Improved Iterative MIMO Signal Detection Accounting for Channel-Estimation Errors

Seyed Mohammad-Sajad Sadough, *Member, IEEE*, Mohammad-Ali Khalighi, *Senior Member, IEEE*, and Pierre Duhamel, *Fellow, IEEE*

***Abstract*—For the case of imperfect channel estimation, we propose an improved detector for multiple-input–multiple-output (MIMO) systems using an iterative receiver. We consider two iterative detectors based on maximum *a posteriori* and soft parallel-interference cancellation and propose, for each case, modifications to the MIMO detector by taking into account the channel-estimation errors. Considering training-based channel estimation, we present some numerical results for the case of Rayleigh block fading, which demonstrate an interesting performance improvement, compared with the classically used *mismatched* detectors. This improvement is obtained without a considerable increase in receiver complexity. It is more advantageous when few pilots are used for channel estimation. Thus, the proposed improved detectors are of particular interest under relatively fast-fading conditions, where it is important to reduce the number of pilots to avoid a significant loss in data rate.**

***Index Terms*—Channel-estimation errors, improved maximum *a posteriori* (MAP) detection, mismatched detection, multiple-input–multiple-output (MIMO), soft parallel-interference cancellation (soft-PIC) detection.**

NOTATION

Upper case bold symbols	matrices;
lower case bold symbols	vectors;
$\mathbb{I}_N$	$N \times N$ identity matrix;
$\mathbb{E}_{\mathbf{x}}[.]$	expectation with respect to the random vector $\mathbf{x}$;
$\det\{.\}$	matrix determinant;
$\mathrm{tr}\{.\}$	matrix trace;
$\lvert . \rvert$	absolute value;
$\lVert . \rVert$	Frobenius norm;
$(.)^T$	vector transpose;
$(.)^\dagger$	Hermitian transpose;
$(.)^*$	complex conjugate.

Manuscript received January 27, 2008; revised October 28, 2008 and January 17, 2009. First published February 27, 2009; current version published August 14, 2009. The review of this paper was coordinated by Prof. E. Bonek.

S. M.-S. Sadough is with the Department of Telecommunications, Faculty of Electrical and Computer Engineering, Shahid Beheshti University, Tehran 1983963113, Iran (e-mail: s_sadough@sbu.ac.ir; sajad.sadough@gmail.com).

M.-A. Khalighi is with the École Centrale Marseille, 13451 Marseille Cedex 20, France, and also with Institut Fresnel, 13397 Marseille Cedex 20, France.

P. Duhamel is with the National Scientific Research Center/Laboratory of Signals and Systems–Supélec, 91192 Gif sur Yvette Cedex, France.

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

Digital Object Identifier 10.1109/TVT.2009.2016103

I. INTRODUCTION

IT is well known that multiple-input–multiple-output (MIMO) systems can provide very high spectral efficiencies in a rich scattering propagation medium [1]. They are, hence, a promising solution for high-speed spectrally efficient reliable wireless communication. One important issue regarding the implementation of MIMO systems is the channel estimation when coherent signal detection is performed at the receiver [2]. To obtain the channel-state information at the receiver (CSIR), one usually used approach consists of sending some known training (also called pilot) symbols from the transmitter, based on which the receiver estimates the channel before proceeding to the detection of data symbols. This method of obtaining the CSIR is usually called pilot-symbol-assisted modulation (PSAM) [3]. Obviously, due to the finite number of pilot symbols and noise, in practice, the receiver can only obtain an *imperfect* (and possibly very bad) estimate of the channel.

A rich literature exists on the impact of imperfect channel estimation on the performance of communication systems that employ multiple antennas. For a MIMO system that uses PSAM for channel estimation, Garg *et al.* showed in [4] that, to compensate for the performance degradation due to imperfect channel estimation, the number of receive antennas should be increased. In [5], the authors investigated the effect of imperfect channel estimation on space-time (ST) decoding and showed that the classical maximum likelihood (ML) detector, which was derived for the case of perfect CSIR, becomes largely suboptimal in the presence of channel-estimation errors. One similar investigation was carried out in [6] and [7] in the case of multicarrier systems based on orthogonal frequency-division multiplexing (OFDM).

To deal with imperfect channel estimation, one suboptimal approach, which is known as *mismatched* detection, consists of using the channel estimate in the detection part as if it was a perfect estimate. It is shown, for instance, in [8] that this scheme greatly degrades the detection performance in the presence of channel-estimation errors.

As an alternative to the aforementioned mismatched approach, Tarokh *et al.* [8] and, recently, Taricco and Biglieri [5] proposed an *improved* ML detection metric under imperfect CSIR and used it in the standard Viterbi algorithm. Recently, in [9], it has been shown that, compared with the mismatched ML, the improved ML metric can increase the achievable outage rates of MIMO-OFDM systems, in particular when only a few training symbols are devoted for channel estimation. Similar observations are reported in [10]–[12] for the case of single-antenna OFDM systems.

Here, we consider iterative (turbo) detection at the receiver, which is an efficient technique when channel coding is used [13], [14]. This scheme is essentially composed of a MIMO detector (also called a demapper) and a soft-input-soft-output (SISO) channel decoder that exchange soft information with each other through several iterations.

One practical concern for the implementation of turbo detectors for MIMO systems is the receiver complexity, which is mainly dominated by that of the MIMO demapper. For instance, in the case of maximum *a posteriori* (MAP) detection [15], which is the optimal solution under perfect CSIR in the sense of bit error rate, this complexity exponentially grows with the number of transmit antennas and the signal constellation size. Thus, suboptimal detection techniques are often preferred to MAP detection. One interesting suboptimal detector is that based on soft parallel-interference cancellation (soft-PIC) and linear minimum mean-square error (MMSE) filtering. This scheme was first proposed in [16] in the context of multiuser detection and was later applied to MIMO systems in [17] and [18], for instance.

Our aim in this paper is to propose an improved iterative detector that takes into account the imperfect channel estimation obtained via training sequences, i.e., using the PSAM approach. To this end, we propose a Bayesian framework based on the *a posteriori* probability density function (pdf) of the perfect channel, which is conditioned on its estimate. This general framework enables us to formulate any detector by considering the average, over the channel uncertainty, of the detector's cost function that would be applied in the case of perfect channel knowledge. As we will see, the improved ML metric in [5] becomes a special case of our general framework.

First, by properly modifying the soft values at the output of the MAP detector, we reduce the impact of channel uncertainty on the SISO decoder performance. Second, we propose an improved soft-PIC detector by taking into account the channel-estimation errors in the formulation of the instantaneous linear MMSE filter, as well as in the interference cancellation part.

The outline of this paper is given as follows. In Section II, we describe our MIMO channel model and our main assumptions concerning data transmission and channel estimation. In Section III, we introduce a general Bayesian framework for improved detection under imperfect channel estimation. In Section IV, we provide the formulation of MAP and soft-PIC detectors in the case of perfect CSIR. Using this material in Section V, we propose improved detectors under imperfect CSIR for the two cases of turbo-MAP and turbo-PIC detection. Section VI illustrates, via simulations, a comparative performance study of the proposed detectors with the classical mismatched detector. The convergence analysis of turbo-PIC based on mismatched and improved soft-PIC detectors is provided in Section VII. Finally, Section VIII concludes this paper.

II. System Model and Channel Estimation

A. MIMO Fading Channel

We consider a single-user MIMO system with M_T transmit and M_R receive antennas $(M_R \geq M_T)$, which transmits over a frequency-nonselective channel, and we refer to it as an $M_T \times M_R$ MIMO channel. At the transmitter, we employ the bit-interleaved coded modulation (BICM) scheme, which is known to be a simple efficient method for exploiting channel time selectivity [19]. The binary data sequence is encoded by a nonrecursive nonsystematic convolutional (NRNSC) code before being interleaved by a quasirandom interleaver. The output bits are gathered in subsequences of B bits and are mapped to complex multilevel quadrature amplitude modulation $(\mathrm{M} = 2^B)$ symbols s before being passed to the ST encoder.

In what follows, unless otherwise mentioned, we consider the ST coding in its simplest form, i.e., just spatial multiplexing of data symbols, which is also known as the Vertical Bell Laboratories Layered Space-Time (V-BLAST) scheme. We present our model for this simple case, and we shall generalize it later to the case of an arbitrary ST code.

Let us denote by $\boldsymbol{x}_k$ the $M_T \times 1$ vector of transmitted data symbols (after the ST encoder) at a sampling time k. Considering the simple spatial multiplexing, we use $\boldsymbol{s}_k$ instead of $\boldsymbol{x}_k$. Assuming a frame of symbols that correspond to L channel uses, which are transmitted over the channel matrix $\mathbf{H}$, the received signal vector $\boldsymbol{y}_k$ of dimension $M_R \times 1$ is given by

$$\boldsymbol{y}_k = \mathbf{H}\boldsymbol{s}_k + \boldsymbol{z}_k, \quad k = 1, \ldots, L \tag{1}$$

where $\boldsymbol{s}_k$ is the $M_T \times 1$ vector of transmitted symbol with average power $E_s \triangleq (1/M_T)\mathbb{E}[\mathrm{tr}\{\boldsymbol{s}_k \boldsymbol{s}_k^\dagger\}]$. We assume that the entries $\{\mathbf{H}\}_{i,j}$ of the random matrix $\mathbf{H}$ are independent identically distributed (i.i.d.) zero-mean circularly symmetric complex Gaussian (ZMCSCG) random variables. Thus, the channel matrix $\mathbf{H}$ is distributed as $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbb{I}_{M_T} \otimes \boldsymbol{\Sigma}_H)$, where

$$\mathcal{CN}(\mathbf{0}, \mathbb{I}_{M_T} \otimes \boldsymbol{\Sigma}_H) = \frac{1}{\pi^{M_R M_T} \det\{\boldsymbol{\Sigma}_H\}^{M_T}} \exp\left\{-\mathrm{tr}\left\{\mathbf{H}\boldsymbol{\Sigma}_H^{-1}\mathbf{H}^\dagger\right\}\right\}. \tag{2}$$

Here, $\mathcal{CN}(.)$ denotes complex Gaussian distribution, $\otimes$ stands for the Kronecker product, and $\boldsymbol{\Sigma}_H$ is the $M_R \times M_R$ covariance matrix of the rows of $\mathbf{H}$. With our assumptions of i.i.d. channel, $\boldsymbol{\Sigma}_H$ is a diagonal matrix with equal diagonal entries σ_h^2. The noise vector $\boldsymbol{z}_k$ is assumed to be a ZMCSCG random vector with covariance matrix $\boldsymbol{\Sigma}_z \triangleq \mathbb{E}(\boldsymbol{z}_k \boldsymbol{z}_k^\dagger) = \sigma_z^2 \mathbb{I}_{M_R}$. Both $\mathbf{H}$ and $\boldsymbol{z}_k$ are assumed to be ergodic and stationary random processes.

We consider the block fading model for the channel time variations. Thus, channel coefficients are assumed to be constant during a block of symbols and change to new independent values from one block to another. We assume that each frame of data symbols corresponds to N_c independent fading blocks. Notice that $N_c = 1$ returns to the quasistatic channel model.

