)}$$

and,

$$p(\mathbf{h}|\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},r_{0:t-1}) = \mathcal{N}_c\left(\mathbf{h};\mathbf{h}_{t-1}^{(j)},\mathbf{\Sigma}_{t-1}^{(j)}\right) \quad \text{(A.20)}$$

Therefore using (A.19) and (A.20), it is straightforward to show that the expression described in (A.17) can be rewritten as :

$$p(r_t|\phi_t,\phi_{0:t-1}^{(j)},\lambda_{\lfloor t/L\rfloor}^{(j)},r_{0:t-1}) = \mathcal{N}_c(r_t;\rho_t^{(j)},\chi_t^{(j)}) \quad \text{(A.21)}$$

where $\rho_t^{(j)} = e^{j\phi_t}\mathbf{s}_t^T\mathbf{h}_{t-1}^{(j)}$ and $\chi_t^{(j)} = \mathbf{s}_t^T\mathbf{\Sigma}_{t-1}^{(j)}\mathbf{s}_t^* + \sigma_{b,\lfloor t/L\rfloor}^{2(j)}$. According to (A.16)-(A.21), an analytical form of (A.15) remains untractable due to the double exponential in (A.21) by $\rho_t^{(j)}$. However, by linearizing the noise term v_k in (1.22), the mean of (A.21) is approximated by :

$$\rho_t^{(j)} \approx \begin{cases} (1+jv_0)\mathbf{s}_0^T\mathbf{h}_1^{(j)} & \text{if } t=0 \\ (1+jv_k)e^{j(\phi_{t-1}^{(j)}+2\pi\epsilon_{\lfloor t/L\rfloor}^{(j)}/N)}\mathbf{s}_t^T\mathbf{h}_{t-1}^{(j)} & \text{otherwise} \end{cases} \quad \text{(A.22)}$$

This approximation holds when the phase noise rate is small and is more accurate than the usual approximation $e^{j\theta_t} \approx 1 + j\theta_t$. Using (A.16), (A.21) and (A.22) and after several algebraic manipulations, (A.15) can be simplified as :

$$p(\phi_t|\phi_{0:t-1}^{(j)}, \lambda_{\lfloor t/L \rfloor}^{(j)}, r_{0:t}) \approx \mathcal{N}(\phi_t; \vartheta_t^{(j)}, \Lambda_t^{(j)}) \tag{A.23}$$

where $\Lambda_t^{(j)} = \frac{\chi_t^{(j)} \sigma_{v,\lfloor t/L \rfloor}^{2(j)}}{|\Gamma_t^{(j)}|^2 \sigma_{v,\lfloor t/L \rfloor}^{2(j)} + \chi_t^{(j)}}$ and

$$\vartheta_t^{(j)} = \begin{cases} \mathrm{Y}_0^{(j)} & \text{if } t = 0 \\ \mathrm{Y}_t^{(j)} + \phi_{t-1}^{(j)} + 2\pi\epsilon_{\lfloor t/L \rfloor}^{(j)}/N & \text{otherwise} \end{cases} \tag{A.24}$$

with $\mathrm{Y}_t^{(j)} = \frac{\Im(\Gamma_t^{(j)*} r_t)\sigma_{v,\lfloor t/L \rfloor}^{2(j)}}{|\Gamma_t^{(j)}|^2 \sigma_{v,\lfloor t/L \rfloor}^{2(j)} + \chi_t^{(j)}}$ ($\Im(.)$ denotes the imaginary part) and

$$\Gamma_t^{(j)} = \begin{cases} \mathbf{s}_0^T \mathbf{h}_{-1}^{(j)} & \text{if } t = 0 \\ e^{j(\phi_{t-1}^{(j)} + 2\pi\epsilon_{\lfloor t/L \rfloor}^{(j)}/N)} \mathbf{s}_t^T \mathbf{h}_{t-1}^{(j)} & \text{otherwise} \end{cases} \tag{A.25}$$

A.3 On the Gaussian Approximation of the Importance Function

As remarked in Section 4.3.2, the accuracy of the Gaussian approximation should be studied since the likelihood function is multimodal and the prior density has an infinite support [Vaswani 07]. In order to study precisely the validity of the Gaussian approximation, Table A.1 presents the gaps between the exact and the approximate cumulative distributions given for the t-th time sample and for the j-th particle by :

$$D_t^{(j)} = \max_{\phi_t} \left| F(\phi_t | \phi_{0:t-1}^{(j)}, \lambda_{\lfloor t/L \rfloor}^{(j)}, r_{0:t}) - F^*(\phi_t | \phi_{0:t-1}^{(j)}, \lambda_{\lfloor t/L \rfloor}^{(j)}, r_{0:t}) \right| \tag{A.26}$$

where $F(.)$ and $F^*(.)$ are the cumulative distributions of the exact and the Gaussian approximate optimal importance function. The exact optimal importance function is obtained using Eq. (A.15) and a sufficiently dense discretization of the state space of ϕ_k.

	βT	10^{-3}			10^{-2}		
	SNR (dB)	0	20	40	0	20	40
Gaps $D_t^{(j)}$	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

Table A.1 – *Study of the Gaussian approximation of the optimal importance function.*

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.

A.4 On the Gaussian approximation of the multicarrier signal

Let us recall that the transmitted multicarrier signal is given, for $t = 0, ..., N-1$, by :

$$s_{n,t} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} d_{n,i} e^{j2\pi it/N} \qquad \text{(A.27)}$$

where $\{d_{n,i}\}_{i=0}^{N-1}$ are a function of user data symbols and spreading codes, so depending of the multicarrier system considered. Since $\{d_{n,i}\}_{i=0}^{N-1}$ are *i.i.d.* variables (1.35) and according to the central limit theorem, we can expect that this multicarrier signal can be approximated as a circular Gaussian distributed random variable.

In order to study precisely the accuracy of the Gaussian approximation, Figure A.1 presents the maximum gaps between the cumulative distribution function (CDF) of both the real and imaginary part of the OFDM signal, denoted by $F(.)$ and the CDF of a Gaussian random variable with zero mean and variance $1/2$, denoted by $G(.)$. This gap is thus defined as $\max_x |F(x) - G(x)|$

For simplicity, an OFDM system is assumed in this appendix and thus $\{d_{n,i}\}_{i=0}^{N-1}$ are directly the user data symbol. Moreover, 2,000,000 OFDM symbols have been generated for each system configuration.

