

Full length article

Decision-directed Kalman particle filter with application to the MIMO phase noise channel

A. Spalvieri^a, L. Reggiani^a, L. Dossi^{b,*}^a DEIB - Politecnico di Milano, Milano, 20133, Italy^b IEIIT - CNR, c/o DEIB - Politecnico di Milano, Milano, 20133, Italy

ARTICLE INFO

Article history:

Received 27 September 2021

Received in revised form 22 April 2022

Accepted 2 September 2022

Available online 17 September 2022

Keywords:

Kalman filter

Particle filter

Bayesian tracking

MIMO fading channel

MIMO phase noise channel

ABSTRACT

Demodulation of data transmitted over time-varying channels with a free running hidden Markov state, like the phase noise channel or the fading channel, requires that the receiver tracks the hidden channel state. The tracking technique adopted in the paper is based on non-data-aided sequential importance sampling, also known as particle filtering.

The paper proposes a new particle filtering framework for data communication receivers based on an importance distribution such that each individual particle becomes a decision-directed Kalman filter relying upon its local symbol-by-symbol hard decisions. In this framework, different particles are left free to take different sequences of decisions. This leaves to the receiver the possibility of exploring different sequences of transmitted modulation symbols. The weight of the particle will be high for those particles that took in the past the correct sequence of decisions, while will be low for those particles that took wrong decisions. In the resampling procedure, particles with high weight will survive, while particles with low weight will be terminated, leaving space to the birth of new particles resampled from the surviving ones.

The crucial point in importance sampling is the choice of the importance distribution and the main novelty of the paper is the proposal of an importance distribution such that the particles of the particle filter become decision-directed Kalman filters. One important benefit brought by our proposed method is that, being non-data-aided, it does not need pilot symbols, thus allowing to preserve the transmission rate. A significant application example, presented and developed in the paper, is constituted by MIMO systems affected by phase noise, where the channel state vector consists of many parameters.

© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In data transmission over a fast time varying channel like, for instance, the phase noise channel or the fading channel, the receiver needs to track the posterior distribution of the free-running state of the channel in order to make feasible the demodulation. Tracking can become particularly challenging in MIMO (Multiple Input Multiple Output) systems, where the channel state vector may consist of many variables [1,2].

In data-aided tracking, the known transmitted modulation symbol is cancelled from the received signal, leading to a linear or linearizable channel model that, in the mentioned cases of fading and phase noise channels, allows to track the posterior probability of the channel state (i.e. the probability distribution of the channel state given the received signal sequence and the known modulation symbol sequence) by means of a parametric

approach as, for instance, the Kalman filter or the Wiener filter. Kalman and Wiener filters have been used widely in the past for the fading channel, as can be seen, for instance, in [3,4], where error-free decisions are assumed and used as if they were known data, and for the phase noise channel, in [5–7] and also [8], where the performance of Kalman filter and Wiener filter are compared.

Non-data-aided tracking is attractive because, compared to data-aided tracking, allows to save the overhead due to pilot symbols. Among non-data-aided techniques, decision-directed methods, that make use of the tentative decisions made at the receive side in place of the truly known data, offer good performance when the probability of error is small [5], that, however, rapidly degrades due to the catastrophic phenomenon of error propagation, when decision errors increase their occurrence rate [8].

In order to better understand the differences between data-aided and non-data-aided tracking, we observe that, while with linear or linearizable channel models the data-aided posterior distribution of the channel state is mono-modal, the non-data-aided tracking posterior distribution of the channel state can

* Corresponding author.

E-mail addresses: arnaldo.spalvieri@polimi.it (A. Spalvieri), luca.reggiani@polimi.it (L. Reggiani), laura.dossi@ieiit.cnr.it (L. Dossi).

be multi-modal, as in the study case of the MIMO phase noise channel proposed here. The possibly multi-modal nature of the posterior distribution makes its tracking even more challenging. One possibility is to resort to non-parametric methods, which have the ability of tracking the multiple modes, an ability that a single decision-directed Kalman or Wiener filter has not. A variety of non-parametric methods have been proposed in the past. For instance, non-data-aided and non-parametric tracking can be performed by the trellis-based BCJR algorithm [9]. Various implementations of trellis-based tracking have been proposed in the literature, see e.g. [10,11], depending on the role that is given to the modulation symbol. In all cases, the state space of the phase noise channel is discretized into bins and the bins are put in correspondence with trellis states. However, with the MIMO channel, the size of the channel state vector grows and the number of states of the trellis explodes exponentially with the size of the channel state vector, making the trellis approach unfeasible.

An alternative non-data-aided and non-parametric approach is based on sequential importance sampling, also known as particle filtering, which has found wide application for the wireless channel: a sample of papers on this research line is [12–15]. The advantage of the particle filter compared to trellis-based methods is that its complexity, i.e. the number of particles, remains under control. At the same time, a key step in importance sampling remains the choice of the importance distribution, which constitutes also the core novelty in this paper.

Compared to the vast literature on non-data-aided demodulation methods and particle filtering for channel demodulation, the following novel contributions characterize this paper:

- The design of a detector for communication channels based on the integration between Kalman filtering and particle filtering, in which through a specific choice of the importance distribution, the individual particle is a Kalman filter that updates its state relying upon the hard symbol-by-symbol decisions locally taken by the particle itself. The use of local decisions seems to be new, compared to the existing literature, e.g. [16]. This leaves to the receiver the possibility of exploring different sequences of transmitted modulation symbols. In the resampling procedure, particles that took the correct sequence of decisions will have high weight and will survive, while particles that took wrong decisions will have low weight and will be terminated, leaving space to the birth of new particles resampled from the surviving ones. This structure allows the feasible implementation of a challenging demodulation, i.e. a non-data-aided parametric tracking of time varying channels with continuous states, which would become easily not treatable with trellis-based approaches (for example with Delayed Decision-Feedback Sequence Estimators). The multiple particles show their capability, w.r.t. a mono-dimensional system, of tracking multi-modal multi-dimensional posterior distributions of the hidden channel state. Nevertheless, when implemented with a single particle, the proposed demodulator reduces to a decision directed Kalman filter structure.
- The analysis of the computational complexity of a practical detector based on this approach and the design of straightforward solutions for the implementation with reduced complexity.
- The application to the 2×2 MIMO Line of Sight channel with independent oscillators affected by Wiener phase noise, with the discussion of the phase ambiguities issue in the state vector. The phase ambiguities have been solved with a reliable solution, validated by simulation results that

show how the proposed reduced complexity implementations of the decision-directed Kalman particle filter outperform data-aided linearized Kalman filter techniques, as discussed in Section 5. This application can be easily extended to a $N \times M$ MIMO channels in presence of phase noise.