B. Pilot-Based Channel Estimation

To estimate the MIMO transmission channel matrix $\mathbf{H}$ at the receiver, which corresponds to each fading block, we send a number of pilot symbols in addition to the ST-encoded data symbols. We devote a number of N_P channel uses to the transmission of pilot vectors $\boldsymbol{s}_{P,i}$, $(i = 1, \ldots, N_P)$ for each one of N_c fading blocks. Considering a given fading block, let us constitute the $(M_T \times N_P)$ matrix $\mathbf{S}_P$ by stacking in its columns the pilot vectors, i.e., $\mathbf{S}_P = [\boldsymbol{s}_{P,1}|, \ldots, |\boldsymbol{s}_{P,N_P}]$. According to (2), during a given channel training interval, we receive

$$\mathbf{Y}_P = \mathbf{H}\mathbf{S}_P + \mathbf{Z}_P. \tag{3}$$

The definition of $\mathbf{Y}_P$ and $\mathbf{Z}_P$ is similar to that of $\mathbf{S}_P$. We denote by E_P the average power of the training symbols defined as

$$E_P \triangleq \frac{1}{N_P M_T} \mathrm{tr}\left\{\mathbf{S}_P \mathbf{S}_P^\dagger\right\}. \tag{4}$$

The least squares estimate of $\mathbf{H}$ is obtained by minimizing $\|\mathbf{Y}_P - \mathbf{H}\mathbf{S}_P\|^2$ with respect to $\mathbf{H}$, which coincides here with the ML estimate [20], i.e.,

$$\widehat{\mathbf{H}}_{\mathrm{ML}} = \mathbf{Y}_P \mathbf{S}_P^\dagger \left(\mathbf{S}_P \mathbf{S}_P^\dagger\right)^{-1}. \tag{5}$$

Let us denote by $\boldsymbol{\mathcal{E}}$ the matrix of the estimation errors. We have

$$\widehat{\mathbf{H}}_{\mathrm{ML}} = \mathbf{H} + \boldsymbol{\mathcal{E}}, \quad \text{with} \quad \boldsymbol{\mathcal{E}} = \mathbf{Z}_P \mathbf{S}_P^\dagger \left(\mathbf{S}_P \mathbf{S}_P^\dagger\right)^{-1}. \tag{6}$$

It is known that the best channel estimate is obtained with mutually orthogonal training sequences, which result in uncorrelated estimation errors. In other words, we should choose $\mathbf{S}_P$ with orthogonal rows such that

$$\mathbf{S}_P \mathbf{S}_P^\dagger = N_P E_P \mathbb{I}_{M_T}. \tag{7}$$

This way, the jth column $\mathcal{E}_j$ of $\boldsymbol{\mathcal{E}}$ has the covariance matrix $\boldsymbol{\Sigma}_{\mathcal{E}}$, i.e.,

$$\boldsymbol{\Sigma}_{\mathcal{E}} = \mathbb{E}\left[\boldsymbol{\mathcal{E}}_j\, \boldsymbol{\mathcal{E}}_j^\dagger\right] = \sigma_{\mathcal{E}}^2\, \mathbb{I}_{M_R}, \quad \text{where } \sigma_{\mathcal{E}}^2 = \frac{\sigma_z^2}{N_P E_P}. \tag{8}$$

Thus, based on (6), the conditional pdf of $\widehat{\mathbf{H}}_{\mathrm{ML}}$, given $\mathbf{H}$, can easily be expressed as

$$p(\widehat{\mathbf{H}}_{\mathrm{ML}}|\mathbf{H}) = \mathcal{CN}(\mathbf{H}, \mathbb{I}_{M_T} \otimes \boldsymbol{\Sigma}_{\mathcal{E}}) \tag{9}$$

where $\mathcal{CN}(\boldsymbol{\Lambda}, \boldsymbol{\Xi})$ denotes a complex Gaussian distribution with mean $\boldsymbol{\Lambda}$ and the covariance matrix $\boldsymbol{\Xi}$. Using the pdf of $\mathbf{H}$ and $(\widehat{\mathbf{H}}_{\mathrm{ML}}|\mathbf{H})$ from (2) and (9), we can derive the *posterior* distribution of the perfect channel matrix, conditioned on its ML estimate, as follows (see Appendix I-A):

$$p(\mathbf{H}|\widehat{\mathbf{H}}_{\mathrm{ML}}) = \mathcal{CN}(\boldsymbol{\Sigma}_\Delta \widehat{\mathbf{H}}_{\mathrm{ML}}, \mathbb{I}_{M_T} \otimes \boldsymbol{\Sigma}_\Delta \boldsymbol{\Sigma}_{\mathcal{E}}) \tag{10}$$

where

$$\boldsymbol{\Sigma}_\Delta = \boldsymbol{\Sigma}_H (\boldsymbol{\Sigma}_{\mathcal{E}} + \boldsymbol{\Sigma}_H)^{-1}. \tag{11}$$

Under the assumptions of mutually orthogonal pilot sequences and i.i.d. channel coefficients, we have $\boldsymbol{\Sigma}_\Delta = \delta\, \mathbb{I}_{M_R}$, and

$$p(\mathbf{H}|\widehat{\mathbf{H}}_{\mathrm{ML}}) = \mathcal{CN}\left(\delta \widehat{\mathbf{H}}_{\mathrm{ML}}, \delta \sigma_{\mathcal{E}}^2\, \mathbb{I}_{M_T} \otimes \mathbb{I}_{M_R}\right) \tag{12}$$

where

$$\delta = \frac{\sigma_h^2}{(\sigma_h^2 + \sigma_{\mathcal{E}}^2)}. \tag{13}$$

The availability of the estimation error distribution is an interesting feature of pilot-assisted channel estimation that we used to derive the posterior distribution (10). For simplicity, we do not specify hereafter the subscript ML for $\widehat{\mathbf{H}}$.

III. Detector Design in the Presence of Channel Estimation Errors

Consider the model (1) and denote by $f(\boldsymbol{y}_k, \boldsymbol{s}_k, \mathbf{H})$ the cost function that would let us decide in favor of a particular $\boldsymbol{s}_k$ at the receiver if the channel was perfectly known. Under a pilot-based channel estimation characterized by the posterior pdf of (10), we propose a detector based on the minimization of a new cost function, which is defined as

$$\begin{aligned}\widetilde{f}(\boldsymbol{y}_k, \boldsymbol{s}_k, \widehat{\mathbf{H}}) &= \int_{\mathbf{H}} f(\boldsymbol{y}_k, \boldsymbol{s}_k, \mathbf{H}) p(\mathbf{H}|\widehat{\mathbf{H}}) d\mathbf{H} \\ &= \mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left[f(\boldsymbol{y}_k, \boldsymbol{s}_k, \mathbf{H})\middle|\widehat{\mathbf{H}}\right]\end{aligned} \tag{14}$$

where we have averaged the cost function f over all realizations of the unknown channel $\mathbf{H}$ conditioned on its available estimate $\widehat{\mathbf{H}}$ by using the distribution (10). We note that the detector that minimizes (14) is an alternative to the suboptimal *mismatched* detector that is based on the minimization of the cost function $f(\boldsymbol{y}_k, \boldsymbol{s}_k, \widehat{\mathbf{H}})$. This latter value is obtained by using the estimated channel $\widehat{\mathbf{H}}$ in the same metric that would be applied if the channel was perfectly known, i.e., $f(\boldsymbol{y}_k, \boldsymbol{s}_k, \mathbf{H})$. The proposed approach in (14) differs from the mismatched detection on the conditional expectation $\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}[.]$, which provides a robust design by averaging the cost function $f(\boldsymbol{y}_k, \boldsymbol{s}_k, \mathbf{H})$ over all realizations of channel-estimation errors.

Consider the problem of detecting the symbol vector $\boldsymbol{s}_k$ from the observation model (1) in the ML sense, i.e., the maximization of the likelihood function, which we denote by $W(\boldsymbol{y}_k|\mathbf{H}, \boldsymbol{s}_k)$. It is well known that under i.i.d. Gaussian noise, the ML detection of $\boldsymbol{s}_k$ leads to minimizing the Euclidean distance metric $\mathcal{D}_{\mathrm{ML}}$. We have

$$\hat{\boldsymbol{s}_k}^{\mathrm{ML}}(\mathbf{H}) = \underset{\boldsymbol{s}_k \in \mathbb{C}^{M_T \times 1}}{\arg\min} \{\mathcal{D}_{\mathrm{ML}}(\boldsymbol{s}_k, \boldsymbol{y}_k, \mathbf{H})\} \tag{15}$$

with

$$\mathcal{D}_{\mathrm{ML}}(\boldsymbol{s}_k, \boldsymbol{y}_k, \mathbf{H}) \triangleq -\log W(\boldsymbol{y}_k|\boldsymbol{s}_k, \mathbf{H}) \propto \|\boldsymbol{y}_k - \mathbf{H}\boldsymbol{s}_k\|^2 \tag{16}$$

where $\propto$ means "is proportional to," and $\mathbb{C}$ is the set of complex numbers.

As an alternative to mismatched detection, we derive, in what follows, a modified likelihood criterion by averaging $W(\boldsymbol{y}_k|\mathbf{H}, \boldsymbol{s}_k)$ over all estimation errors. The modified likelihood function that we denote by $\widetilde{W}(\boldsymbol{y}_k|\widehat{\mathbf{H}}, \boldsymbol{s}_k)$ is obtained

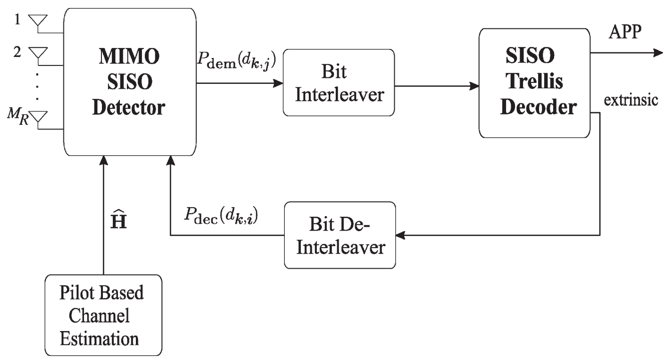

Fig. 1. Structure of the iterative MIMO receiver.

by using the pdf (10) and the formulation provided in (14). We have

$$\begin{aligned}\widetilde{W}(\boldsymbol{y}_k|\widehat{\mathbf{H}},\boldsymbol{s}_k) &= \int\limits_{\mathbf{H}\in\mathbb{C}^{M_R\times M_T}} W(\boldsymbol{y}_k|\mathbf{H},\boldsymbol{s}_k)p(\mathbf{H}|\widehat{\mathbf{H}})\,d\mathbf{H}\\ &=\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left[W(\boldsymbol{y}_k|\mathbf{H}_k,\boldsymbol{s}_k)\big|\widehat{\mathbf{H}}\right]. \qquad (17)\end{aligned}$$

After some algebraic manipulations, which are provided in Appendix I-B, we obtain

$$\widetilde{W}(\boldsymbol{y}_k|\widehat{\mathbf{H}},\boldsymbol{s}_k)=\mathcal{CN}(\mathbf{m}_{\mathcal{M}},\boldsymbol{\Sigma}_{\mathcal{M}}) \qquad (18)$$

where

$$\mathbf{m}_{\mathcal{M}}=\delta\widehat{\mathbf{H}}\boldsymbol{s}_k \quad \boldsymbol{\Sigma}_{\mathcal{M}}=\boldsymbol{\Sigma}_z+\delta\boldsymbol{\Sigma}_{\varepsilon}\|\boldsymbol{s}_k\|^2. \qquad (19)$$

Now, using (16), we can obtain the improved ML decision metric in the presence of imperfect channel estimation, which we denote by $\mathcal{D}_{\mathcal{M}}(\boldsymbol{s}_k,\boldsymbol{y}_k,\widehat{\mathbf{H}})$, i.e.,

$$\begin{aligned}\mathcal{D}_{\mathcal{M}}(\boldsymbol{s}_k,\boldsymbol{y}_k,\widehat{\mathbf{H}}) &\triangleq -\log\widetilde{W}(\boldsymbol{y}_k|\widehat{\mathbf{H}},\boldsymbol{s}_k)\\ &=M_R\log\pi\left(\sigma_z^2+\delta\sigma_{\mathcal{E}}^2\|\boldsymbol{s}_k\|^2\right)\\ &\quad+\frac{\|\boldsymbol{y}_k-\delta\widehat{\mathbf{H}}\boldsymbol{s}_k\|^2}{\sigma_z^2+\delta\sigma_{\mathcal{E}}^2\|\boldsymbol{s}_k\|^2}. \qquad (20)\end{aligned}$$

Note that the metric independently proposed in [5] and [8] for the case of ML detection can, in fact, be considered as a special case of the general framework (14), which leads to (20).