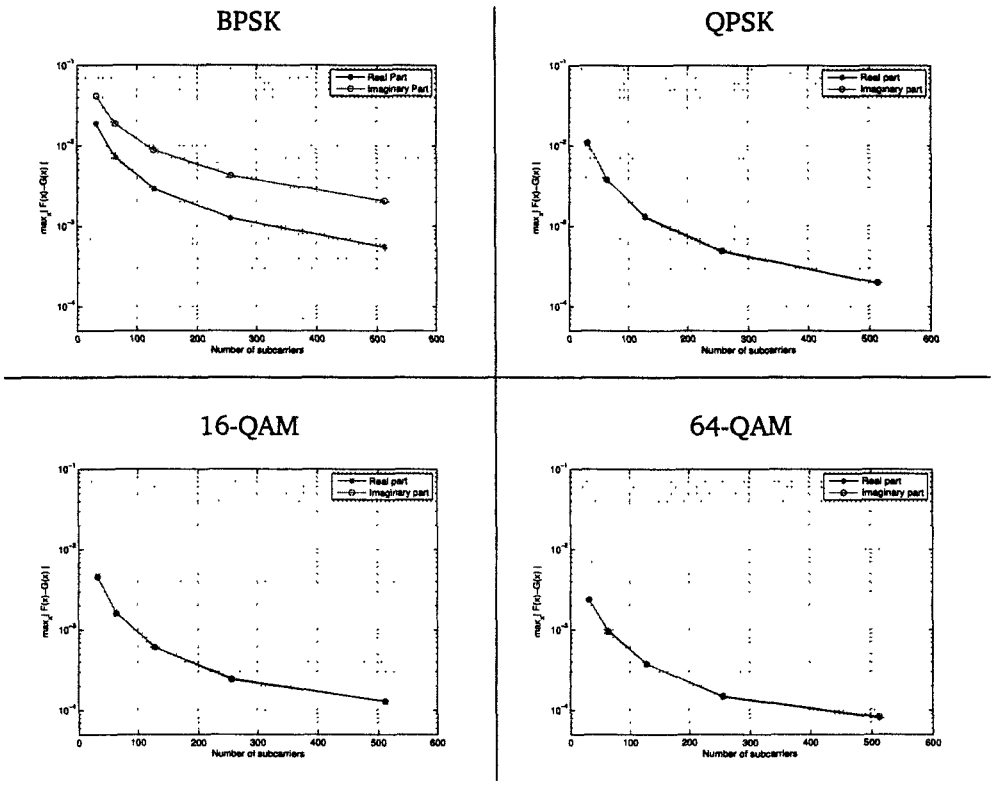

Figure A.1 – *Maximum gaps between the CDF of both the real and imaginary part of the OFDM signal and the CDF of a Gaussian random variable with zero mean and variance* $1/2$.

From these figures, it can be seen logically that the accuracy of a Gaussian approximation for the multicarrier signal increases with both the number of subcarriers and

the size of the constellation. Moreover, the approximation accuracy for the real and imaginary part of the multicarrier is identical except for the BPSK case. In fact, this difference with the use of BPSK can be simply explained by Eq. A.27. Indeed, when BPSK is used, both $s_{n,0}$ and $s_{n,N/2}$ have no imaginary part, *i.e.* :

$$s_{n,0} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} d_{n,i}$$
$$s_{n,N/2} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} d_{n,i}(-1)^i$$

As a consequence, with BPSK modulation, the Gaussian approximation of the real and imaginary part of the OFDM signal can be summarized as follows :

$$\Re(s_{n,t}) \sim \begin{cases} \mathcal{N}(0,1) & \text{if } t = 0 \text{ and } t = N/2 \\ \mathcal{N}(0,1/2) & \text{otherwise} \end{cases} \tag{A.28}$$

$$\Im(s_{n,t}) \sim \begin{cases} \delta(\Im(s_{n,t})) & \text{if } t = 0 \text{ and } t = N/2 \\ \mathcal{N}(0,1/2) & \text{otherwise} \end{cases} \tag{A.29}$$

These adaptations have to be taken into account in the proposed multicarrier receiver if the BPSK modulation is used.

From these results, it can be concluded that the Gaussian distribution $\mathcal{N}(0,1/2)$ represents an accurate model for both the real and imaginary part of the transmitted multicarrier signal, except in the BPSK case for $t = 0$ and $t = N/2$ where we have to use Eqs. (A.28) and (A.29).

A.5 Derivation of the PHN optimal importance function for the multicarrier signal estimation problem

The optimal importance function for phase distortion $\phi_{n,t}$ sampling in the proposed particle filter requires the derivation of the *p.d.f.* defined in (5.30). This p.d.f can be rewritten as :

$$p(\phi_{n,t}|\phi_{n,0:t-1}^{(j)}, \epsilon^{(j)}, r_{0:t}) = \frac{p(r_k|\phi_{n,t}, \phi_{n,0:t-1}^{(j)}, \epsilon^{(j)}, r_{0:t-1}) p(\phi_{n,t}|\phi_{n,0:t-1}^{(j)}, \epsilon^{(j)})}{\int p(r_k|\phi_{n,t}, \phi_{n,0:t-1}^{(j)}, \epsilon^{(j)}, r_{0:t-1}) p(\phi_{n,t}|\phi_{n,0:t-1}^{(j)}, \epsilon^{(j)}) d\phi_{n,t}} \quad \text{(A.30)}$$

with

$$p(\phi_{n,t}|\phi_{n,0:t-1}^{(j)}, \epsilon^{(j)}) = \begin{cases} \mathcal{N}(\phi_{n,t}; 0, \sigma_v^2) & \text{if } t = 0 \\ \mathcal{N}(\phi_{n,t}; \phi_{n,t-1}^{(j)} + 2\pi\epsilon^{(j)}/N, \sigma_v^2) & \text{otherwise} \end{cases} \quad \text{(A.31)}$$

and,

$$p(r_{n,t}|\phi_{n,t}, \phi_{n,0:t-1}^{(j)}, \epsilon^{(j)}, r_{n,0:t-1}) = \int p(r_{n,t}|\mathbf{u}_{n,t}, \phi_{n,t}, \phi_{n,0:t-1}^{(j)}, \epsilon^{(j)}, r_{n,0:t-1}) \times p(\mathbf{u}_{n,t}|\phi_{n,0:t-1}^{(j)}, \epsilon^{(j)}, r_{0:t-1}) d\mathbf{u}_{n,t} \quad \text{(A.32)}$$