With regard to the core contribution of the paper, i.e. the proposal of the importance distribution in this particle filtering, we found inspiration, more than from the already cited bibliography, from two approaches that have been developed in very different contexts, integrated here in the novel detector: one is the ensemble Kalman filter (see [17] for a comprehensive presentation), which can be seen as a particle filter where each one of the particles is actually a Kalman filter [18], and the other one is the dual decision-feedback equalizer in [19]. Although here the state is free-running, while in [19] the state is data-driven, our proposal has a crucial feature in common with [19]: also in this paper many independent decision-directed tracking algorithms are implemented in parallel, each one based on its own decisions (actually, only two in the example proposed in [19], but the extension to many is straightforward). Notice that this is substantially different from [13], where the decision is global, and from [15], which is not decision-directed.

The outline of the paper is as follows: Section 2, after a brief review on the application of particle filtering to demodulation of signals transmitted over channels with time-varying states, introduces the main novel contribution of the paper, i.e. the decision-directed Kalman particle filtering technique, showing the importance function that leads to this specific form of particle filtering. In Section 3 we discuss the implementation complexity, the specific resampling technique used for the simulation results and the options for the complexity reduction. In Section 4 we present the model of the MIMO phase noise channel with independent oscillators used as the application case in order to evaluate the performance of the proposed method by simulation results, which are finally reported and discussed in Section 5. Finally, conclusions are drawn in Section 6.

2. Channel tracking and demodulation by decision directed Kalman particles

We consider a transmission channel model with hidden input modulation symbol sequence $x_1^N = \{x_i, i = 1, \dots, N\}$ made of i.i.d. symbols x_i up to time N , drawn from the symbol set \mathcal{X} with size equal to $N_{\mathcal{X}}$. We assume that the input symbol sequence x_1^N is independent from the first-order Markov free-running channel state sequence $s_0^N = \{s_i, i = 0, 1, \dots, N\}$, hence their joint probability distribution $p(s_0^N, x_1^N)$ is

$$p(s_0^N, x_1^N) = p(s_0) \prod_{k=1}^N p(s_k | s_{k-1}) p(x_k). \quad (1)$$

In [20] the reader can find further information on the topics addressed here, i.e. the model assumptions and Bayesian algorithms for nonlinear tracking problems, including particle filters.

We make the assumption of memoryless channel given the state, i.e. we assume that the conditional probability distribution $p(y_1^N | s_1^N)$, where $y_1^N = \{y_i, i = 1, \dots, N\}$ is the sequence observed at the channel output, is

$$p(y_1^N | s_1^N) = \prod_{k=1}^N p(y_k | s_k), \quad (2)$$

where

$$p(y_k | s_k) = \sum_{x_k \in \mathcal{X}} p(y_k | s_k, x_k) p(x_k). \quad (3)$$

Ideal non-data-aided demodulation consists of computing the conditional probability distribution

$$\begin{aligned} p(x_k|y_1^N) &= \int_{s_k \in S_k} p(s_k, x_k|y_1^N) ds_k \\ &= \int_{s_k \in S_k} p(s_k|y_1^N) p(x_k|s_k, y_k) ds_k \\ &\propto \int_{s_k \in S_k} \frac{p(s_k|y_1^{k-1})}{p(s_k)} p(s_k|y_k^N) p(x_k|s_k, y_k) ds_k, \end{aligned} \quad (4)$$

where S_k is the state space at time k . The conditional distributions $p(s_k|y_1^{k-1})$ and $p(s_k|y_k^N)$ appearing in (4) bring in the conditioning sequence the memory of the channel. They can be worked out by forward-backward Bayesian tracking. The conditional distributions $p(s_k|y_1^{k-1})$ and $p(s_k|y_k^N)$ are often non-treatable, meaning that they cannot be expressed in a parametric form. When this happens one is forced to renounce to exact Bayesian tracking and to look for a non-parametric approximation of these distributions. The approximation that we consider is based on the popular concept of *importance sampling*, i.e.

$$p(s_0^k|y_1^k) \approx q(s_0^k|y_1^k) = \sum_{i=1}^P w_k^{(i)} \delta(s_0^k - s_0^{k(i)}), \quad (5)$$

where $\delta(\cdot)$ is the Dirac delta, P is the number of important samples, $s_0^{k(i)}$ is the i th important sample drawn from the importance distribution $g(s_0^k|y_1^k)$ and

$$w_k^{(i)} = \frac{1}{P} \cdot \frac{p(s_0^{k(i)}|y_1^k)}{g(s_0^{k(i)}|y_1^k)} \quad (6)$$

is his weight. The marginal distribution $q(s_k|y_1^k)$ is obtained by the saturation of the multivariate $q(s_0^k|y_1^k)$, hence

$$p(s_k|y_1^k) \approx q(s_k|y_1^k) = \sum_{i=1}^P w_k^{(i)} \delta(s_k - s_k^{(i)}). \quad (7)$$

The form of the importance distribution $g(s_0^N|y_1^N)$ in (6)

$$g(s_0^N|y_1^N) = g(s_0) \prod_{k=1}^N g(s_k|s_{k-1}, y_k) \quad (8)$$

leads to the *sequential* implementation of importance sampling. In *sequential* importance sampling the P important samples are called *particles* [20]. At the k th step of Bayesian tracking, particles and weights evolve as

$$s_k^{(i)} \sim g(s_k|s_{k-1}^{(i)}, y_k), \quad (9)$$

$$w_k^{(i)} \propto \frac{w_{k-1}^{(i)} p(y_k|s_k^{(i)}) p(s_k^{(i)}|s_{k-1}^{(i)})}{g(s_k^{(i)}|s_{k-1}^{(i)}, y_k)}, \quad (10)$$

leading to the k th distribution (7). For the demodulation we write

$$\begin{aligned} p(s_k|y_1^{k-1}) &= \int_{s_k} p(s_k|s_{k-1}) p(s_{k-1}|y_1^{k-1}) ds_{k-1} \\ &\approx \sum_{i=1}^P w_{k-1}^{(f,i)} p(s_k|s_{k-1}^{(f,i)}), \end{aligned} \quad (11)$$

$$p(s_k|y_k^N) \approx \sum_{l=1}^P w_k^{(b,l)} \delta(s_k - s_k^{(b,l)}), \quad (12)$$

where f and b in the superscript stand for *forward* and *backward*, respectively, and, substituted in (4), this leads to

$$p(x_k|y_1^N) \approx q(x_k|y_1^N) \propto$$

$$\sum_{i=1}^P \sum_{l=1}^P w_k^{(b,l)} w_{k-1}^{(f,i)} \frac{p(s_k^{(b,l)}|s_{k-1}^{(f,i)}) p(x_k|y_k, s_k^{(b,l)})}{p(s_k^{(b,l)})}. \quad (13)$$