We note that, under near-perfect CSIR, which is obtained when the number of pilot symbols tends to infinity, i.e. $N_P\to\infty$, we have

$$\lim_{N_P\to\infty}\frac{\mathcal{D}_{\mathcal{M}}(\boldsymbol{s}_k,\boldsymbol{y}_k,\widehat{\mathbf{H}})}{\mathcal{D}_{\rm ML}(\boldsymbol{s}_k,\boldsymbol{y}_k,\widehat{\mathbf{H}})}=1. \qquad (21)$$

Consequently, the improved metric of (18) becomes equivalent to the mismatched metric for negligible estimation errors, i.e., for $\sigma_{\mathcal{E}}^2\to 0$.

IV. Iterative MIMO Detection Under Perfect CSIR

At the receiver, we perform iterative symbol detection and channel decoding. As shown in Fig. 1, the receiver principally consists of a MIMO detector (also called demapper) and a SISO channel decoder that exchange *extrinsic* soft information with each other. Here, we consider this soft information in the form of log-likelihood ratio (LLR).

The SISO decoder is based on the Max-Log-MAP algorithm, as described in [21] and [22]. We briefly present, in the following sections, the principle of MIMO detection based on two approaches of MAP and soft-PIC, assuming that perfect CSIR is available at the receiver. This step constitutes the bases for the next section, where we present improved detectors for the case of imperfect CSIR.

A. MAP Detection

Let d_k^m be the mth $(m=1,2,\ldots,BM_T)$ bit that corresponds to the symbol vector $\boldsymbol{s}_k$, transmitted at the kth time slot. We denote by $L(d_k^m)$ the LLR of the bit d_k^m at the output of the MIMO detector. Conditioned to perfect CSIR, $L(d_k^m)$ is given by

$$L\left(d_k^m\right)=\log\frac{P_{\rm dem}\left(d_k^m=1|\boldsymbol{y}_k,\mathbf{H}\right)}{P_{\rm dem}\left(d_k^m=0|\boldsymbol{y}_k,\mathbf{H}\right)} \qquad (22)$$

where $P_{\rm dem}(d_k^m|\boldsymbol{y}_k,\mathbf{H})$ denotes the probability of transmission of d_k^m. Let $\mathcal{S}$ be the set of all possibly transmitted symbol vectors $\boldsymbol{s}_k$. We partition $\mathcal{S}$ into $\mathcal{S}_0^m$ and $\mathcal{S}_1^m$, for which the mth bit of $\boldsymbol{s}_k$ is equal to "0" or "1," respectively. We have

$$L\left(d_k^m\right)=\log\frac{\sum\limits_{\boldsymbol{s}_k\in\mathcal{S}_1^m}e^{-\mathcal{D}_{\rm ML}(\boldsymbol{s}_k,\boldsymbol{y}_k,\mathbf{H})}\prod\limits_{\substack{n=1\\n\neq m}}^{BM_T}P_{\rm dec}^1\left(d_k^n\right)}{\sum\limits_{\boldsymbol{s}_k\in\mathcal{S}_0^m}e^{-\mathcal{D}_{\rm ML}(\boldsymbol{s}_k,\boldsymbol{y}_k,\mathbf{H})}\prod\limits_{\substack{n=1\\n\neq m}}^{BM_T}P_{\rm dec}^0\left(d_k^n\right)} \qquad (23)$$

where $P_{\rm dec}^1(d_k^n)$ and $P_{\rm dec}^0(d_k^n)$ are *a priori* information that comes from the SISO decoder. The summation in (23) is taken over the product of the likelihood $W(\boldsymbol{y}_k|\boldsymbol{s}_k,\mathbf{H})=\exp\{-\mathcal{D}_{\rm ML}(\boldsymbol{s}_k,\boldsymbol{y}_k,\mathbf{H})\}$, given a symbol vector $\boldsymbol{s}_k$, and the a priori probability on this symbol (the term $\prod P_{\rm dec}$), which is fed back from the SISO decoder at the previous iteration. In this latter term, the *a priori* probability of the bit d_k^m itself has been excluded to respect the exchange of *extrinsic* information between the channel decoder and the demapper. In addition, note that this term assumes independent coded bits d_k^n, which is true for random interleaving of large size. At the first iteration, where no *a priori* information is available on bits d_k^n, the probabilities $P_{\rm dec}^0(d_k^n)$ and $P_{\rm dec}^1(d_k^n)$ are set to 1/2.

B. Soft-PIC Detection

The computational complexity of the MAP detector becomes prohibitively large for large-sized signal constellations and/or for a large number of transmit antennas, because each of the sets $\mathcal{S}_1^m$ and $\mathcal{S}_0^m$ in (23) contains $2^{(BM_T-1)}$ vectors $\boldsymbol{s}_k$. For such cases, the suboptimal soft-PIC detector would make a good compromise between complexity and performance [17], [23]. In what follows, we recall the formulation of soft-PIC.

The general block diagram in Fig. 1 still applies to the turbo-PIC detector. Here, to detect a symbol transmitted from a given antenna, we first make use of the soft information available from the SISO channel decoder to reduce and, hopefully, cancel the interfering signals that arise from other transmit antennas. At the first iteration, where this information is not available, we perform a classical MMSE filtering.

Let us consider the transmitted vector $\boldsymbol{s}_k = [s_k^1, \ldots, s_k^{M_T}]^T$ at time k and assume that we are interested in the detection of its ith symbol s_k^i. We start by evaluating the parameters $\hat{s}_k^j$ and $\sigma^2_{s_k^j}$ for the interfering symbols $s_k^j, j \neq i$ from the SISO decoder as follows:

$$\hat{s}_k^j = \mathbb{E}\left[s_k^j\right] = \sum_{j=1}^{2^B} s_k^j P\left[s_k^j\right] \tag{24}$$

$$\sigma^2_{s_k^j} = \mathbb{E}\left[\left|s_k^j\right|^2\right] = \sum_{j=1}^{2^B} \left|s_k^j\right|^2 P\left[s_k^j\right] \tag{25}$$

where $P[s_k^j]$ is the probability of the transmission of s_k^j and is evaluated using the probabilities $P_{\mathrm{dec}}(d_k^{j,n})$ at the decoder output. We have

$$P\left[s_k^j\right] = K \prod_{n=1}^{B} P_{\mathrm{dec}}\left(d_k^{j,n}\right)$$

where K is a normalization factor. We further introduce the following definitions. $\underline{\mathbf{H}}_i$ is the $(M_R \times (M_T - 1)$ matrix constructed from $\mathbf{H}$ by discarding its ith column, i.e., $\mathbf{h}_i$. We also define the $(M_T - 1) \times 1$ vectors as

$$\underline{\mathbf{s}}_k^i \triangleq \left[s_k^1, s_k^2, \ldots, s_k^{i-1}, s_k^{i+1}, \ldots, s_k^{M_T}\right]^T$$
$$\underline{\hat{\mathbf{s}}}_k^i \triangleq \left[\hat{s}_k^1, \hat{s}_k^2, \ldots, \hat{s}_k^{i-1}, \hat{s}_k^{i+1}, \ldots, \hat{s}_k^{M_T}\right]^T$$

where $\hat{s}_k^j$ are estimated in (24).

Now, given the received signal vector $\boldsymbol{y}_k$, a soft interference cancellation is performed on $\boldsymbol{y}_k$ to detect the symbol s_k^i by subtracting to $\boldsymbol{y}_k$ the estimated signals of the other transmit antennas as

$$\underline{\boldsymbol{y}}_k^i = \boldsymbol{y}_k - \underline{\mathbf{H}}_i \underline{\hat{\boldsymbol{s}}}_k^i = \mathbf{h}_i s_k^i + \underline{\mathbf{H}}_i \underline{\boldsymbol{s}}_k^i - \underline{\mathbf{H}}_i \underline{\hat{\boldsymbol{s}}}_k^i + \boldsymbol{z}_k \quad \text{for} \quad i = 1, \ldots, M_T. \tag{26}$$

Except under perfect prior information on the symbols, which leads to $\hat{s}_k^j = s_k^j$, there remains a residual interference in $\underline{\boldsymbol{y}}_k^i$. To further reduce this interference, an instantaneous linear MMSE filter $\mathbf{w}_k^i$ is applied to $\underline{\boldsymbol{y}}_k^i$ to minimize the mean-square value of the error e_k^i defined as

$$e_k^i = s_k^i - r_k^i \tag{27}$$

where the filter output r_k^i is equal to

$$r_k^i = \mathbf{w}_k^i \underline{\boldsymbol{y}}_k^i. \tag{28}$$

Here, $\mathbf{w}_k^i$ is obtained as

$$\mathbf{w}_k^i = \underset{\mathbf{w}_k^i \in \mathbb{C}^{M_R}}{\arg\min}\, \mathbb{E}_{\boldsymbol{s}_k, \boldsymbol{z}_k}\left[\left|s_k^i - \mathbf{w}_k^i \underline{\boldsymbol{y}}_k^i\right|^2\right]. \tag{29}$$

By invoking the orthogonality principle [24], the coefficients of the MMSE filter $\mathbf{w}_k^i$ are given by

$$\mathbf{w}_k^i = \mathbf{h}_i^\dagger \left[\mathbf{h}_i \mathbf{h}_i^\dagger + \frac{\underline{\mathbf{H}}_i (\boldsymbol{\Lambda}_{k,i} - \widetilde{\boldsymbol{\Lambda}}_{k,i}) \underline{\mathbf{H}}_i^\dagger}{\sigma^2_{s_k^i}} + \frac{\sigma_z^2}{\sigma^2_{s_k^i}} \mathbb{I}_{M_R}\right]^{-1} \tag{30}$$

where

$$\begin{aligned}
\boldsymbol{\Lambda}_{k,i} &= \mathbb{E}\left[\underline{\boldsymbol{s}}_k^i \underline{\boldsymbol{s}}_k^{i\,\dagger}\right] \\
&\approx \mathrm{diag}\Bigg(\mathbb{E}\left[\left|s_k^1\right|^2\right], \ldots, \mathbb{E}\left[\left|s_k^{i-1}\right|^2\right] \\
&\qquad \mathbb{E}\left[\left|s_k^{i+1}\right|^2\right], \ldots, \mathbb{E}\left[\left|s_k^{M_T}\right|^2\right]\Bigg) \\
\widetilde{\boldsymbol{\Lambda}}_{k,i} &= \underline{\hat{\boldsymbol{s}}}_k^i \underline{\hat{\boldsymbol{s}}}_k^{i\,\dagger} \\
&\approx \mathrm{diag}\left(\left|\hat{s}_k^1\right|^2, \ldots, \left|\hat{s}_k^{i-1}\right|^2, \left|\hat{s}_k^{i+1}\right|^2, \ldots, \left|\hat{s}_k^{M_T}\right|^2\right).
\end{aligned}$$

Note that the off-diagonal entries in $\boldsymbol{\Lambda}_{k,i}$ and $\widetilde{\boldsymbol{\Lambda}}_{k,i}$ have been neglected to reduce the complexity without causing significant performance loss [18].