The summand in (A.32) is the product of the likelihood *p.d.f.* and the posterior *p.d.f.* of the multicarrier signal $\mathbf{s}_{n,t}$ which are respectively given by :

$$p(r_{n,t}|\mathbf{u}_{n,t}, \phi_{n,t}, \phi_{n,0:t-1}^{(j)}, \epsilon^{(j)}, r_{0:t-1}) = \mathcal{N}_c\left(r_k; e^{j\phi_{n,t}}\mathbf{h}_n^T\mathbf{u}_{n,t}, \sigma_b^2\right) \quad \text{(A.33)}$$

and,

$$p(\mathbf{u}_{n,t}|\phi_{n,0:t-1}^{(j)}, \epsilon^{(j)}, r_{0:t-1}) = \mathcal{N}_c\left(\mathbf{u}_{n,t}; \mathbf{u}_{n,t-1|t-1}, \mathbf{\Sigma}_{n,t-1|t-1}^{(j)}\right) \quad \text{(A.34)}$$

Therefore using (A.33) and (A.34) , it is straightforward to show that the expression described in (A.32) can be rewritten as :

$$p(r_{n,t}|\phi_{n,t}, \phi_{n,0:t-1}^{(j)}, \epsilon^{(j)}, r_{n,0:t-1}) = \mathcal{N}_c(r_{n,t}; \rho_t^{(j)}, \chi_t^{(j)}) \quad \text{(A.35)}$$

where $\rho_t^{(j)} = e^{j\phi_{n,t}^{(j)}}\mathbf{h}_n^T\mathbf{s}_{n,t|t-1}^{(j)}$ and $\chi_t^{(j)} = \mathbf{h}_n^T\mathbf{\Sigma}_{n,t|t-1}^{(j)}\mathbf{h}_n^* + \sigma_b^2$. According to (A.31)-(A.35), an analytical form for (A.30) remains untractable due to the double exponential in (A.35). However, by linearizing $e^{jv_{n,t}}$ where the noise term $v_{n,t}$ is defined by (1.22), the mean of (A.35) is approximated by :

$$\rho_t^{(j)} \approx \begin{cases} (1 + jv_{n,0})\mathbf{h}_n^T(\mathbf{u}_{n,0|-1}^{(j)} + \mathbf{f}_{n,0}) & \text{if } t = 0 \\ (1 + jv_{n,t})e^{j(\phi_{n,t-1}^{(j)} + 2\pi\epsilon^{(j)}/N)}\mathbf{h}_n^T(\mathbf{u}_{n,t|t-1}^{(j)} + \mathbf{f}_{n,t}) & \text{otherwise} \end{cases} \quad \text{(A.36)}$$

This approximation holds when the phase distortion power is small and is more accurate than the usual approximation $e^{j\phi_{n,t}} \approx 1 + j\phi_{n,t}$. Using (A.31), (A.35) and (A.36)

and after several algebraic manipulations, the numerator of (A.30) can be simplified as :

$$p(r_t|\phi_{n,t},\phi_{n,0:t-1}^{(j)},\epsilon^{(j)},r_{0:t-1})p(\phi_{n,t}|\phi_{n,0:t-1}^{(j)},\epsilon^{(j)}) \approx \Upsilon_t^{(j)} e^{-\frac{1}{2\Lambda_t^{(j)}}\left[\frac{|r_{n,t}-\Gamma_t^{(j)}|^2}{|\Gamma_t^{(j)}|^2\sigma_v^2+\chi_t^{(j)}}-|\gamma_t^{(j)}|^2\right]} \times\mathcal{N}(\phi_{n,t};\mu_{n,t}^{(j)},\Lambda_t^{(j)}) \quad \text{(A.37)}$$

where

$$\mu_{n,t}^{(j)} = \begin{cases} \gamma_0^{(j)} & \text{if } t=0 \\ \gamma_t^{(j)} + \phi_{n,t-1}^{(j)} + 2\pi\epsilon^{(j)}/N & \text{otherwise} \end{cases} \quad \text{(A.38)}$$

and $\Upsilon_t^{(j)} = \sqrt{\Lambda_t^{(j)}/(2\pi\chi_t^{(j)}\sigma_v^2)}$, $\gamma_t^{(j)} = \frac{\Im(\Gamma_t^{(j)*}r_{n,t})\sigma_v^2}{|\Gamma_t^{(j)}|^2\sigma_v^2+\chi_t^{(j)}}$ (with $\Im(.)$ the imaginary part), $\Lambda_t^{(j)} = \frac{\chi_t^{(j)}\sigma_v^2}{|\Gamma_t^{(j)}|^2\sigma_v^2+\chi_t^{(j)}}$ and,

$$\Gamma_t^{(j)} = \begin{cases} \mathbf{h}_n^T(\mathbf{u}_{n,0|-1}^{(j)} + \mathbf{f}_{n,0}) & \text{if } t=0 \\ e^{j(\phi_{n,t-1}^{(j)}+2\pi\epsilon^{(j)}/N)}\mathbf{h}_n^T(\mathbf{s}_{n,t|t-1}^{(j)} + \mathbf{f}_{n,t}) & \text{otherwise} \end{cases} \quad \text{(A.39)}$$

Therefore from (A.37), it is obvious that :

$$\int p(r_t|\phi_{n,t},\phi_{n,0:t-1}^{(j)},\epsilon^{(j)},r_{0:t-1})p(\phi_{n,t}|\phi_{n,0:t-1}^{(j)},\epsilon^{(j)})d\phi_{n,t} \approx \Xi_t^{(j)} e^{-\frac{1}{2\Lambda_t^{(j)}}\left[\frac{|r_{n,t}-\Gamma_t^{(j)}|^2}{|\Gamma_t^{(j)}|^2\sigma_v^2+\chi_t^{(j)}}-|\gamma_t^{(j)}|^2\right]} \quad \text{(A.40)}$$