In order to reduce the complexity of the above computation (13), among the several options, we highlight here two solutions that easily follow from the decoding process and will be exploited widely in the simulation results. The former, denoted as Best Particle Selection (BPS), is based on the selection of the best 2 particles, one from the forward and one from the backward recursions, before the application of (13); as a consequence, the double sum is reduced to a single term,

$$p(x_k|y_1^N) \propto \frac{p(s_k^{(b,1)}|s_{k-1}^{(f,1)}) p(x_k|y_k, s_k^{(b,1)})}{p(s_k^{(b,1)})}, \quad (14)$$

assuming the particles ordered at each step from the highest to the lowest weights. The latter, denoted as Direct State Recombination (DSR), is based on the estimate of the state, \hat{s}_k , derived from the forward and backward particles, for instance

$$\hat{s}_k = \frac{1}{2} \left(\sum_{i=1}^P w_{k-1}^{(f,i)} s_k^{(f,i)} + \sum_{i=1}^P w_k^{(b,i)} s_k^{(b,i)} \right), \quad (15)$$

which can be used to compute the symbol probabilities as if \hat{s}_k were the true hidden state:

$$p(x_k|y_1^N) \approx p(x_k|y_k, \hat{s}_k). \quad (16)$$

Let us consider now the following importance distribution

$$g(s_0) = p(s_0),$$

$$g(s_k|s_{k-1}^{(i)}, y_k) = u(\mu(s_{k-1}^{(i)}, y_k), \rho; s_k), \quad k > 0, \quad (17)$$

where $u(\mu, \rho; x)$ is an uniform distribution of the random vector x over the hypersphere of radius ρ centered on μ . With the importance distribution (17), the iteration for the weights is

$$w_k^{(i)} \propto w_{k-1}^{(i)} p(y_k|s_k^{(i)}) p(s_k^{(i)}|s_{k-1}^{(i)}) \quad (18)$$

independently of ρ . For $\rho \rightarrow 0$, the importance distribution becomes

$$g(s_k|s_{k-1}^{(i)}, y_k) = \delta(\mu(s_{k-1}^{(i)}, y_k) - s_k), \quad k > 0. \quad (19)$$

The delta-type importance distribution leads to a deterministic update of the state of particles, i.e.

$$s_k^{(i)} = \mu(s_{k-1}^{(i)}, y_k). \quad (20)$$

Importance sampling leaves us the freedom of defining the entire importance distribution. In our case, having restricted our attention to a delta-type distribution, now we have to define the deterministic function $\mu(\cdot)$. In the following we show that we take the function $\mu(\cdot)$ that makes each particle a decision-directed Kalman filter. To this aim, let us consider the state space model

$$s_k = F s_{k-1} + v_{k-1}, \quad (21)$$

$$y_k = H(x_k) s_k + n_k, \quad (22)$$

where the uppercase character denotes the matrices, v_k and n_k are jointly Gaussian independent vectors with white power spectral density and zero mean, the covariance matrix of the process noise v_k is Q , the channel noise n_k is the familiar Additive White Gaussian Noise (AWGN) with covariance matrix $\sigma^2 I$, the matrix F is the state transition matrix while the matrix $H(x_k)$ is the time-varying measurement matrix that depends here on the specific transmitted modulation symbol x_k . For instance, in the SISO

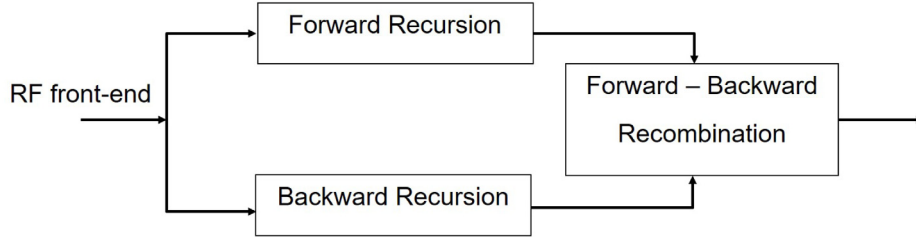


Fig. 1. Block diagram of the overall structure of the proposed algorithm.

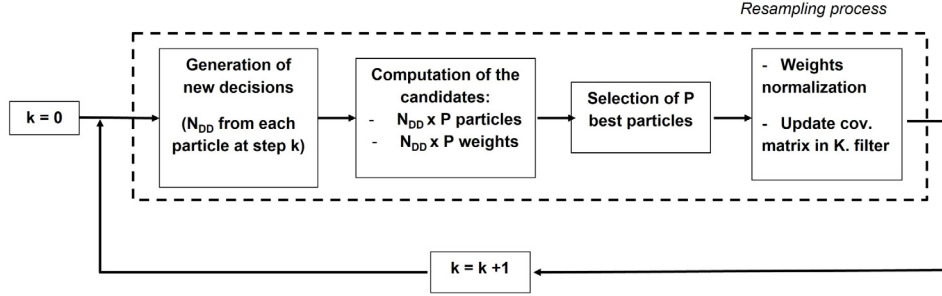


Fig. 2. Block diagram of the forward (backward) recursion.

narrow-band fading channel $H(x_k) = x_k$ and s_k is the value of the amplitude of the channel tap.

Given the state space model (21), (22), we take for the function $\mu(\cdot)$ of (20) the state update equation of the Kalman filter algorithm

$$\begin{aligned} s_k^{(i)} &= \mu(s_{k-1}^{(i)}, y_k) \\ &= F s_{k-1}^{(i)} + K_k^{(i)} (y_k - H(\hat{x}_k^{(i)}) F s_{k-1}^{(i)}), \end{aligned} \quad (23)$$

$$K_k^{(i)} = \Sigma_k^{(i)} (H(\hat{x}_k^{(i)})^H (H(\hat{x}_k^{(i)}) \Sigma_k^{(i)} (H(\hat{x}_k^{(i)})^H + \sigma^2 I)^{-1}). \quad (24)$$

Note that in the above equation each particle has its own data-dependent measurement matrix $H(\hat{x}_k^{(i)})$. This is the core of our proposed receiver, which makes use of a measurement matrix that, being based on the following local decisions

$$\hat{x}_k^{(i)} = \arg \min_{x \in \mathcal{X}} \|y_k - H(x) s_{k-1}^{(i)}\|^2, \quad (25)$$

can be different particle-by-particle.