At the first decoding iteration, we have no prior information available on the transmitted data, i.e., $\boldsymbol{\Lambda}_{k,i} = \sigma^2_{s_k^i} \mathbb{I}_{M_T - 1}$, and $\widetilde{\boldsymbol{\Lambda}}_{k,i} = \mathbf{0}_{M_T - 1}$. Consequently, (30) reduces to

$$\mathbf{w}_k^i = \mathbf{h}_i^\dagger \left[\mathbf{H}\mathbf{H}^\dagger + \frac{\sigma_z^2}{\sigma^2_{s_k^i}} \mathbb{I}_{M_R}\right]^{-1} \tag{31}$$

which is no more than the linear MMSE detector for s_k^i.

Before being passed to the SISO decoder, the detected symbols r_k at the output of the MMSE filter in (28) should be converted to LLR. This method is done by assuming a Gaussian distribution for the residual interference after soft-PIC detection. Details on the LLR conversion can be found in [16]. For brevity, the discussion will be presented only for the case of the improved detector in Section V-B.

1) Simplified Soft-PIC Detector: It is clear from (30) and (38) that the MMSE filter coefficients change as a function of time and must be recomputed for each time symbol. This condition motivates us to propose here a simplified version of the soft-PIC detector to further reduce the receiver complexity.

Consider the filter expression (30). We may assume that, after the second iteration, the soft estimates of transmitted symbols are reliable, and hence, the interference cancellation is almost perfectly performed. In other words, we assume that $\underline{\hat{\boldsymbol{s}}}_k^i = \underline{\boldsymbol{s}}_k^i$, and as a result, $\widetilde{\boldsymbol{\Lambda}}_{k,i} = \boldsymbol{\Lambda}_{k,i}$. With this approximation, the expression of $\mathbf{w}_k^i$ in (30) simplifies to the following expression [17], [25]:

$$\mathbf{w}_k^i = \mathbf{h}_i^\dagger \left[\mathbf{h}_i \mathbf{h}_i^\dagger + \frac{\sigma_z^2}{\sigma^2_{s_k^i}} \mathbb{I}_{M_R}\right]^{-1} = \frac{1}{\mathbf{h}_i^\dagger \mathbf{h}_i + \frac{\sigma_z^2}{\sigma^2_{s_k^i}}} \mathbf{h}_i^\dagger \tag{32}$$

where the filter coefficients no longer depend on k because $\sigma^2_{s^i_k}$ is constant for all k. Note that the last expression is nothing more than a simple matched filter. In addition, note that, with this approximation, we do not need to update the MMSE filter for each sample time, because we effectively need to calculate $\mathbf{w}^i_k$ once for a given realization of $\mathbf{H}$, i.e., for the underlying fading block. Furthermore, the matrix inversion in (30) is replaced by a scalar inversion. Obviously, this simplification causes degradation to the receiver performance. However, this performance loss would be justified with regard to the complexity reduction, and hopefully, due to iterative detection, it would be acceptable [26], [27].

Note that, unless otherwise mentioned, by turbo-PIC under perfect CSIR, we mean the exact formulation of the soft-PIC detector, i.e., (30).

C. Generalization to the Case of Linear Dispersion ST Codes

We have considered the simple spatial multiplexing (i.e., V-BLAST) scheme at the transmitter. We show here that the detectors' formulations, which are presented for the case of spatial multiplexing, can also be applied to the more general case of linear dispersion (LD) codes [28] with a slight modification. The reader is referred to [28] for more details on the LD codes' formulation.

Let $\boldsymbol{s}$ of dimension $Q \times 1$ be the vector of data symbols prior to ST coding. By LD coding, these symbols are mapped into a matrix $\boldsymbol{X}$ of dimension $M_T \times T_u$ with T_u the number of channel uses. For an encoded matrix $\boldsymbol{X}$, we receive the matrix $\boldsymbol{Y}$ of dimension $M_R \times T_u$. To obtain a general formulation for the receiver, we separate the real and the imaginary parts of the entries of $\boldsymbol{s}$, $\boldsymbol{Y}$, and $\boldsymbol{X}$ and stack them rowwise in vectors $\breve{\mathrm{s}}$, $\breve{\mathbf{X}}$, and $\breve{\mathbf{Y}}$ of dimension $2Q \times 1$, $2M_T T_u \times 1$, and $2M_R T_u \times 1$, respectively. Vectors $\breve{\mathrm{s}}$ and $\breve{\mathbf{Y}}$ are related through an equivalent channel matrix $\breve{\mathbf{H}}_{eq}$ of dimension $(2M_R T_u \times 2Q)$, which depends on $\mathbf{H}$ and the employed LD code (see [28]). We have

$$\breve{\mathbf{Y}} = \breve{\mathbf{H}}_{eq}\, \breve{\mathrm{s}} + \breve{\mathbf{Z}} \tag{33}$$

where $\breve{\mathbf{Z}}$ is the vector of real AWGN of variance $\sigma^2_z/2$. We see that, in the expressions of the detectors, we only have to consider $\breve{\mathbf{H}}_{eq}$, $\breve{\mathrm{s}}$, and $\breve{\mathbf{Y}}$, instead of $\mathbf{H}$, $\boldsymbol{s}$, and $\boldsymbol{y}$, respectively.

V. Improved Detection Under Imperfect CSIR

We propose here modifications to the MIMO detector to mitigate the impact of imperfect channel estimation on the receiver performance. We first consider the MAP and, then, the soft-PIC detector in the following sections. For soft-PIC, we consider both cases of exact and simplified detectors.

A. Improved MAP Detection

We notice that the metric $\mathcal{D}_{\mathrm{ML}}(\boldsymbol{s}_k, \boldsymbol{y}_k, \mathbf{H})$ involved in (23) requires the perfect channel matrix $\mathbf{H}$, of which the receiver has *solely* an estimate $\widehat{\mathbf{H}}$. We propose to use the decoding metric $\mathcal{D}_{\mathcal{M}}(\boldsymbol{s}_k, \boldsymbol{y}_k, \widehat{\mathbf{H}})$ of (20) for the evaluation of the LLRs in (23). This step leads us to derive a new demapping rule adapted from the imperfect CSIR. This way, we calculate $L(d^m_k)$ from (22), where, for instance, the nominator is calculated as follows:

$$\begin{aligned} P&\left(d^m_k = 1 | \boldsymbol{y}_k, \widehat{\mathbf{H}}\right) \\ &= \sum_{\boldsymbol{s}_k \in \mathcal{S}^m_1} \exp\left\{-\mathcal{D}_{\mathcal{M}}(\boldsymbol{s}_k, \boldsymbol{y}_k, \widehat{\mathbf{H}})\right\} \prod_{\substack{n=1\\ n\neq m}}^{BM_T} P^1_{\mathrm{dec}}\left(d^n_k\right) \\ &= \frac{1}{\pi e^{M_R}} \sum_{\boldsymbol{s}_k \in \mathcal{S}^m_1} \frac{1}{\sigma^2_z + \delta\sigma^2_\varepsilon \|\boldsymbol{s}_k\|^2} \\ &\quad \times \exp\left\{\frac{\|\boldsymbol{y}_k - \delta\widehat{\mathbf{H}}\boldsymbol{s}_k\|^2}{\sigma^2_z + \delta\sigma^2_\varepsilon\|\boldsymbol{s}_k\|^2}\right\} \prod_{\substack{n=1\\ n\neq m}}^{BM_T} P^1_{\mathrm{dec}}\left(d^n_k\right). \end{aligned} \tag{34}$$

B. Improved Soft-PIC Detection

As shown in (26) and (30), we need the channel $\mathbf{H}$ for both interference canceling and MMSE filtering. The receiver has only an imperfect channel estimate $\widehat{\mathbf{H}}$; thus, the suboptimal *mismatched* solution consists of replacing $\underline{\mathbf{H}}_i$ and $\mathbf{h}_i$ in (26) and (30) by their estimates $\widehat{\underline{\mathbf{H}}}_i$ and $\widehat{\mathbf{h}}_i$, respectively. As the first step toward a realistic design, we make use of the available channel estimate $\widehat{\mathbf{H}}$ for interference cancellation. That is, (26) is rewritten as

$$\underline{\boldsymbol{y}}^i_k = \boldsymbol{y}_k - \widehat{\underline{\mathbf{H}}}_i \hat{\underline{\boldsymbol{s}}}^i_k = \mathbf{h}_i s^i_k + \underline{\mathbf{H}}_i \underline{\boldsymbol{s}}^i_k - \widehat{\underline{\mathbf{H}}}_i \hat{\underline{\boldsymbol{s}}}^i_k + \boldsymbol{z}_k \quad \text{for} \quad i = 1, \ldots, M_T \tag{35}$$

where $\widehat{\underline{\mathbf{H}}}_i$ is the $(M_R \times (M_T - 1))$ matrix constructed from $\widehat{\mathbf{H}}$ by discarding its ith column, i.e., $\widehat{\mathbf{h}}_i$. We now propose a novel *improved* PIC detector under imperfect CSIR. We note that (35) naturally depends on the unknown channel matrix $\mathbf{H}$, of which the receiver has only an imperfect estimate available. Instead of replacing the unknown channel by its estimate (i.e., the mismatched approach), we use the posterior distribution (12) and make two modifications to the detector described in Section IV-B as follows.