Finally, the optimal importance sampling for $\phi_{n,t}$ defined in (A.30) can be approximated by :

$$p(\phi_{n,t}|\phi_{n,0:t-1}^{(j)},\epsilon^{(j)},r_{0:t}) \approx \mathcal{N}(\phi_{n,t};\mu_{n,t}^{(j)},\Lambda_t^{(j)}) \quad \text{(A.41)}$$

A.6 SUITABILITY OF THE AR MODELING FOR THE PROCESS $u_{n,t}$

To evaluate the suitability of the AR modeling for the process $u_{n,t}$, we have made several comparisons between theorical and empirical results. We have have chosen for the OFDM parameters : $N = 64$,. $N_{cp} = 0$ and $P = 4$. The set of pilot tones are arbitrarily chosen as $\Omega_p = \begin{bmatrix} 4 & 19 & 43 & 59 \end{bmatrix}$. A bias of $\tau_0 = 10^{-13}$ was used to condition the Yule-Walker equation for the AR model. From figure A.2, we can see that the autocorrelation function obtained by generating correlated variate from the AR model closely matches the theorical autocorrelation function.For Figs. A.3 and A.4, we have simulated 50,000 OFDM symbols to obtain the empirical variance of the driving noise defined as :

$$\sigma^{2\ empiric}_{b_{n,t}} = E[|u_{n,t} + \sum_{i=1}^{k-1} a_{n,i} u_{n,t-i}|^2] \tag{A.42}$$

Fig. A.3 compares the evolution of the theorical (given by Eq. (5.16)) and empirical variance of the error term vs the index of subcarrier. Plot of the mean square error between the theorical and empirical variance of this driving noise is depicted in Fig. A.4. From this two figures, we can see that the AR modeling of the process $u_{n,t}$ presented in this section gives an accurately time evolution for this unknown data signal process.

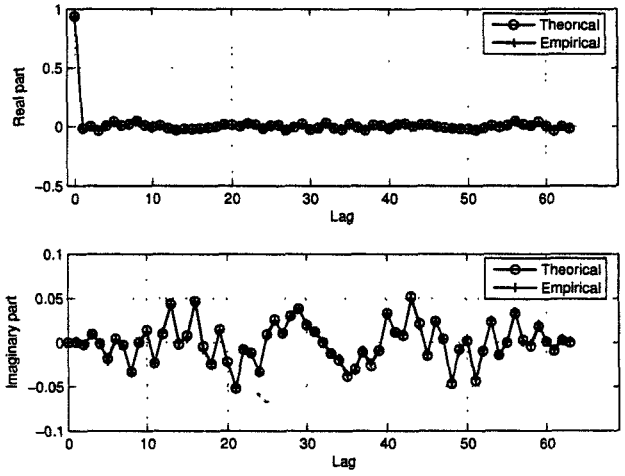

Figure A.2 – *Theorical and empirical autocorrelation function for the AR model.*

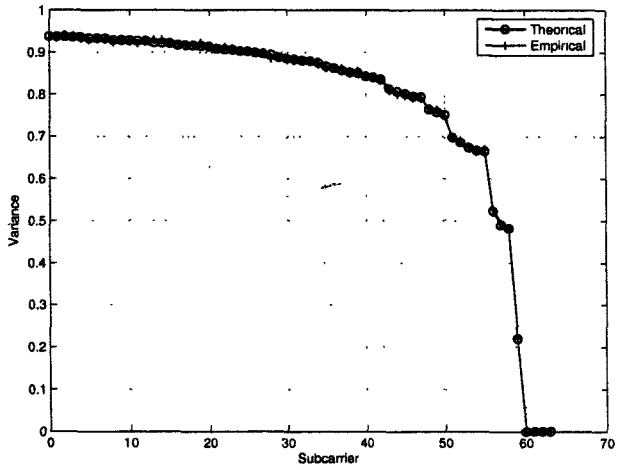

Figure A.3 – *Theorical and empirical variance of the driving noise of the AR process.*

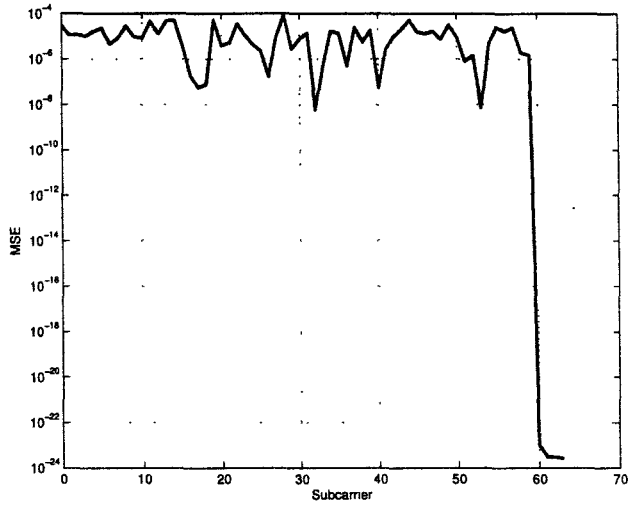

Figure A.4 – *Mean square error between theorical and empirical variance of the driving noise of the AR model.*

REFERENCES

[Andrieu *et al.* 03a] Andrieu, C., & Doucet, A. *Online expectation-maximization type algorithms for parameter estimation in general state space models.* Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03). Apr. 2003.

[Andrieu *et al.* 05] Andrieu, C., Doucet, A., & Tadic, V. *On-Line Parameter Estimation in General State-Space Models.* Proc. IEEE CDC-ECC, pp. 332–337. Dec. 2005.

[Andrieu *et al.* 03b] Andrieu, C., Freitas, N. D., Doucet, A., & Jordan, M. I. *An introduction to MCMC for Machine Learning.* Machine Learning, Vol. 50, n° 1-2, pp. 5–43, 2003.

[Arulampalam *et al.* 02] Arulampalam, M., Maskell, S., Gordon, N., & Clapp, T. *A tutorial on particle filters for online nonlinear/non-gaussian Bayesian tracking.* IEEE Trans. Signal Process., Vol. 50, n° 2, pp. 174–188, Feb. 2002.

[Baddour *et al.* 05] Baddour, K., & Beaulieu, C. *Autoregressive modeling for fading channel simulation.* IEEE Transactions on Wireless Communications, Vol. 4, pp. 1650 – 1662, July 2005.