Finally the covariance matrix of the state prediction error evolves according to

$$\Sigma_{k+1}^{(i)} = F(I - K_k^{(i)} H(\hat{x}_k^{(i)})) \Sigma_k^{(i)} F^H + Q. \quad (26)$$

3. Implementation

A practical detector based on the principles presented in Section 2 has a structure that is comparable to a BCJR detector [9], composed by forward recursion, backward recursion and recombination stages as shown in Fig. 1, with a number of states equal to the number of particles P .

The proposed detector is able to return hard estimates of the symbols or their a-posteriori likelihoods (soft detections). In detail, we can distinguish the following fundamental operations in the detector:

- *Forward and backward recursions*, generating the sets of states $s_k^{(f,l)}$ and $s_k^{(b,l)}$ and associated particle weights $w_k^{(f,l)}$ and $w_k^{(b,l)}$ for $k = 1, \dots, N$ and $l = 1, \dots, P$. Each step k of the forward (or, equivalently, backward) recursion is characterized by the following operations, as outlined in the

block diagram in Fig. 2: for each state or particle with index l , a tentative symbol $\hat{x}_k^{(l)}$ is found according to (25). In order to improve performance, this step can be extended to the identification of the most likely $N_{DD} \geq 1$ tentative symbols. In fact, this allows to have a process with $P > 1$ particles; with $P = 1$ the particle filter reduces to a decision directed Kalman filter that tracks the hidden state. In this case, it is well known that decision errors can severely affect tracking performance.

- The *resampling* process, responsible of the generation of many particles from one particle during the forward and backward recursions, as shown in Fig. 2: at each of the N_{DD} tentative symbols, it is associated a new weight $w_k^{(f,l)}$ according to (18) and the corresponding next deterministic states are computed according to (23). The final set of new $N_{DD} \times P$ states has to be reduced to P and this is obtained selecting the first P new states according to the computed weights $w_k^{(f,l)}$. In other words, for everyone of the N_{DD} decisions, one new particle is generated, its weight computed, then the weights of all the particles are ranked and, after ranking, only the P particles with higher weights are kept alive for the next step $k+1$. Finally we remark that, for each of the selected P states, the next covariance matrices are computed according to (26); in order to reduce the number of operations, the update of the covariance matrices can be avoided adopting, for all the blocks, the covariance matrix reached after a number of symbols sufficiently high for a convergence (e.g. through the use of a preamble in practice).
- The *recombination* stage, which is the most intensive part of the decoding process, reported in Fig. 3. At each step k , the a-posteriori probability of each symbol is given by (13), which includes, for each possible symbol x_k , the computation of P^2 terms composed by the 2 particle weights, 2 probabilities that do not depend on x_k and 1 depending on x_k . In the implementation, in order to reduce the number of computations, the list of symbols considered in the recombination step can be reduced to the symbols selected in the forward and backward recursions during the decision-directed process and the others are set to a zero a-posteriori probability; this option will be referred as Limited Symbol List Recombination (LSLR). We observe that the LSLR can be

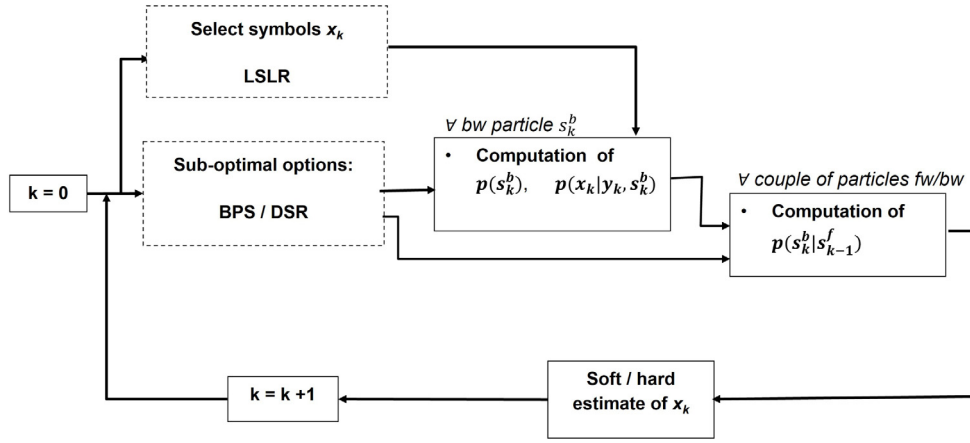


Fig. 3. Block diagram of the recombination stage. The dashed blocks are optional.

Table 1

Computational complexity of the recombination stage in the decoding process: Number of operations per symbol in the log-domain.

	(LSLR)	Sums	LogSumExp	Euclidean distances
Full complexity		$P^2(3 + N_X)$	$P - 1 + N_X(P^2 - 1)$	$P(1 + N_X + P)$
	+ LSLR	$P^2(3 + 2N_{DD})$	$P - 1 + 2N_{DD}(P^2 - 1)$	$P(1 + 2N_{DD} + P)$
DSR		$2P$	$3P - 2$	N_X
	+ LSLR	$2P$	$3P - 2$	$2N_{DD}$
BPS		$(3 + N_X)$	$P - 1$	$(2 + N_X)$
	+ LSLR	$(3 + 2N_{DD})$	$P - 1$	$(2 + 2N_{DD})$

used jointly with the BPS and DSR simplification options, (14) and (15). In Table 1, we summarize the number of operations per symbol in the log-domain for the recombination stage of the full and reduced complexity decoders; as the probabilities are Gaussian, the complexity is reduced to the computation of a weighted Euclidean distance in the log-domain, as reported in the last column of the table.

We note that the evaluation of the overall computational complexity must take into account also the computation of the forward and backward parts with the resampling process, with a corresponding number of operations (sums and Euclidean distances) proportional to $2P \times N_{DD}$, and the computations involved in the steps (23), (24), (26), whose exact number of operations depends on the specific state transition and measurement matrices F and H of the considered process.