The first modification that we propose concerns the design of the filter $\mathbf{w}^i_k$ in (29). The cost function $f(\boldsymbol{y}_k, s^i_k, \mathbf{H}) = \mathbb{E}_{\boldsymbol{s}_k, \boldsymbol{z}_k}[|s^i_k - \mathbf{w}^i_k \underline{\boldsymbol{y}}^i_k|^2]$ is a function of the perfect channel $\mathbf{H}$ (via $\underline{\boldsymbol{y}}^i_k$); thus, we propose a modified filter $\widetilde{\mathbf{w}}^i_k$, which was chosen to minimize the average of the mean-square error over all realizations of channel estimation errors. Using (12), we propose the following filter design:

$$\begin{aligned} \widetilde{\mathbf{w}}^i_k &= \underset{\widetilde{\mathbf{w}}^i_k \in \mathbb{C}^{M_R}}{\arg\min}\ \mathbb{E}_{\mathbf{H}, \boldsymbol{s}_k, \boldsymbol{z}_k}\left[\left|s^i_k - \widetilde{\mathbf{w}}^i_k \underline{\boldsymbol{y}}^i_k\right|^2 \middle| \widehat{\mathbf{H}}\right] \\ &= \underset{\widetilde{\mathbf{w}} \in \mathbb{C}^{M_R}}{\arg\min}\ \mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left[\mathbb{E}_{\boldsymbol{s}_k, \boldsymbol{z}_k}\left[\left|s^i_k - \widetilde{\mathbf{w}}^i_k \underline{\boldsymbol{y}}^i_k\right|^2\right]\right] \end{aligned} \tag{36}$$

where, in the latter expression, we have assumed the independence between $\mathbf{H}$, $\boldsymbol{s}_k$, and $\boldsymbol{z}_k$.

From (36) and after invoking the orthogonality principle [24], we obtain

$$\widetilde{\mathbf{w}}^i_k = \left(\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left[\mathbb{E}_{\boldsymbol{s}_k, \boldsymbol{z}_k}\left[s^i_k \underline{\boldsymbol{y}}^{i\,\dagger}_k\right]\right]\right)\left(\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left[\mathbb{E}_{\boldsymbol{s}_k, \boldsymbol{z}_k}\left[\underline{\boldsymbol{y}}^i_k \underline{\boldsymbol{y}}^{i\,\dagger}_k\right]\right]\right)^{-1}. \tag{37}$$

After some algebraic manipulations in Appendix II-A, we get the modified filter $\widetilde{\mathbf{w}}_k^i$, directly as a function of $\widehat{\underline{\mathbf{H}}}_i$ and $\widehat{\mathbf{h}}_i$, as follows:

$$\widetilde{\mathbf{w}}_k^i = \overline{\mathbf{R}}_{s_k^i \underline{\boldsymbol{y}}_k^i} \overline{\mathbf{R}}_{\underline{\boldsymbol{y}}_k^i}^{-1} \tag{38}$$

where

$$\overline{\mathbf{R}}_{s_k^i \underline{\boldsymbol{y}}_k^i} = \delta\sigma_{s_k^i}^2 \widehat{\mathbf{h}}_i^\dagger + (\delta - 1)\mathbf{m}_{k,i} \widehat{\underline{\mathbf{H}}}_i^\dagger \tag{39}$$

with $\mathbf{m}_{k,i} = \hat{s}_k^i \hat{\underline{\boldsymbol{s}}}_k^{i\,\dagger}$. δ is given by (13), and

$$\begin{aligned}\overline{\mathbf{R}}_{\underline{\boldsymbol{y}}_k^i} =& \delta^2\sigma_{s_k^i}^2 \widehat{\mathbf{h}}_i\widehat{\mathbf{h}}_i^\dagger + \delta^2\widehat{\underline{\mathbf{H}}}_i\mathbf{\Lambda}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger + (\delta^2 - \delta)\widehat{\mathbf{h}}_i\mathbf{m}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger \\ &+ (\delta^2 - \delta)\widehat{\underline{\mathbf{H}}}_i\mathbf{m}_{k,i}^\dagger\widehat{\mathbf{h}}_i^\dagger + (1 - 2\delta)\widehat{\underline{\mathbf{H}}}_i\widetilde{\mathbf{\Lambda}}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger \\ &+ \left(\sigma_z^2 + (1-\delta)\sigma_{s_k^i}^2 + (1-\delta)\mathrm{tr}(\mathbf{\Lambda}_{k,i})\right)\mathbb{I}_{M_R}.\end{aligned} \tag{40}$$

To get more insight on the proposed detector, let us consider the ideal case where perfect channel knowledge is available at the receiver, i.e., $\widehat{\mathbf{H}} = \mathbf{H}$, and $\sigma_{\mathcal{E}}^2 = 0$. We note that, in this case, $\delta = 1$, and the posterior pdf (12) reduces to a Dirac delta function; consequently, the two filters $\widetilde{\mathbf{w}}_k^i$ and $\mathbf{w}_k^i$ coincide. Similarly, under near-perfect CSIR, which is obtained either when $\sigma_{\mathcal{E}}^2 \to 0$ or when $N_P \to \infty$, we have $\delta \to 1$, and the filter $\widetilde{\mathbf{w}}_k^i$ gives a similar expression as $\mathbf{w}_k^i$ in (30). However, in the presence of estimation errors, the proposed improved and mismatched detectors become different due to the inherent averaging in (36), which provides a robust design that adapts itself to the channel estimate available at the receiver.

Our second modification concerns the application of the derived filter $\widetilde{\mathbf{w}}_k^i$ to the received signal $\underline{\boldsymbol{y}}_k^i$. By applying the modified filter $\widetilde{\mathbf{w}}_k^i$ to $\underline{\boldsymbol{y}}_k^i$ in (35), we have

$$r_k^i = \widetilde{\mathbf{w}}_k^i\underline{\boldsymbol{y}}_k^i = \widetilde{\mathbf{w}}_k^i\mathbf{h}_i s_k^i + \widetilde{\mathbf{w}}_k^i\underline{\mathbf{H}}_i\underline{\boldsymbol{s}}_k^i - \widetilde{\mathbf{w}}_k^i\widehat{\underline{\mathbf{H}}}_i\hat{\underline{\boldsymbol{s}}}_k^i + \widetilde{\mathbf{w}}_k^i\boldsymbol{z}_k. \tag{41}$$

The latter equation is a function of the perfect channel of which the receiver has only an imperfect estimate. Again, instead of replacing the perfect channel by its estimate (which is the mismatched approach), we propose to average the filter output r_k^i by using the *a posteriori* distribution (12) as follows:

$$\widetilde{r}_k^i = \mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left[r_k^i\right] = \underbrace{\delta\widetilde{\mathbf{w}}_k^i\widehat{\mathbf{h}}_i}_{\mu_{k,i}} s_k^i + \underbrace{\delta\widetilde{\mathbf{w}}_k^i\widehat{\underline{\mathbf{H}}}_i\underline{\boldsymbol{s}}_k^i - \widetilde{\mathbf{w}}_k^i\widehat{\underline{\mathbf{H}}}_i\hat{\underline{\boldsymbol{s}}}_k^i + \widetilde{\mathbf{w}}_k^i\boldsymbol{z}_k}_{\eta_{k,i}} \tag{42}$$

where we have used $\mathbb{E}_{\mathbf{h}_i|\widehat{\mathbf{h}}_i}[\mathbf{h}_i] = \delta\widehat{\mathbf{h}}_i$, $\mathbb{E}_{\underline{\mathbf{H}}_i|\widehat{\underline{\mathbf{H}}}_i}[\underline{\mathbf{H}}_i] = \delta\widehat{\underline{\mathbf{H}}}_i$, and $\eta_{k,i}$ is the interference-plus-noise ratio that affects the output of the instantaneous MMSE filter $\widetilde{r}_k^i$. Based on (42), it is clear that the output of the improved MMSE filter can be viewed as an equivalent AWGN channel with s_k^i at its input, i.e.,

$$\widetilde{r}_k^i = \mu_{k,i}s_k^i + \eta_{k,i}. \tag{43}$$

It is shown in [16] and [19] that, under perfect CSIR, a similar expression of $\eta_{k,i}$ is well approximated by a zero-mean Gaussian random variable with variance $\sigma_{\eta_{k,i}}^2$. The parameters $\mu_{k,i}$ and $\sigma_{\eta_{k,i}}^2$ are calculated at each time slot by using the symbol statistics. The exact derivation of the variance $\sigma_{\eta_{k,i}}^2$ is provided in Appendix II-B, for the more general case of the improved detector.

Based on (42), we can calculate the LLRs on the corresponding bits of the detected symbols at the output of the MMSE filter, which will be used by the SISO channel decoder. Let $d_k^{i,m}$ denote the mth $(m = 1, \ldots, B)$ bit that corresponds to s_k^i. The LLR on $d_k^{i,m}$ is given by

$$\begin{aligned}L\left(d_k^{i,m}\right) &= \log\frac{P_{\mathrm{dem}}\left(d_k^{i,m} = 1|\widetilde{r}_k^i, \mu_{k,i}\right)}{P_{\mathrm{dem}}\left(d_k^{i,m} = 0|\widetilde{r}_k^i, \mu_{k,i}\right)} \\ &= \log\frac{\sum\limits_{s_k^i \in \underline{\mathcal{S}}_1^m}\exp\left\{-\frac{\left|\widetilde{r}_k^i - \mu_{k,i}s_k^i\right|^2}{\sigma_{\eta_{k,i}}^2}\right\}\prod\limits_{\substack{n=1\\n\neq m}}^{B}P_{\mathrm{dec}}^1\left(d_k^{i,n}\right)}{\sum\limits_{s_k^i \in \underline{\mathcal{S}}_0^m}\exp\left\{-\frac{\left|\widetilde{r}_k^i - \mu_{k,i}s_k^i\right|^2}{\sigma_{\eta_{k,i}}^2}\right\}\prod\limits_{\substack{n=1\\n\neq m}}^{B}P_{\mathrm{dec}}^0\left(d_k^{i,n}\right)}.\end{aligned} \tag{44}$$

Note that, in contrast to the case of MAP detection where, in (23), $\mathcal{S}_1^m$ and $\mathcal{S}_0^m$ are of size $2^{(M_T B - 1)}$, here, the cardinality of the sets $\underline{\mathcal{S}}_1^m$ and $\underline{\mathcal{S}}_0^m$ is only equal to 2^{B-1}.

Finally, note that adopting the aforementioned improved reception scheme does not considerably increase the complexity compared to the mismatched approach. More precisely, comparing the mismatched and the improved filter expressions [i.e., comparing (30) after replacing the perfect channel by its estimate and (38)] shows that the improved filter requires only a few number of extra matrix additions and multiplications, which does not have an important impact on the receiver complexity; the required additional operations are one vector by matrix multiplication and one vector addition for calculating $\overline{\mathbf{R}}_{s_k^i\underline{\boldsymbol{y}}_k^i}$, as well as two matrix multiplications, two scalar by matrix multiplications, and two matrix additions for calculating $\overline{\mathbf{R}}_{\underline{\boldsymbol{y}}_k^i}$.