[Beal 03] Beal, M. J. *Variational algorithms for approximate Bayesian inference.* Ph.D. Thesis, Gatsby Computational Neuroscience Unit, University College London, 2003. URL http://www.cse.buffalo.edu/faculty/mbeal/thesis/.

[Bingham 90] Bingham, J. *Multicarrier modulation for data transmission: an idea whose time has come.* IEEE Communications Magazine, Vol. 28, n° 5, pp. 5–14, May 1990.

[Bolic *et al.* 04] Bolic, M., Djuric, P., & Hong, S. *Resampling Algorithms for Particle Filters: A Computational Complexity Perspective.* EURASIP Journal on Applied Signal Processing, Vol. 15, pp. 2267–2277, 2004.

[Booth *et al.* 01] Booth, J., Hobert, J., & Jank, W. *A survey of Monte Carlo algorithms for maximizing the likelihood of a two-stage hierarchical model.* Statistical Modelling, Vol. 1, pp. 333–349, May 2001.

[Cappé *et al.* 05] Cappé, O., & Moulines, E. *On the Use of Particle Filtering for Maximum Likelihood Parameter Estimation.* European Signal Processing Conference (EUSIPCO'05). Turkey, Sep. 2005.

[Casas *et al.* 02] Casas, R. A., Biracree, S. L., & Youtz, A. E. *Time Domain Phase Noise Correction for OFDM signals.* IEEE Trans. Broadcast., 2002.

[Casella *et al.* 96] Casella, G., & Robert, C. *Rao-Blackwellisation of sampling schemes.* Biometrika, Vol. 83, pp. 81–94, Mar. 1996.

[Celeux *et al.* 85] Celeux, G., & Diebolt, J. *The SEM algorithm: a Probabilistic Teacher Algorithm Derived from the EM Algorithm for the Mixture Problem.* Computational Statistics Quaterly, Vol. 2, pp. 73–82, 1985.

[Cérou *et al.* 01] Cérou, F., Gland, F. L., & Newton, N. *Stochastic particle methods for linear tangent filtering equations.* Optimal Control and PDE's - Innovations and Applications. In honor of Alain Bensoussan on the occasion of his 60th birthday, edited by J.-L. Menaldi, E. Rofman, *et al.*, pp. 231–240. IOS Press, 2001.

[Chang 66] Chang, R. *Synthesis of band-limited orthogonal signals for multichannel data transmission.*. Bell Systems Technical Journal, Vol. 46, pp. 1775–1796, Dec. 1966.

[Chen *et al.* 00] Chen, R., & Liu, J. S. *Mixture Kalman filters.* Journal of the Royal Statistical Society, Vol. 62, pp. 493–508, 2000.

[Chopin 02] Chopin, N. *A sequential particle filter method for static models.* Biometrika 89, pp. 539–552, 2002.

[Chouly *et al.* 93] Chouly, A., Brajal, A., & Jourdan, S. *Orthogonal multicarrier techniques applied to direct sequence spread spectrum CDMA systems.*. Proc. IEEE GLOBECOM, pp. 1723–1728. Houston, USA, Nov. 1993.

[Cimini 85] Cimini, L. *Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing.*. IEEE Transactions on Communications, Vol. 33, pp. 665–675, Jul. 1985.

[Costa *et al.* 02] Costa, E., & Pupolin, S. *M-QAM-OFDM System Performance in the Presence of a Nonlinear Amplifier and Phase Noise.* IEEE Transactions on Communications, Vol. 50, n° 3, pp. 462–472, Mar. 2002.

[Crisan *et al.* 02] Crisan, D., & Doucet, A. *A survey of convergence results on particle filtering methods for practitioners.* IEEE Transactions on Signal Processing, Vol. 50, n° 3, pp. 95–108, 2002.

[DaSilva *et al.* 93] DaSilva, V., & Sousa, E. *Performance of orthogonal CDMA codes for quasi-synchronous communication system.*. Proc. IEEE ICUPC, pp. 995–999. Ottawa, Canda, Oct. 1993.

[Delyon *et al.* 99] Delyon, B., Lavielle, M., & Moulines, E. *Convergence of a Stochastic Approximation version of the EM algorithm.* The Annals of Statistics, Vol. 27, pp. 94–128, 1999.

[Dempster *et al.* 77] Dempster, A. P., Laird, N. M., & Rubin, D. B. *Maximum Likelihood from Incomplete Data via the* EM *Algorithm.* Journal of the Royal Statistical Society. Series B (Methodological), Vol. 39, n° 1, pp. 1–38, 1977.

[Devroye 86] Devroye, L. *Non-Uniform Random Variate Generation.* New-York: Springer, 1986.

[Djuric *et al.* 03] Djuric, P. M., Kotecha, J. H., Zhang, J., Huang, Y., Ghirmai, T., Bugallo, M. F., & MÃ-guez, J. *PArticle Filtering.* IEEE Signal Processing Magazine, Vol. 20, n° 5, pp. 19–38, Sep. 2003.

[Douc. *et al.* 05] Douc., R., Cappé, O., & Moulines, E. *Comparison of resampling schemes for particle filtering.* Proc. of Image and Signal Processing (ISPA), pp. 64–69. Sep. 2005.

[Doucet *et al.* 01] Doucet, A., Freitas, N. D., & Gordon, N. *Sequential Monte Carlo Methods in Practice.* Springer-Verlag, 2001.

[Doucet *et al.* 00] Doucet, A., Godsill, S., & Andrieu, C. *On sequential Monte Carlo sampling methods for Bayesian filtering.* Statist. Comput., Vol. 10, pp. 197–208, 2000.

[Doucet *et al.* 03] Doucet, A., & Tadic, V. *Parameter estimation in general state-space models using particle methods : New trends in statistical information processing.* Ann. Inst. Stat. Math., Vol. 55, n° 2, pp. 409–422, 2003.

[Fazel *et al.* 93] Fazel, K., & Papke, L. *On the performance of convolutionally-coded CDMA/OFDM for mobile communication system.*. Proc. IEEE PIMRC, pp. 468–472. Yokohama, Japan, Sep. 1993.

[Fearnhead 01] Fearnhead, P. *Markov chain Monte Carlo, sufficient statistics and particle filter.* J. Comput. Graph. Stat., Vol. 11, n° 4, pp. 848–862, Jul. 2001.