4. Application to the MIMO LoS phase noise channel

The proposed method is hereafter applied to the 2×2 phase noise MIMO Line of Sight (LoS) channel with independent oscillators affected by Wiener phase noise. We assume that the state vector evolves over time as a discrete-time vector Wiener process

$$s_{k+1} = s_k + q_k,$$

where q_k is the k th sample of a multi-dimensional white Gaussian noise process with covariance matrix $Q = \gamma^2 I$ and zero mean vector. The reader can find additional details about the channel model in [7], where the MIMO phase noise is linearized and tracked by a data-aided linearized Kalman. In place of using transmitted symbols in the Kalman filter algorithm as in [7], here each one of the Kalman particles (filters) uses the decision that it takes based on its state.

There is an issue regarding the specific case of the MIMO phase noise channel that has been not considered in [7], whose consideration is necessary in order to perform the operations in

Section 2 and describe exhaustively the process for obtaining the numerical results in Section 5. In the 2×2 MIMO case, what our algorithm tracks is the state vector

$$s = (\phi_{t;1}, \phi_{t;2}, \phi_{r;1}, \phi_{r;2})^T,$$

where $\phi_{t;1}, \phi_{t;2}$ are the phases of the two transmitters, $\phi_{r;1}, \phi_{r;2}$, are the phases of the two receivers and the time index is omitted here for the sake of brevity. The output of the channel contains the following vector

$$\phi = (\phi_{t;1} + \phi_{r;1}, \phi_{t;2} + \phi_{r;1}, \phi_{t;1} + \phi_{r;2}, \phi_{t;2} + \phi_{r;2})^T. \quad (27)$$

Writing the output phase vector in the matrix-vector form

$$\phi = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \cdot s, \quad (28)$$

one immediately realizes that the matrix in (28) has rank 3, i.e. there is one degree of freedom in the determination of s , given a generic realization of ϕ , as can be seen also in [21]. Specifically, it is easy to see that, for a generic s , also

$$s_{\Delta} = s + \Delta \cdot (1, 1, -1, -1)^T, \quad (29)$$

provides the same ϕ after the application of (28). In other words, the direction in the s -space associated with the null eigenvalue of the matrix is the one spanned by the vector $v = (1, 1, -1, -1)^T$ and consequently:

- Vector s and vector

$$s + \Delta \cdot (1, 1, -1, -1)^T,$$

$\forall \Delta$, are indistinguishable.

Besides this ambiguity, there are the following other two ambiguities that are due to the periodicity of 2π of the phase:

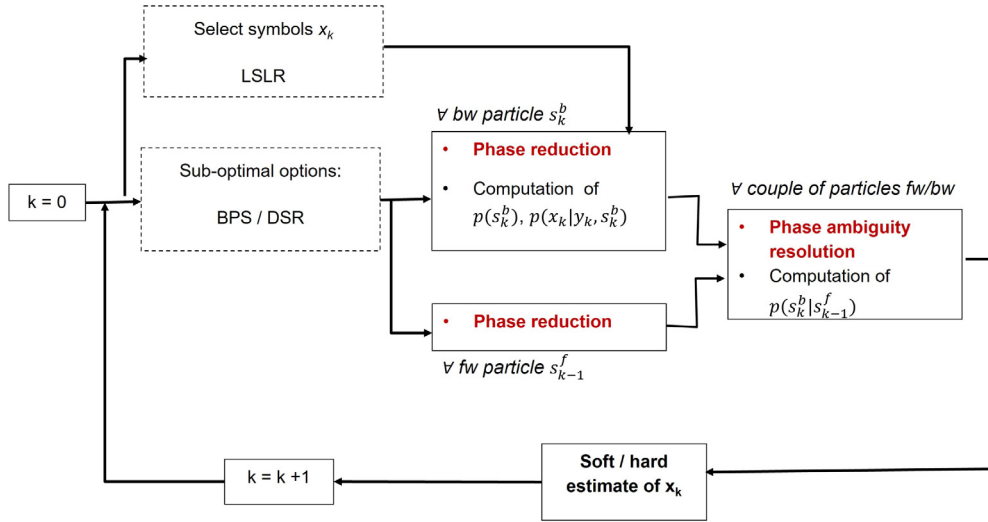


Fig. 4. Block diagram of the recombination stage when the particles contain phase values. W.r.t. Fig. 3, the bold red characters denote the additional steps for the resolution of phase ambiguities.

- Vector s and vector

$$s + k\pi \cdot (1, 1, 1, 1)^T,$$

$\forall k$ integer, are indistinguishable.

- Vector s and vector

$$s + 2\pi \cdot (k_1, k_2, k_3, k_4)^T,$$

$\forall (k_1, k_2, k_3, k_4)$ integers, are indistinguishable.

All these ambiguities must be considered and solved reliably in the forward-backward recombination of two states containing the MIMO phase information, in particular in the computation of $p(s_k^{(b,l)} | s_{k-1}^{(f,i)})$ in (13). Our approach is to find, among all the ambiguous pairs, the pair $(s_k^{(b,l)}, s_{k-1}^{(f,i)})$ that minimizes the Euclidean distance between them and to use only this pair in the recombination. In the practical implementation, this specific step is achieved by means of two steps, highlighted in bold red characters in the blocks of Fig. 4:

- In each state vector s from the forward or backward recursions the component along the space direction associated with the vector $v = (1, 1, -1, -1)^T$ is nullified, i.e.:

$$s_0 = s - \langle s, v \rangle v, \quad (30)$$

where the operator $\langle a, b \rangle$, defined as $\langle a, b \rangle = a^T b$, denotes the scalar product between two real vectors a and b . Then the elements of s_0 are expressed in the interval $[-\pi, \pi]$. This operation is referred as *phase reduction* in the blocks of Fig. 4.

- Each couple of forward and backward phase vectors is subject to the following tests for the *phase ambiguity resolution*: (i) from the forward (or the backward) s_0 the three vectors $[s_{0-} = s_0 - (1, 1, 1, 1)^T, s_0, s_{0+} = s_0 + (1, 1, 1, 1)^T]$ are generated, (ii) each element of the vectors $[s_{0-}, s_0, s_{0+}]$ is summed to the values $[-2\pi, 0, 2\pi]$ and the closest to the corresponding value of the backward one is selected, (iii) among the three resulting vectors, the one with the minimum Euclidean distance w.r.t. the backward one is selected.

5. Simulation results

Here we report the simulation results of the proposed method for the 2×2 MIMO channel with independent oscillators affected by Wiener phase noise, comparing our results with those of the pilot-aided linearized Kalman filter approach of [7].