1) Modification for the Simplified Soft-PIC Detector: We have presented a simplified formulation for the turbo-PIC detector; thus, in this section, we present the improved detector formulation for the case of imperfect CSIR. For the first iteration, the formulation is like the case of exact soft-PIC and remains unchanged. For the next iterations, it can easily be verified that replacing $\widetilde{\mathbf{\Lambda}}_{k,i}$ by $\mathbf{\Lambda}_{k,i}$ in (38) does not, unfortunately, lead to a compact expression that involves a simple scalar inversion. However, with a slight modification in the interference cancellation part, we can also derive a simple improved detector similar to (32). This result is achieved by cancelling the residual interference by using (26) instead of (35) as

$$\underline{\boldsymbol{y}}_k^i = \mathbf{h}_i s_k^i + \underline{\mathbf{H}}_i\underline{\boldsymbol{s}}_k^i - \underline{\mathbf{H}}_i\hat{\underline{\boldsymbol{s}}}_k^i + \boldsymbol{z}_k. \tag{45}$$

By following a similar procedure as in Section V-B, it can then be shown that the improved MMSE filter of (38) reduces to

$$\begin{aligned}\widetilde{\mathbf{w}}_k^i = \delta\widehat{\mathbf{h}}_i^\dagger\Bigg[&\delta^2\widehat{\mathbf{h}}_i\widehat{\mathbf{h}}_i^\dagger + (1-\delta) \\ &\times\left(1 + \frac{\mathrm{tr}(\mathbf{\Lambda}_{k,i} - \widetilde{\mathbf{\Lambda}}_{k,i})}{\sigma_{s_k^i}^2} + \frac{\sigma_z^2}{(1-\delta)\sigma_{s_k^i}^2}\right) \\ &\times\mathbb{I}_{M_R} + \frac{\delta^2\widehat{\underline{\mathbf{H}}}_i(\mathbf{\Lambda}_{k,i} - \widetilde{\Lambda}_{k,i})\widehat{\underline{\mathbf{H}}}_i^\dagger}{\sigma_{s_k^i}^2}\Bigg]^{-1}.\end{aligned} \tag{46}$$

Now, by setting $\widetilde{\mathbf{\Lambda}}_{k,i}$ to be equal to $\mathbf{\Lambda}_{k,i}$ in (46), we obtain the following simplified expression for the MMSE filter of the improved detector:

$$\begin{aligned}\widetilde{\mathbf{w}}_k^i &= \delta\widehat{\mathbf{h}}_i^\dagger \left[\delta^2\widehat{\mathbf{h}}_i\widehat{\mathbf{h}}_i^\dagger + (1-\delta)\mathbb{I}_{M_R} + \frac{\sigma_z^2}{\sigma_{s_k^i}^2}\mathbb{I}_{M_R}\right]^{-1} \\ &= \frac{\delta}{\delta^2\widehat{\mathbf{h}}_i^\dagger\widehat{\mathbf{h}}_i + (1-\delta) + \frac{\sigma_z^2}{\sigma_{s_k^i}^2}}\widehat{\mathbf{h}}_i^\dagger. \end{aligned} \quad (47)$$

The modification of the interference cancellation and the calculation of the LLRs are the same, as provided by (42) and (44), respectively.

VI. Numerical Results

In this section, we provide some numerical results to show the performance improvement by using the proposed modified detectors in the presence of channel-estimation errors. The performance is evaluated in terms of receiver bit error rates (BERs).

For channel coding in our BICM scheme, we consider the rate 1/2 NRNSC code of constraint length $K = 7$, defined in octal form by $(133, 171)_8$. Uncorrelated Rayleigh fading channel is considered, and each frame of symbols corresponds to N_c fading blocks. Channel coefficients are kept constant during each fading block and are changed to new independent realizations from one block to the next. For each fading block, we devote N_P channel uses to the transmission of pilot sequences. Throughout the simulations, each frame is composed of $L = 128$ channel uses for data symbols plus a total number of N_cN_P channel uses for pilot transmission. Data symbols belong to QPSK or 16-QAM constellations with Gray labeling. We use mutually orthogonal QPSK pilot sequences for channel estimation, and the average pilot-symbol power is set to be equal to the average data-symbol power.

The interleaver is pseudorandom, which operates over the entire frame of size $N_I = LBM_T$ bits (excluding pilots, obviously). Moreover, the number of receiver iterations is set to 5. The signal-to-noise ratio (SNR) is considered in the form of E_b/N_0 and includes the receiver antenna array gain M_R. We assume that the noise variance is known at the receiver.

A. Case of Turbo-MAP Detector

Let us first consider the turbo-MAP detector. Fig. 2 shows BER curves of the mismatched and improved receivers for the case of QPSK modulation, $M_T = 2$ and $M_R = 2$, which we denote by a (2×2) MIMO system. The number of channel uses for pilot transmission is $N_P \in \{2, 4, 8\}$. As a reference, we have also presented the BER curve in the case of perfect CSIR. We notice that the required SNR to attain the BER of 10^{-5} with $N_P = 2$ pilots is reduced by about 0.4 dB for the improved detector compared with the mismatched detector. By increasing N_P, the channel-estimation error becomes less important, and the difference of the performances of the two detectors decreases; the achieved gain in SNR at BER $= 10^{-5}$ is about 0.2 and 0.1 dB for $N_P = 4$ and $N_P = 8$, respectively.

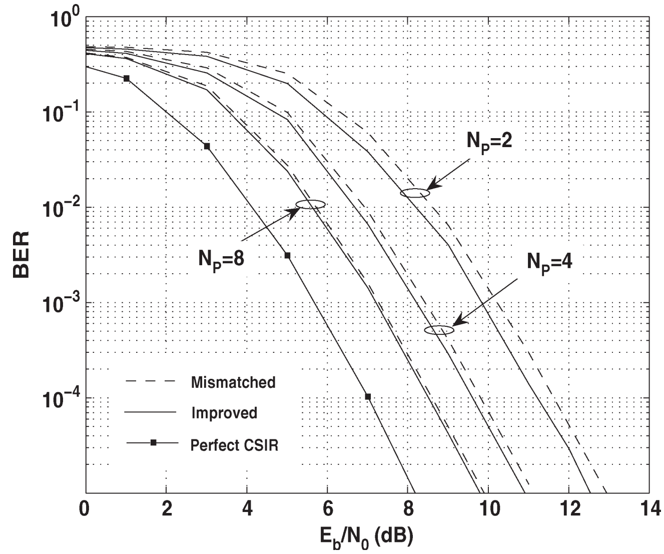

Fig. 2. BER performance of the improved and the mismatched turbo-MAP with the 2×2 MIMO with the V-BLAST ST scheme, i.i.d. Rayleigh fading with $N_c = 4$, QPSK modulation, and training sequence length of $N_P \in \{2, 4, 8\}$.

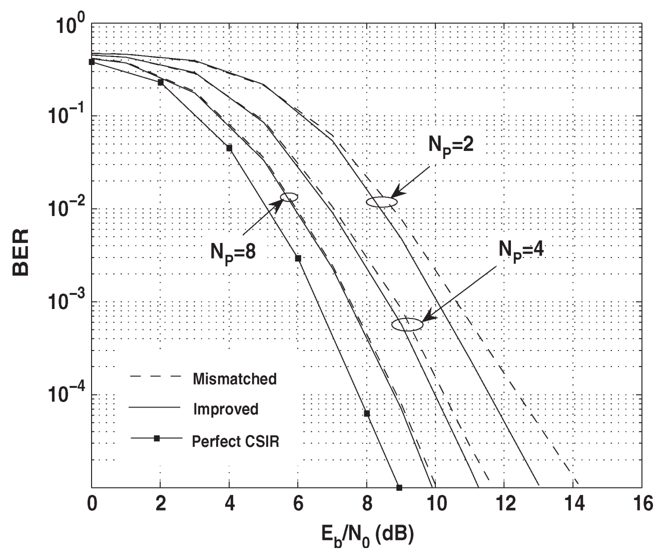

Fig. 3. BER performance of the improved and the mismatched turbo-PIC with the $2\times$ 2) MIMO with the V-BLAST ST scheme, i.i.d. Rayleigh fading with $N_c = 4$, QPSK modulation, and training sequence length of $N_P \in \{2, 4, 8\}$.

In fact, the performance loss of the mismatched receiver with respect to the improved receiver becomes insignificant for $N_P \geq 8$.

B. Case of the Turbo-PIC Detector

For the case of the turbo-PIC detector, Fig. 3 shows BER curves of the mismatched and improved receivers under the same conditions as in Fig. 2. We see that the gain in the SNR of the improved detector to attain the BER of 10^{-5} is now about 1.4, 0.5, and 0.2 dB, respectively, for $N_P = 2, 4,$ and 8. In fact, the gain is more important than for turbo-MAP.

It is interesting to see the evolution of BER versus the number of iterations. We have shown in Fig. 4, for $N_P = 2$, the corresponding BER curves for the cases of mismatched and

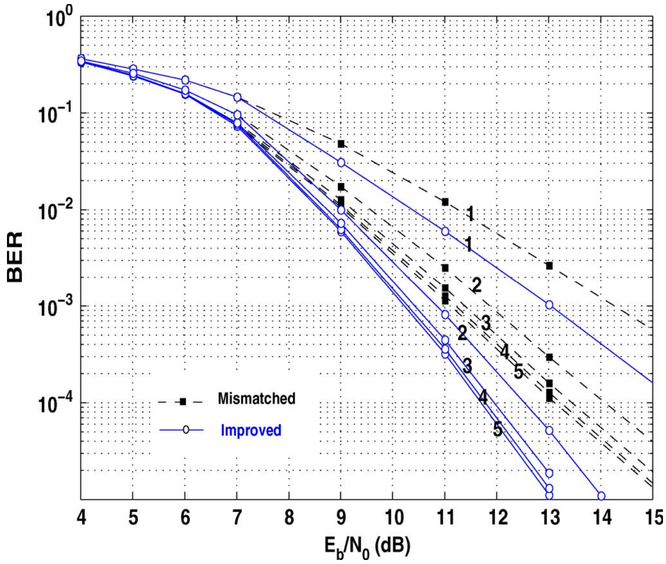

Fig. 4. BER performance of the improved and the mismatched turbo-PIC through iterations. $N_P = 2$. Other parameters are the same as in Fig. 3. Numbers on the curves indicate iteration numbers.

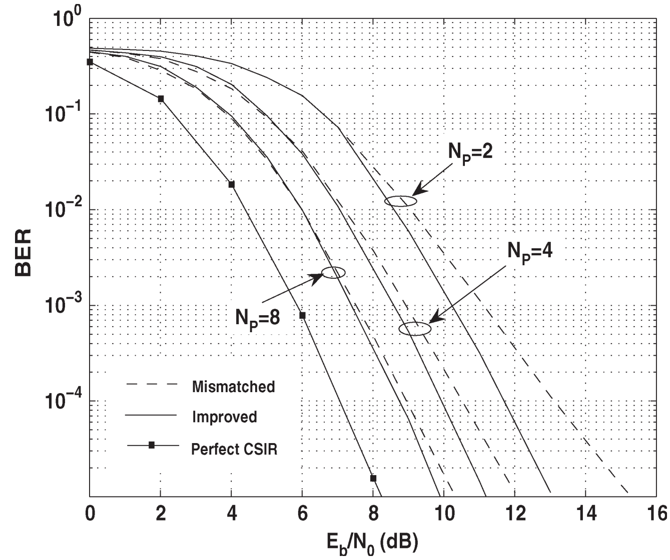

Fig. 5. BER performance of the improved and the mismatched turbo-PIC with the 2×2 MIMO with the GLD ST scheme, i.i.d. Rayleigh fading with $N_c = 4$, QPSK modulation, and training sequence length of $N_P \in \{2, 4, 8\}$.