[Garnier *et al.* 02] Garnier, C., Clavier, L., Delignon, Y., Loosvelt, M., & Boulinguez, D. *Multiple access for 60 GHz mobile ad hoc networks*. Proc. IEEE VTC Spring, Vol. 3, pp. 1517–1521. May 2002.

[Ghogho *et al.* 04] Ghogho, M., & Swami, A. *Carrier frequency synchronization for OFDM systems*. Chapter 8, in *Signal Processing for Mobile Communications Handbook*, Ed. M. Ibnkahla, CRC, 2004.

[Gilks *et al.* 01] Gilks, W., & Berzuini, C. *Following a moving target - Monte Carlo inference for dynamic Bayesian models*. Journal of the Royal Statistical Society. Series B, Vol. 63, pp. 127–146, 2001.

[Gordon *et al.* 93] Gordon, N., Salmond, D., & Smith, A. *Novel approach to nonlinear/non-gaussian Bayesian state estimation*. IEE Proceedings-F, Vol. 140, pp. 107–113, Apr. 1993.

[Hara *et al.* 96] Hara, S., & Prasad, R. *DS-CDMA, MC-CDMA and MT-CDMA for mobile multimedia communications*. Proc. IEEE Vehicular TEchnology Conference (VTC'96), pp. 1106–1110. Atlanta, USA, Apr./May 1996.

[Haykin 91] Haykin, S. *Adaptive Filter Theory, 2nd ed.* Englewood Cliffs, NJ: Prentice-Hall, 1991.

[Hol *et al.* 06] Hol, J., Schon, T., & Gustafsson, F. *On Resampling Algorithms for Particle Filters*. Nonlinear Statistical Signal Processing Workshop. Cambridge, United-Kingdom, Sep. 2006.

[Hrycak *et al.* 06] Hrycak, T., & Matz, G. *Low-Complexity Time-Domain* ICI *Equalization for* OFDM *Communications Over Rapidly Varying Channels*. Asilomar Conference on Signals, Systems and Computers, pp. 1767–1771. 2006.

[Jordan 98] Jordan, M. I. *Learning in Graphical Models*. Kluwer, Series D: Behavioural and Social Sciences, 1998.

[Julier *et al.* 97] Julier, S. J., & Uhlmann, J. K. *A New Extension of the Kalman Filter to Nonlinear Systems*. Proc. SPIE Int. Soc. Opt. Eng., Vol. 3068, pp. 182–193. Orlando, (FL-USA), 1997.

[Kaiser 96] Kaiser, S. *Trade-off between channel coding and spreading in multi-carrier CDMA systems*. Proc. IEEE International Symposium on Spread Spectrum Techniques & Applications (ISSSTA'96), pp. 1366–1370. Mainz, Germany, Sep. 1996.

[Kalman *et al.* 61] Kalman, R., & Bucy, R. *New results in linear filtering and prediction theory*. Journal of Basic Engineering, Transactions ASME Series D, , n° 83, pp. 95–108, 1961.

[Kondo *et al.* 93] Kondo, S., & Milstein, L. *On the use of multicarrier direct sequence spread spectrum systems*. Proc. IEEE MILCOM, pp. 52–56. Boston, USA, Oct. 1993.

[Kong *et al.* 94] Kong, A., Liu, J., & Wong, W. *Sequential imputations and Bayesian missing data problems*. Journal of the American Statistical Association, Vol. 89, pp. 278–288, 1994.

[Lin *et al.* 06] Lin, D., Pacheco, R., Lim, T., & Hatzinakos, D. *Joint estimation of Channel Response, Frequency Offset and Phase Noise in OFDM*. IEEE Transactions on Signal Processing, Vol. 54, n° 9, pp. 3542–3554, Sep. 2006.

[Lin *et al.* 07] Lin, D. D., & Lim, T. J. *The variational inference approach to joint data detection and phase noise estimation in OFDM*. IEEE Transactions on Signal Processing, Vol. 55, n° 5, pp. 1862–1874, May 2007.

[Lin *et al.* 05] Lin, D. D., Zhao, Y., & Lim, T. J. OFDM *Phase Noise Cancellation via Approximate Probabilistic Inference*. IEEE WCNC, pp. 27–32. Mar. 2005.

[Liu *et al.* 98] Liu, J., & Chen, R. *Sequential Monte Carlo Methods for Dynamic Systems*. Journal of the American Statistical Association, Vol. 93, n° 443, pp. 1032–1044, 1998.

[Liu *et al.* 01] Liu, J., & West, M. *Combined parameter and state estimation in simulation-based filtering.* Sequential Monte Carlo in Practice, edited by A. Doucet, J. de Freitas, *et al.*, pp. 197–223. New-York : Springer-Verlag, 2001.

[Musso *et al.* 01] Musso, C., Oudjane, N., & Gland, F. L. *Improving regularized particle filters.* Sequential Monte Carlo in Practice, edited by A. Doucet, J. de Freitas, *et al.*, pp. 247–271. New-York : Springer-Verlag, 2001.

[Natarajan *et al.* 04] Natarajan, B., Wu, Z., Nassar, C., & Shattil, S. *Large set of CI scpreading codes for high-capacity MC-CDMA.* IEEE Transactions on Communications, Vol. 52, n° 11, pp. 1862–1866, Nov. 2004.

[Neal *et al.* 98] Neal, R. M., & Hinton, G. E. *A view of the EM algorithm that justifies incremental, sparse, and other variants.* in *Learning in graphical models*, M. I. Jordan, Ed. Norwell, MA : Kluwer, 1998.

[Nikitopoulos *et al.* 05] Nikitopoulos, K., & Polydoros, A. *Phase-impairment effects and compensation algorithms for* OFDM *systems.*. IEEE Trans. Commun., Vol. 53, n° 4, pp. 698–707, Apr. 2005.

[Nocedal *et al.* 99] Nocedal, J., & Wright, S. J. *Numerical optimization.* New York: Springer-Verlag, 1999.

[Olsson *et al.* 06] Olsson, J., Cappé, O., Douc, R., & Moulines, E. *Sequential Monte Carlo smoothing with application to parameter estimation in non-linear state space models.* Technical report, Lund University, 2006. URL http://arxiv.org/abs/math/0609514.

[Papavasiliou 05] Papavasiliou, A. *A uniformly convergent adaptive particle filter.* J. Appl. Probab., Vol. 42, n° 4, pp. 1053–1068, 2005.