The simulations are performed with symbols x_k from a 64-QAM constellation, i.e. $N_X = 64$, and two examples of 2×2 MIMO LoS channel described in Section 4: the MIMO channel A, with matrix M_A characterized by the two singular values $\lambda_1 = \lambda_2 = \sqrt{2}$,

$$M_A = \begin{bmatrix} 1 & -j \\ -j & 1 \end{bmatrix}, \quad (31)$$

and the MIMO channel B, with matrix M_B characterized by the two singular values $\lambda_1 = \sqrt{3.93}$, $\lambda_2 = \sqrt{0.068}$,

$$M_B = \begin{bmatrix} 1 & e^{-j\pi/12} \\ e^{-j\pi/12} & 1 \end{bmatrix}. \quad (32)$$

Fig. 5 reports an example which exemplifies the advantage provided by the non-parametric method in tracking a possibly multi-modal posterior distribution. The figure shows a Wiener realization of phase noise (dashed line) with power $\gamma^2 = -38$ dB and the evolution of the posterior distribution inferred by the forward particle weights (colored tones), for the channel matrix M_B (32), 64-QAM modulation, i.e. $N_X = 64$, $E_S/N_0 = 31$ dB and $P = 64$ particles.

Coming to the Symbol Error Rate (SER) results, the Monte Carlo simulations are performed using blocks of 2×1282 symbols, passed through the MIMO channel with phase noise and AWGN till to the accumulation of at least 200 symbol errors or a maximum number of blocks equal to 400. In order to obtain more fair and comparable results from the different tracking techniques, at each value of γ^2 the results of the different solutions are simulated with the same realization of phase noise: a common set of 1000 vectors of phase noise samples at the transmitter and 1000 vectors at the receiver is used, scaled with the correct power γ^2 . Each figure reports the results obtained by the different methods: the channel tracking and demodulation with the full complexity particle filter (13) with $P = 16$ and $P = 1$, referred as 'K. PF, $P = 16$ ' and 'K. PF, $P = 1$ ' in the legend, the DSR option (14) with $P = 16$, referred as 'K. PF, $P = 16$,

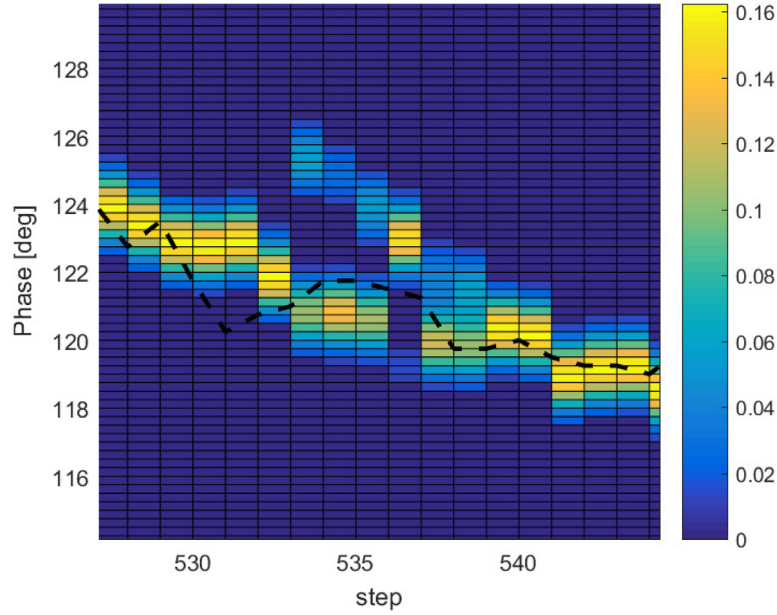


Fig. 5. Example of the time evolution of the posterior distribution that shows multi-modality. The dashed line is the time evolution of the hidden free running channel state. The colored tones represent the posterior distribution, from less probable (dark blue) to more probable (yellow). Around step $k = 535$ the posterior distribution is bi-modal. Details about the system model and about how the posterior distribution has been inferred can be found in Section 5.

DSR' and the BPS (15) (16) with $P = 16$, referred as 'K. PF, $P = 16$, BPS'. In addition, referred as 'FW-BW K. + pilots', the results when the MIMO phase noise is tracked by a data-aided linearized Kalman filter [7] for a pilot symbols rate 1/10 are also reported, taking into account, for a fair comparison, the energy of the pilot symbols in the computation of the same E_S/N_0 . In order to limit the complexity and the corresponding simulation time, the option LSLR has been adopted for all the simulations; from tests made in subsets of the reported SNRs and phase noise powers, this option does not affect significantly the final SER w.r.t. the full complexity detectors.

Fig. 6 reports a preliminary study on the impact of the parameter N_{DD} , responsible for the multiple decisions process; with the channel matrix M_B (32), a final $N_{DD} = 4$, with a satisfactory trade-off between computational load and performance, has been selected for the rest of the simulations. Similar results have been obtained for the channel matrix M_A (31) as well.

Fig. 7 shows the SER as a function of γ^2 for the channel matrix M_A and $E_S/N_0 = \{19, 23\}$ dB. Similarly, Fig. 8 shows the SER as a function of γ^2 for the channel matrix M_B and $E_S/N_0 = \{31, 35\}$ dB. From the simulation results, we can observe that:

- for phase noise power levels not too severe, depending on the channel matrix and on the value of E_S/N_0 , i.e. $\gamma^2 < -34$ dB for the channel matrix M_A and $E_S/N_0 = 19$ and $\gamma^2 < -36$ dB for the channel matrix M_B , decision-directed Kalman particle filter implementations with $P = 16$, either full or reduced complexity solutions, outperform the tracking by the data-aided linearized Kalman filter;
- for a particle filter with $P = 16$ the reduced complexity implementation DSR shows approximately the same performance as the full complexity solution;
- at decreasing values of the phase noise power, even channel tracking and demodulation by the particle filter with only $P = 1$, that approximates the performance achievable by the reduced complexity implementation BPS with $P = 16$, shows better results than data-aided Kalman filter;
- as the E_S/N_0 increases and the phase noise power level reduces, the advantage of the proposed schemes tends to