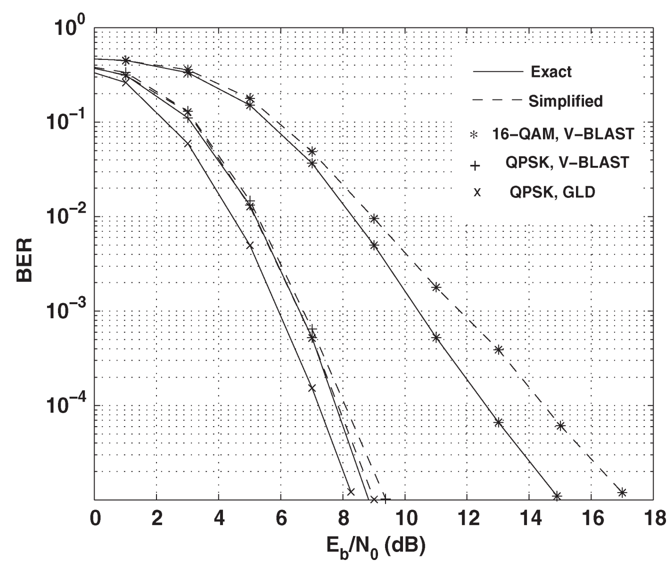

Fig. 6. BER performance of the exact and the simplified turbo-PIC implementations with 2×2 MIMO, i.i.d. Rayleigh fading with $N_c = 4$, and perfect CSIR.

improved detectors for up to five iterations of the receiver. We notice that the two detectors have almost the same convergence trend, and the major improvement is obtained after the second iteration for both detectors. As a result, if, due to complexity reduction, we only process two receiver iterations, then we still have a considerable performance gain by using the improved detector.

We have also considered the case of a 2×2 system with a larger sized signal constellation, i.e., the 16-QAM (results are not shown due to space limits). The obtained gain by using the improved detector is less considerable compared with the case of QPSK modulation. In fact, even with perfect CSIR, for larger constellation sizes, turbo-PIC becomes more suboptimal, except in the presence of high (time or space) diversity. Consequently, the improved turbo-PIC does not offer a considerable gain compared to the mismatched detector.

C. Case of Turbo-PIC Detector With ST Coding

We have considered the simple spatial multiplexing at the transmitter. It is interesting to study the performance of the improved detector for *more powerful* ST codes. We consider here as the ST scheme the optimized *golden code* (GLD), as presented in [30] for the case of two transmit antennas and $M_R \geq M_T$, which offers full-rate and full-diversity with nonvanishing determinant. With this ST scheme, each vector of four symbols $\boldsymbol{s} = [\, s^1 \quad s^2 \quad s^3 \quad s^4 \,]^T$ is mapped into a (2 × 2) matrix $\mathbf{X}$ as described as follows:

$$\mathbf{X} = \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha[s^1 + \theta s^2] & \alpha[s^3 + \theta s^4] \\ \gamma\overline{\alpha}[s^3 + \overline{\theta} s^4] & \overline{\alpha}[s^1 + \overline{\theta} s^2] \end{bmatrix} \tag{48}$$

where

$$\theta = \frac{1 + \sqrt{5}}{2}, \qquad \alpha = 1 + j(1 - \theta), \quad \overline{\theta} = 1 - \theta$$
$$\overline{\alpha} = 1 + j(1 - \overline{\theta}), \quad \gamma = j, \qquad j = \sqrt{-1}.$$

Fig. 5 contrasts the BER curves of the receiver for the cases of improved and mismatched detectors. It is shown that the performance gain by using the improved detector can be quite considerable in this case. The SNR gains at a BER of 10^{-5} are of 2.3, 0.9, and 0.4 dB for $N_P = 2, 4,$ and 8, respectively.

D. Case of the Simplified Turbo-PIC Detector

Let us now consider the case of the simplified soft-PIC and study the performance gain by using the improved detector. Remember that, here, both the mismatched and the improved simplified detectors need to calculate the MMSE filter once per channel realization and not for each channel use, as in the case for the exact implementation of the soft-PIC.

First, in Fig. 6, we have compared the performances of the exact and the simplified turbo-PIC for the case of perfect

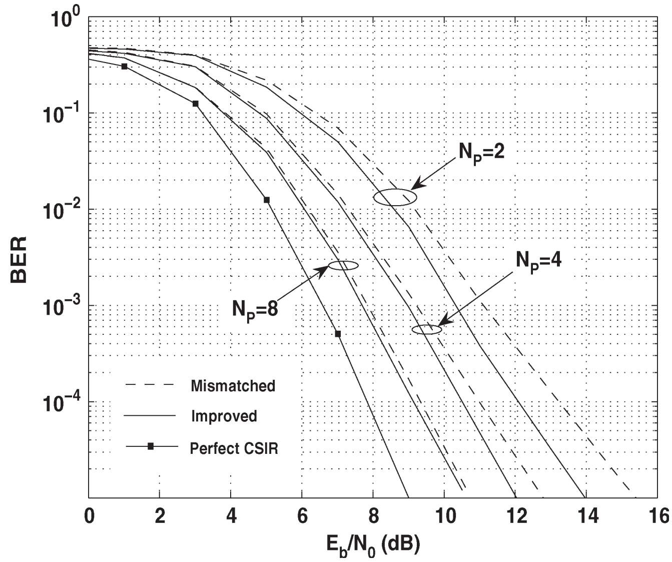

Fig. 7. BER performance of the improved and the mismatched simplified turbo-PIC with 2×2 MIMO with the V-BLAST ST scheme, i.i.d. Rayleigh fading with $N_c = 4$, QPSK modulation, and training sequence length of $N_P \in \{2, 4, 8\}$.

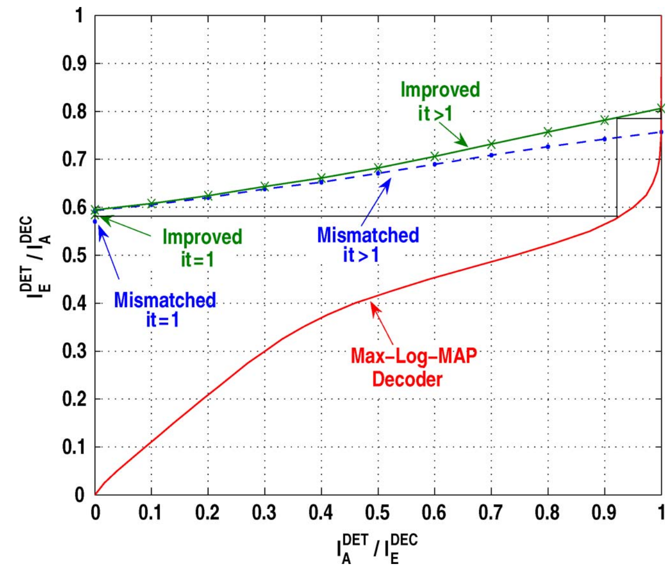

Fig. 8. EXIT charts analysis of the improved and the mismatched turbo-PIC with 2×2 MIMO with the GLD ST scheme, i.i.d. Rayleigh fading with $N_c = 4$, QPSK modulation, and training sequence length of $N_P = 2$. $E_b/N_0 = 5$ dB. "it" stands for iteration number.

CSIR, where a 2×2 system is considered with the simple V-BLAST (QPSK modulation and 16-QAM) and with the GLD scheme (QPSK modulation). We notice that the performance degradation by using the simplified soft-PIC detector at the BER of 10^{-5} is about 0.6 and 0.7 dB, respectively, for the V-BLAST and GLD schemes with a QPSK modulation and more than 3 dB for the V-BLAST scheme with 16-QAM. In effect, for large constellation sizes, the performance degradation by the simplification made in soft-PIC becomes considerable.

Now, consider Fig. 7, which compares the performances of the mismatched and the improved simplified detectors under the same conditions as in Fig. 3, i.e., for a V-BLAST scheme with QPSK modulation. We notice that the gain in SNR by using the improved detector is as important as in the case of the exact formulation of the soft-PIC and is about 1.4, 0.7, and 0.2 dB for $N_P = 2, 4$, and 8, respectively.

VII. Impact on the Convergence of Turbo-PIC

It is interesting to study the influence of the proposed modification on the convergence of the turbo-PIC detector. For this purpose, we consider the extrinsic-information transfer (EXIT) charts, which are simple efficient tools for the convergence analysis of iterative receivers [31], [32]. They are based on the flow of the extrinsic information exchanged between the SISO blocks in an iterative scheme and give insight into the convergence behavior of the receiver. In the following discussion, we briefly introduce this tool and refer to the provided references for more details.

In the EXIT chart analysis, the LLRs input to a SISO block are assumed to be uncorrelated and to follow a Gaussian distribution, with its mean being related to its variance [31]. Note that these assumptions are not perfectly satisfied in practice. As a result, EXIT charts do not exactly predict the convergence behavior of the receiver. For turbo-PIC, *a posteriori* (and not extrinsic) LLRs are fed from the decoder to the MIMO detector. However, it is quite logical to also consider the aforementioned assumptions for the *a posteriori* LLRs.

Let us use the subscripts $._{\mathrm{A}}$ (for *a priori*) and $._{\mathrm{E}}$ (for extrinsic) to denote the variables at the input and output of a SISO block, respectively. Note that, for the MIMO detector, the subscript $._{\mathrm{E}}$ corresponds to extrinsic LLRs, whereas for the decoder, it corresponds to *a posteriori* LLRs. In addition, let us denote by I_{A} and I_{E} the mutual information at the input and output of a SISO block, respectively. The behavior of the iterative detector is determined by associating $I_{\mathrm{E}}^{\mathrm{DET}} \Rightarrow I_A^{\mathrm{DEC}}$ and, inversely, $I_{\mathrm{E}}^{\mathrm{DEC}} \Rightarrow I_{\mathrm{A}}^{\mathrm{DET}}$, where the superscripts $.^{\mathrm{DET}}$ and $.^{\mathrm{DEC}}$ refer to the MIMO detector and the channel decoder, respectively.

We have shown in Fig. 8 the EXIT charts of the soft-PIC detector for two cases of mismatched and improved receivers, as well as that of the Max-Log-MAP decoder.[1] The exact formulation of the soft-PIC detector is considered. To better see the difference between the two detectors, we have considered the GLD ST scheme and set N_P to 2 and E_b/N_0 to a relatively low value, i.e., 5 dB. The iterative process is characterized by a trace between the EXIT curves of the MIMO detector and the decoder. This is shown in the figure for the improved detector. Notice that, at the first iteration, where there is no *a priori* information on the transmitted symbols available, i.e., $I_{\mathrm{A}}^{\mathrm{DET}} = 0$, we perform an MMSE detection, and the EXIT characteristic of the soft-PIC is specified by a single point. For the next iterations, where we perform soft interference cancellation, the EXIT curve is obtained for $0 \leq I_{\mathrm{A}}^{\mathrm{DET}} \leq 1$. As a result, the EXIT trace starts at the point that corresponds to the first iteration of the detector and proceeds according to the curve corresponding to the next iterations. Note that the obtained trajectory provides

[1]Note that it is not really correct to use the term "EXIT" chart for the SISO decoder in turbo-PIC, because it delivers *a posteriori* LLRs at its output. In addition, note the particular form of the corresponding curve in Fig. 8, which is due to the same reason.

an asymptotic (or an approximate) description of the receiver convergence. In particular, for the assumption of uncorrelated LLRs, this trajectory would correspond to the case of perfect interleaving with infinite interleaver size.