[Petrovic *et al.* 04a] Petrovic, D., Rave, W., & Fettweis, G. *Common Phase Error due to Phase Noise in* OFDM *- Estimation and Suppression.* Proc. IEEE PIMRC. 2004.

[Petrovic *et al.* 04b] Petrovic, D., Rave, W., & Fettweis, G. *Intercarrier Interference due to Phase Noise in OFDM - Estimation and Suppression.* Proc. IEEE VTC Fall. Sep. 2004.

[Piazzo *et al.* 02] Piazzo, L., & Mandarini, P. *Analysis of phase noise effects in OFDM modems.* IEEE Transactions on Communications, Vol. 50, n° 10, pp. 1696–1705, Oct. 2002.

[Pitt *et al.* 99] Pitt, M. K., & Shephard, N. *Filtering via Simulation: Auxiliary Particle Filters.* Journal of the American Statistical Association, Vol. 94, n° 446, pp. 590–599, Jun. 1999.

[Pollet *et al.* 95] Pollet, T., van Bladel, M., & Moeneclaey, M. *BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise.* IEEE Transactions on Communications, Vol. 43, pp. 191–193, Feb. 1995.

[Popovic 99] Popovic, B. *Spreading Sequences for multicarrier CDMA systems.* IEEE Transactions on Communications, Vol. 47, n° 6, pp. 918–926, Jun. 1999.

[Poyiadjis *et al.* 06] Poyiadjis, G., Singh, S. S., & Doucet, A. *Gradient-free maximum likelihood parameter estimation with particle filters.* Proc. American Control Conference. Jun. 2006.

[Proakis 95] Proakis, J. *Digital Communications.* Mc Graw-Hill, third edition, 1995.

[Rappaport 02] Rappaport, T. *Wireless Communications: Principles & Practice.* Prentice-Hall, 2nd Ed., 2002.

[Reiners *et al.* 94] Reiners, C., & Rohling, H. *Multicarrier transmission techniques in cellular mobile communications systems.* Proc. IEEE VTC, pp. 1645–1649. Stockholm, Sweden, Jun. 1994.

[Robert *et al.* 99] Robert, C., & Casella, G. *Monte Carlo Statistical Methods*. New York: Springer-Verlag, 1999.

[Robertson *et al.* 95] Robertson, P., & Kaiser, S. *Analysis of the effects of phase noise in orthogonal frequency division multiplexing (OFDM) systems*. Proc. IEEE ICC, pp. 1652–1657. 1995.

[Rohling *et al.* 96a] Rohling, H., & Grünheid, R. *OFDM transmission technique with flexible subcarrier allocation*. Proc. International Conference on Telecommmunciations (ICT'96), pp. 567–571. Istanbul, Turkey, Apr. 1996.

[Rohling *et al.* 96b] Rohling, H., & Grünheid, R. *Performance of an OFDM-TDMA mobile communication system*. Proc. IEEE Vehicular TEchnology Conference (VTC'96), pp. 567–571. Atlanta, USA, Apr./May 1996.

[Ryden 94] Ryden, T. *Consistent and asymptotically normal parameter estimates for hidden Markov models*. Ann. stat., Vol. 22, pp. 1884–1895, 1994.

[Schnell 94] Schnell, M. *Hadamard codewords as orthogonal spreading sequences in synchronous DS-CDMA systems for mobile radio channels*. Proc. ISSSTA, Vol. 2, pp. 505–509. Jul. 1994.

[Schon *et al.* 05] Schon, T., Gustafsson, F., & Nordlund, P. *Marginalized particle filters for mixed linear/nonlinear state-space models*. IEEE Transactions on Signal Processing, Vol. 53, pp. 2279–2289, Jul. 2005.

[Sheng *et al.* 98] Sheng, S., & Brodersen, R. *Low-Power CMOS Wireless Communications: A Wideband CDMA System Design*. Kluwer Academic Publisher, 1998.

[Smidl *et al.* 06] Smidl, V., & Quinn, A. *The variational Bayes method in signal processing*. Springer, 2006.

[Sorenson 88] Sorenson, H. *Recursive estimation for nonlinear dynamic systems*. Bayesian Analysis of Time Series and Dynamic Models, New York : Dekker, ed. J.C. Spall, 1988.

[Steedam *et al.* 01] Steedam, H., & Moeneclaey, M. *The effect of carrier phase jitter on* MC-DS-CDMA. Proc. IEEE ICC, Vol. 6, pp. 1881–1884. Jun. 2001.

[Tarighat *et al.* 03] Tarighat, A., & Sayed, A. *An optimum OFDM receiver exploiting cyclic prefix for improved data estimation*. Proc. ICASSP'03, Vol. 4, pp. 217–220. Hong-Kong, China, Apr. 2003.

[ANSI T1E1.4 Committee Contribution 93] ANSI T1E1.4 Committee Contribution. *The DWMT: a multicarrier transceiver for ADSL using M-band wavelets*. Tech. Rep., ANSI, 1993.

[ETSI Normalization Committee 96] ETSI Normalization Committee. *Digital broadcasting systems for television, sound and data services*. Norme ETSI, ETS 300 744, 1996.

[ETSI Normalization Committee 97] ETSI Normalization Committee. *Radio broadcasting systems, digital audio broadcasting (DAB) to mobile, portable and fixed receivers*. Norme ETSI, ETS 300 401, 1997.

[ETSI Normalization Committee 99] ETSI Normalization Committee. *High performance radio local area networks (HIPERLAN) Type 2*. Norm ETSI, ETR101683114, 1999.

[Tichavsky 95] Tichavsky, P. *Posterior Cramér-Rao bounds for adaptive harmonic retrieval*. IEEE Transactions on Signal Processing, Vol. 43, pp. 1299–1302, May 1995.

[Tichavsky *et al.* 98] Tichavsky, P., Muravchik, C., & Nehorai, A. *Posterior Cramér-Rao bounds for discrete-time nonlinear filtering*. IEEE Transactions on Signal Processing, Vol. 46, n° 5, pp. 1386–1396, May 1998.

[Tomba 98] Tomba, L. *On the effect of Wiener phase noise in OFDM systems*. IEEE Transactions on Communications, Vol. 46, pp. 580–583, May 1998.