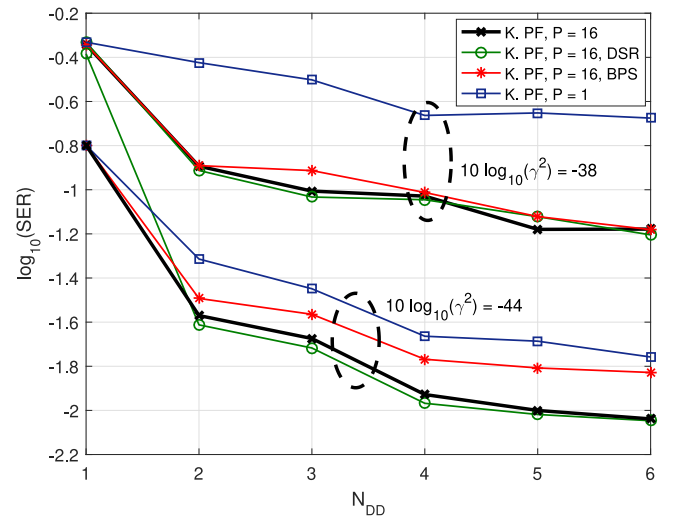


Fig. 6. Symbol Error Rate for the channel matrix M_B (32), 64-QAM, $E_S/N_0 = 31$ dB, phase noise power values $\gamma^2 = \{-38, -44\}$ dB as a function of N_{DD} .

increase w.r.t. the data-aided approaches since the decisions tend to be all or almost all correct and the system exploits this advantage for improving the phase noise tracking with symbol-by-symbol updates.

These considerations highlight how the proposed reduced complexity implementations of decision-directed Kalman particle filter outperform data-aided linearized Kalman filter technique as, even if data-aided techniques take advantage from the knowledge of transmitted pilot symbols, they ignore the information coming from the samples between two consecutive pilots, while the proposed non-data-aided techniques reveal the capability of exploiting the information from all the samples.

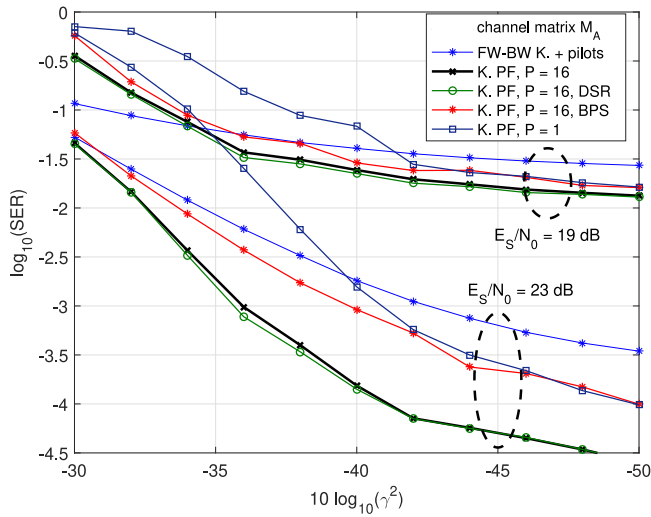


Fig. 7. Symbol Error Rate for the channel matrix M_A (31), 64-QAM and $E_S/N_0 = \{19, 23\}$ dB as a function of the phase noise power γ^2 in dB.

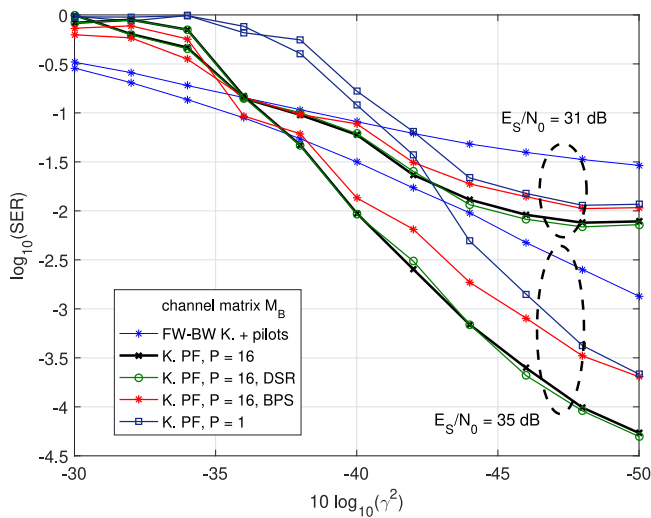


Fig. 8. Symbol Error Rate for the channel matrix M_B (32), 64-QAM and $E_S/N_0 = \{31, 35\}$ dB as a function of the phase noise power γ^2 in dB.

6. Conclusions

The paper proposes a non-data-aided method based on sequential importance sampling (particle filtering) for tracking the time variation of a hidden continuous channel state vector. W.r.t. pilot-aided methods, the non-data-aided feature allows substantial savings in the transmission rate preserving a competitive performance, as shown in the relevant example of the MIMO phase noise channel. The reason behind the improved performance is that our proposed method makes use of all the received signal samples for extracting information about the hidden channel state vector. An important feature of the proposed method is that it can track a multi-modal multi-dimensional posterior distribution of the hidden channel state: what happens when the posterior distribution becomes multi-modal is that the set of particles splits into subsets, which track the different modes. Finally, the proposed method can take benefit from a forward-backward operation, even if the complexity of the recombination stage can become large; sub-optimal methods for reducing the complexity of the recombination of the forward and backward parts, which perform as the full complexity one at least in the