[Tomba *et al.* 99] Tomba, L., & Krzymien, W. *Sensitivity of the* MC-CDMA *access scheme to carrier phase noise and frequency offset.*. IEEE Trans. Veh. Technol., Vol. 48, n° 5, pp. 1657–1665, Sep. 1999.

[van Lint *et al.* 93] van Lint, J., & Wilson, R. *A course in combinatorics*. New York: Cambridge University Press, 1993.

[van Trees 68] van Trees, H. L. *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968.

[Vaswani 07] Vaswani, N. PF-EIS & PF-MT: *New Particle Filtering Algorithms for Multimodal Observation Likelihoods and Large Dimensional State Spaces*. Proc. IEEE ICASSP, pp. 1193–1196. Honolulu, Hawaii, Apr. 2007.

[Wang *et al.* 00] Wang, Z., & Giannakis, G. *Wireless multicarrier communications*. IEEE Signal Processing Magazine, Vol. 17, pp. 29–48, May 2000.

[Wei *et al.* 90] Wei, G., & Tanner, M. *A Monte Carlo Implementation of the EM Algorithm and the Poor Man's Data Augmentation Algorithms*. Journal of the American Statistical Association, Vol. 85, pp. 699–704, 1990.

[Weinstein *et al.* 71] Weinstein, S. B., & Ebert, P. M. *Data transmission by frequency division multiplexing using the discrete Fourier transform*. IEEE Trans. Commun., Vol. 19, n° 5, pp. 628–634, Oct. 1971.

[Wu *et al.* 03a] Wu, S., & Bar-Ness, Y. *A new phase noise mitigation method in OFDM systems with simultaneous CPE and ICI correction*. Proc. MCSS. Sep. 2003.

[Wu *et al.* 03b] Wu, S., & Bar-Ness, Y. *OFDM Channel Estimation in the Presence of Frequency Offset and Phase Noise*. Proc. IEEE ICC, Vol. 5, pp. 3366–3370. May 2003.

[Wu *et al.* 04] Wu, S., & Bar-Ness, Y. OFDM *Systems in the presence of phase noise : consequences and solutions*. IEEE Trans. Commun., Vol. 52, pp. 1988–1996, Nov. 2004.

[Yee *et al.* 05] Yee, D., Reilly, J., & Kirubarajan, T. *Channel Equalization and Phase Noise Suppression for* OFDM *systems in a time-varying frequency channel using Particle Filtering*. Proc. IEEE Workshop ICASSP. Mar. 2005.

[Yee *et al.* 93] Yee, N., Linnartz, J., & Fettweis, G. *Multicarrier CDMA in indoor wireless radio networks.*. Proc. IEEE PIMRC, pp. 109–113. Yokohama, Japan, Sep. 1993.

[Zaritskii *et al.* 75] Zaritskii, V., Svetnik, V., & Shimelevich, L. *Monte Carlo technique in problems of optimal data processing*. Automation and remote control, pp. 95–103, 1975.

Titre *Méthodes séquentielles de Monte-Carlo pour les systèmes multiporteuses en présence de distorsions de phase.*

Résumé Dans le contexte d'une demande croissante de débits de communications de plus en plus élevés, les systèmes multiporteuses ont suscité un grand intérêt dans la communauté scientifique depuis ces dernières années et sont désormais employées dans de nombreux systèmes de communication. Malheureusement, ce type de système est extrêmement sensible aux distorsions de phase, comme le bruit de phase et le décalage fréquentiel de la porteuse, engendrées par l'instabilité des oscillateurs locaux. Le but de cette thèse est donc de concevoir un récepteur capable de compenser ces perturbations. Notre approche à ce problème non-linéaire est basée sur l'inférence Bayésienne et plus particulièrement sur les méthodes séquentielles de Monte-Carlo, appelées également filtrage particulaire. En premier lieu, nous proposons un estimateur conjoint du canal et des distorsions de phase grâce à une séquence d'apprentissage. Ensuite, nous traitons le problème de l'estimation des données en présence de distorsions de phase. Les filtres particulaires proposés sont implémentés efficacement en combinant le principe d'échantillonnage par importance, un schéma de sélection, une technique de réduction de variance d'estimée et surtout un nouvel estimateur "on-line" de paramètres utilisant des algorithmes d'espérance-maximisation stochastique parallèles. De plus, dans le but toujours d'augmenter l'efficacité et la robustesse de nos estimateurs, nous proposons une modélisation originale du signal multiporteuses dans le domaine temporel par un processus autorégréssif dans le cas où des porteuses nulles ou pilotes sont présentes.

Mots-clés Systèmes multiporteuses, Multiplexage par division en fréquence orthogonales (OFDM), Inférence Bayésienne, estimation conjointe, canal, bruit de phase, décalage fréquentiel, méthodes de Monte-Carlo, filtrage particulaire, estimation des paramètres.

Title *Sequential Monte Carlo receiver for multicarrier systems in the presence of phase distortions.*

Abstract Multicarrier transmission systems have aroused great interest in recent years as a potential solution to the problem of transmitting high data rate over a frequency selective fading channel. Nowadays, multicarrier modulation is being selected as the transmission scheme for the majority of new communication systems. However, multicarrier systems are very sensitive to phase noise and carrier frequency offset caused by the oscillator instabilities. In this thesis, a general receiver for compensating the phase distortions effects in multicarrier systems is proposed. Our approach to this non-linear problem is based on Bayesian inference using sequential Monte Carlo filtering also referred to as particle filtering. First, the problem of channel estimation in the presence of phase noise and carrier frequency offset is addressed. Then, a particle filter is proposed to include the joint signal, phase noise and carrier frequency offset estimation. The proposed sequential Monte Carlo filters are efficiently implemented by combining sequential importance sampling, a selection scheme, a variance reduction technique and especially a new on-line parameter estimation based on parallel stochastic expectation maximization algorithms. Moreover in order to improve the estimation accuracy, an original autoregressive modeling of the time-domain multicarrier signal including either pilot or null-subcarriers is also proposed. Extensive simulation study is provided to illustrate the efficiency and the robustness of the proposed algorithms in comparison with those of existing schemes.

Keywords Multicarrier systems, orthogonal frequency-division multiplexing (OFDM), Bayesian inference, joint estimation, channel, phase noise, carrier frequency offset, Monte-Carlo methods, particle filtering, parameter estimation.