MIMO phase noise example, have been proposed and validated by means of extensive simulations. Finally, we point out that the proposed methodology, here applied to the MIMO channel with phase noise, can be applied to other channel models as, for instance, those characterized by time-variant fading.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] C. Xiantao, K. Xu, S. Li, Compensation of phase noise in uplink massive MIMO OFDM systems, *IEEE Trans. Wireless Commun.* 18 (3) (2019) 1764–1778, <http://dx.doi.org/10.1109/TWC.2019.2897089>.
- [2] X. Liu, W. Wan, X. Song, X. Gao, G. FettweisLiu, Sparse channel estimation via hierarchical hybrid message passing for massive MIMO-OFDM systems, *Accept. Publ. IEEE Trans. Wirel. Commun.* (2021).
- [3] C. Kominakis, C. Fragouli, A.H. Sayed, R.D. Wesel, Multi-input multi-output fading channel tracking and equalization using Kalman estimation, *IEEE Trans. Signal Process.* 50 (5) (2002) 1065–1076, <http://dx.doi.org/10.1109/78.995063>.
- [4] W. Chen, R. Zhang, Kalman-filter channel estimator for OFDM systems in time and frequency-selective fading environment, in: 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, QC, Canada, 2004, p. iv, <http://dx.doi.org/10.1109/ICASSP.2004.1326842>.
- [5] H. Mehrpouyan, A.A. Nasir, S.D. Blostein, T. Eriksson, G.K. Karagiannidis, T. Swenson, Joint estimation of channel and oscillator phase noise in MIMO systems, *IEEE Trans. Signal Process.* 60 (9) (2012) 4790–4807.
- [6] A. Spalvieri, L. Barletta, Pilot-aided carrier recovery in the presence of phase noise, *IEEE Trans. Commun.* 59 (7) (2011) 1966–1974, <http://dx.doi.org/10.1109/TCOMM.2011.051311.100047>.
- [7] L. Reggiani, L. Dossi, L. Barletta, A. Spalvieri, Extended Kalman filter for MIMO phase noise channels with independent oscillators, *IEEE Commun. Lett.* 22 (6) (2018) 1200–1203, <http://dx.doi.org/10.1109/LCOMM.2018.2791590>.
- [8] L. Barletta, M. Magarini, A. Spalvieri, Bridging the gap between Kalman filter and Wiener filter in carrier phase tracking, *IEEE Photonics Technol. Lett.* 25 (11) (2013) 1035–1038, <http://dx.doi.org/10.1109/LPT.2013.2259476>.
- [9] L.R. Bahl, J. Cocke, F. Jelinek, J. Raviv, Optimal decoding of linear codes for minimizing symbol error rate, *IEEE Trans. Inf. Theory* IT-20 (2) (1974) 284–287.
- [10] S. Galan, M. Peleg, S. Shamai, On iterative phase trellis based noncoherent detection of coded MPSK in a noisy phase regime, in: MELECON '98, 9th Mediterranean Electrotechnical Conference Proceedings (Cat. No. 98CH36056), Tel-Aviv, Israel, 1998, pp. 834–838, <http://dx.doi.org/10.1109/MELCON.1998.699342>, 2.
- [11] S. Pecorino, S. Mandelli, L. Barletta, M. Magarini, A. Spalvieri, Bootstrapping iterative demodulation and decoding without pilot symbols, *J. Lightwave Technol.* 33 (17) (2015) 3613–3622, <http://dx.doi.org/10.1109/JLT.2015.2448109>.
- [12] K. Huber, S. Haykin, Application of particle filters to MIMO wireless communications, in: IEEE International Conference on Communications ICC '03, vol. 4, Anchorage, AK, USA, 2003, pp. 2311–2315, <http://dx.doi.org/10.1109/ICC.2003.1204296>.
- [13] W.H. Chin, D.B. Ward, A.G. Constantinides, Semi-blind MIMO channel tracking using auxiliary particle filtering, in: IEEE Global Telecommunications Conference GLOBECOM '02, Vol. 1, Taipei, Taiwan, 2002, pp. 322–325, <http://dx.doi.org/10.1109/GLOCOM.2002.1188093>.
- [14] J.H. Kotecha, P.M. Djuric, Gaussian sum particle filtering, *IEEE Trans. Signal Process.* 51 (10) (2003) 2602–2612, <http://dx.doi.org/10.1109/TSP.2003.816754>.
- [15] L. Barletta, M. Magarini, S. Pecorino, A. Spalvieri, Upper and lower bounds to the information rate transferred through first-order Markov channels with free-running continuous state, *IEEE Trans. Inform. Theory* 60 (7) (2014) 3834–3844.
- [16] J. Miguez, P.M. Djuric, Blind equalization of frequency-selective channels by sequential importance sampling, *IEEE Trans. Signal Process.* 52 (10) (2004) 2738–2748, <http://dx.doi.org/10.1109/TSP.2004.834>.
- [17] G. Evensen, The ensemble Kalman filter: theoretical formulation and practical implementation, *Ocean Dyn.* 53 (2003) 343–367.
- [18] A.S. Stordal, H.A. Karlsen, G. Nævdal, H.J. Skaug, B. Vallès, Bridging the ensemble Kalman filter and particle filters: the adaptive Gaussian mixture filter, *Comput. Geosci.* 15 (2) (2011) 293–305, <http://dx.doi.org/10.1007/s10596-010-9207-1>.

- [19] J.W.M. Bergmans, J.O. Voorman, H.W. Wong-Lam, Dual decision feedback equalizer, *IEEE Trans. Commun.* 45 (5) (1997) 514–518, <http://dx.doi.org/10.1109/26.592548>.
- [20] V. Syrjälä, T. Levanen, T. Ihalainen, M. Valkama, A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, *IEEE Trans. Signal Processing* 50 (2) (2002) 174–188.
- [21] A. Tarable, G. Montorsi, S. Benedetto, S. Chinnici, An EM-based phase-noise estimator for MIMO systems, in: *IEEE International Conference on Communications, ICC'13, Budapest, Hungary, 2013*, pp. 3215–3219.



Arnaldo Spalvieri received his degree in Electronic Engineering from the University of Ancona (Italy) in 1985. From 1986 to 1992 he was in the modem lab. of Telettra (now Alcatel-Lucent), working at the design of modems for high capacity digital radio systems. From 1992 to 1998 he was with the Dipartimento di Elettronica e Informazione at Politecnico di Milano as an assistant professor, and from 1998 he was appointed associate professor. His major research activity is in the field of digital transmission systems for telecommunications. From 1992 he has been a consultant of the industry, including Alcatel-Lucent, STMicroelectronics, Andrew, Carlo Gavazzi Space. In 2006 he co-founded the company Binary Core, a spinoff of Politecnico di Milano active in the area of design on FPGA of mo-demodulators for point-to-point terrestrial radio and for digital video broadcasting.



Luca Reggiani Luca Reggiani received the PhD in Electronics and Communications Engineering in 2001 from Politecnico di Milano (Italy). He has collaborated with several industries and Universities in the field of wireless communications and magnetic recording, as a consultant or within Italian and European research programs. He is a co-founder of two startup companies in the ICT field and his research interests include mobile systems, wireless sensor networks, UAV applications, information theory.



Laura Dossi received the degree in Electronic Engineering from Politecnico di Milano (Italy) in 1985. She had been with IBM in Vimercate, Milano, for two years, and then with Telettra, Telecommunication Central Laboratory, Milano. Since June 1988 she has been a researcher at the Consiglio Nazionale delle Ricerche (CNR) at the Istituto di Elettronica e Ingegneria Informatica e delle Telecomunicazioni (IEIIT) c/o Politecnico di Milano. From 1998 she has taught as a lecturer at the Politecnico di Milano. Her research interests concern radio channel modeling, wireless communications, mmwave propagation.