We notice that the EXIT curve of the improved detector (for any iteration) lays above that of the mismatched detector. Interestingly, the distance between the curves of improved and mismatched detectors becomes much more considerable by increasing the *a priori* information $I_{\mathrm{A}}^{\mathrm{DET}}$. This result confirms the lower BER after convergence for the improved detector. Note that the final BER depends on the intersection point of the EXIT curves of the detector and the decoder. The closer the corresponding $I_{\mathrm{E}}^{\mathrm{DEC}}$ is to one, the lower the BER becomes. The convergence rate is almost the same for the two detectors, because their EXIT curves have almost the same slope. We notice that both improved and mismatched detectors require about only two iterations to achieve the decoder output mutual information very close to one. In other words, by using the improved soft-PIC detector, in practice, no improvement is achieved in terms of the receiver convergence rate. This result is in accordance with the results in Fig. 4.

VIII. Conclusion and Discussions

By introducing a Bayesian approach that characterizes the channel-estimation process, we have proposed a general detector design that takes into account the imperfect channel estimate. Using this design, we derived two improved iterative MIMO receivers based on MAP and soft-PIC detectors that mitigate the impact of channel uncertainty on the detection performance. The cases of the exact and the simplified implementations of the soft-PIC detector were both treated. The formulation of the improved detector was provided for the simple case of the V-BLAST scheme. We have also provided its extension to the more general case of LD codes. We showed that the classically used mismatched detector becomes suboptimal once it is compared with the proposed detectors, in particular when only few pilots are devoted for channel estimation. The important point is that the performance improvement is obtained while imposing no considerable increase on the receiver's complexity. For ST schemes designed by imposing more constraints on the coding rate and/or diversity, e.g., the full-rate full-diversity GLD scheme, the receiver is more sensitive to the channel-estimation errors, and the improved detector provides even more important gains. By analyzing the convergence behavior of the iterative receiver, we showed that the improved soft-PIC detector does not bring a significant improvement in the convergence rate compared with the mismatched detector.

The amount of performance improvement provided by the improved detector is a function of the channel-estimation error variance. In the presented results, we have noticed that a relatively small performance improvement is achieved, except for a very small number of pilot symbols. Thus, the proposed improved detector is particularly interesting under the conditions of relatively fast fading. In such cases, it is important to transmit as few pilots as possible to avoid a significant loss in spectral efficiency. In other words, we can use few pilot symbols and employ the improved detector, hence reducing the loss in data rate due to pilot insertion.

One should also note that, in the presented results, we have set equal average power for pilot and data symbols. Obviously, if we allocate less power to pilots, the improved detector will provide larger gains with respect to the mismatched detector. In other words, by employing the improved detector, we can reduce the pilot's power and dedicate more power to data symbols, hence improving the system performance in terms of channel capacity. Again, this result is of particular interest in relatively fast-fading channels.

Finally, note that, although we have considered LS channel estimation in this work, the extension of the proposed detector to the MMSE channel estimation is rather straightforward.

Appendix I

A. Derivation of the a Posteriori Probability (10)

The following theorem is derived in [33].

Theorem 1.1: Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be circularly symmetric complex Gaussian random vectors with zero means and full-rank covariance matrices $\mathbf{\Sigma}_{ij} = \mathbb{E}[\mathbf{x}_i\mathbf{x}_j^\dagger]$. Then, the conditional random vector $\mathbf{x}_1|\mathbf{x}_2 \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{\Sigma})$ is circularly symmetric complex Gaussian with mean $\boldsymbol{\mu} = \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{x}_2$ and covariance matrix $\mathbf{\Sigma} = \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$.

We denote by $\mathbf{H}_i$ and $\widehat{\mathbf{H}}_i$ the ith column of matrices $\mathbf{H}$ and $\widehat{\mathbf{H}}$, respectively, and set $\mathbf{x}_1 = \mathbf{H}_i$ and $\mathbf{x}_2 = \widehat{\mathbf{H}}_i$. Based on (2) and (9), we have $\mathbf{\Sigma}_{11} = \mathbf{\Sigma}_{12} = \mathbf{\Sigma}_H$ and $\mathbf{\Sigma}_{22} = \mathbf{\Sigma}_H + \mathbf{\Sigma}_{\mathcal{E}}$ in Theorem 1.1. According to the theorem, the conditional pdf of $\mathbf{H}_i|\widehat{\mathbf{H}}_i$ is a circularly symmetric complex Gaussian distribution with

$$\text{mean} = \mathbf{\Sigma}_\Delta \widehat{\mathbf{H}}_i$$

where

$$\mathbf{\Sigma}_\Delta \triangleq \mathbf{\Sigma}_H(\mathbf{\Sigma}_H + \mathbf{\Sigma}_{\mathcal{E}})^{-1} \tag{49}$$

$$\begin{aligned}\text{covariance matrix} &= \mathbf{\Sigma}_H - \mathbf{\Sigma}_H(\mathbf{\Sigma}_H + \mathbf{\Sigma}_{\mathcal{E}})^{-1}\mathbf{\Sigma}_H^\dagger \\ &= \mathbf{\Sigma}_\Delta\mathbf{\Sigma}_{\mathcal{E}}.\end{aligned} \tag{50}$$

The equivalence in (50) is shown by left multiplying both sides of $(\mathbf{\Sigma}_{\mathcal{E}} + \mathbf{\Sigma}_H) - \mathbf{\Sigma}_H^\dagger = \mathbf{\Sigma}_{\mathcal{E}}$ by $\mathbf{\Sigma}_H(\mathbf{\Sigma}_{\mathcal{E}} + \mathbf{\Sigma}_H)^{-1}$.

Assuming the same covariance matrix for all columns of $\mathbf{H}$ and $\widehat{\mathbf{H}}$, we obtain the *a posteriori* pdf (10).

B. Evaluation of the Likelihood Function (17)

To evaluate the conditional expectation in (17), we use the following theorem from [34].

Theorem 1.2: For a circularly symmetric complex random vector $\boldsymbol{u} \sim \mathcal{CN}(\mathbf{m}, \mathbf{\Sigma})$ with mean $\mathbf{m} = \mathbb{E}[\boldsymbol{u}]$, covariance matrix $\mathbf{\Sigma} = \mathbb{E}[\boldsymbol{u}\boldsymbol{u}^\dagger] - \mathbf{m}\mathbf{m}^\dagger$, and a Hermitian matrix $\mathbf{A}$ such that $\mathbb{I} + \mathbf{\Sigma}\mathbf{A} > 0$, we have

$$\mathbb{E}_{\boldsymbol{u}}\left[\exp\{-\boldsymbol{u}^\dagger\mathbf{A}\boldsymbol{u}\}\right] = \frac{\exp\left\{-\mathbf{m}^\dagger\mathbf{A}(\mathbb{I} + \mathbf{\Sigma}\mathbf{A})^{-1}\mathbf{m}\right\}}{\det\{\mathbb{I} + \mathbf{\Sigma}\mathbf{A}\}}. \tag{51}$$

Let us define $\boldsymbol{u} = \boldsymbol{y}_k - \mathbf{H}\boldsymbol{s}_k$. Using the *a posteriori* distribution of (10) and after some algebra, we can derive the conditional pdf of $\boldsymbol{u}$, given $\boldsymbol{s}_k$ and $\widehat{\mathbf{H}}$ as $\boldsymbol{u}|(\boldsymbol{s}_k, \widehat{\mathbf{H}}) \sim \mathcal{CN}(\mathbf{m}_u, \mathbf{\Sigma}_u)$, where $\mathbf{m}_u = \boldsymbol{y}_k - \boldsymbol{\Sigma}_\Delta \widehat{\mathbf{H}}\boldsymbol{s}_k$, and $\boldsymbol{\Sigma}_u = \boldsymbol{\Sigma}_\Delta \boldsymbol{\Sigma}_\mathcal{E} \|\boldsymbol{s}_k\|^2$. We further define $\mathbf{A} = \boldsymbol{\Sigma}_z^{-1}$. By applying Theorem 1.2, (17) is written as (52), shown at the bottom of page. $\boldsymbol{\Sigma}_z$, $\boldsymbol{\Sigma}_\Delta$ and $\boldsymbol{\Sigma}_\mathcal{E}$ are diagonal matrices; thus, the latter equation is rewritten as

$$\begin{aligned}
&\widetilde{W}(\boldsymbol{y}_k|\widehat{\mathbf{H}}, \boldsymbol{s}_k)\\
&= \frac{\exp\left\{-(\boldsymbol{y}_k - \delta\widehat{\mathbf{H}}\boldsymbol{s}_k)^\dagger \left(\boldsymbol{\Sigma}_z + \delta\boldsymbol{\Sigma}_\mathcal{E}\|\boldsymbol{s}_k\|^2\right)^{-1}(\boldsymbol{y}_k - \delta\widehat{\mathbf{H}}\boldsymbol{s}_k)\right\}}{\det\left\{\pi\left(\boldsymbol{\Sigma}_z + \delta\boldsymbol{\Sigma}_\mathcal{E}\|\boldsymbol{s}_k\|\right)\right\}}\\
&= \mathcal{CN}\left(\delta\widehat{\mathbf{H}}\boldsymbol{s}_k, \boldsymbol{\Sigma}_z + \delta\boldsymbol{\Sigma}_{calE}\|\boldsymbol{s}_k\|^2\right). \qquad (53)
\end{aligned}$$

APPENDIX II
DETAILS ON THE FORMULATION OF THE IMPROVED SOFT-PIC

A. Derivation of the Improved MMSE Filter (38)

The inner expectations involved in (37) can easily be evaluated from (35) as

$$\mathbf{R}_{s_k^i \underline{\boldsymbol{y}}_k^i} = \mathbb{E}_{\boldsymbol{s}_k, \boldsymbol{z}_k}\left[s_k^i {\underline{\boldsymbol{y}}_k^i}^\dagger\right] = \sigma_{s_k^i}^2 \mathbf{h}_i^\dagger + \mathbf{m}_{k,i}(\underline{\mathbf{H}}_i - \widehat{\underline{\mathbf{H}}}_i)^\dagger \qquad (54)$$

$$\begin{aligned}
\mathbf{R}_{\underline{\boldsymbol{y}}_k^i} &= \mathbb{E}_{\boldsymbol{s}_k, \boldsymbol{z}_k}\left[\underline{\boldsymbol{y}}_k^i {\underline{\boldsymbol{y}}_k^i}^\dagger\right]\\
&= \sigma_{s_k^i}^2 \mathbf{h}_i\mathbf{h}_i^\dagger + \underline{\mathbf{H}}_i \boldsymbol{\Lambda}_{k,i}\underline{\mathbf{H}}_i^\dagger + \mathbf{h}_i\mathbf{m}_{k,i}\underline{\mathbf{H}}_i^\dagger + \underline{\mathbf{H}}_i\mathbf{m}_{k,i}^\dagger\mathbf{h}_i^\dagger\\
&\quad - \mathbf{h}_i\mathbf{m}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger - \widehat{\underline{\mathbf{H}}}_i\mathbf{m}_{k,i}^\dagger\mathbf{h}_i^\dagger - \underline{\mathbf{H}}_i\widetilde{\boldsymbol{\Lambda}}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger - \widehat{\underline{\mathbf{H}}}_i\widetilde{\boldsymbol{\Lambda}}_{k,i}\underline{\mathbf{H}}_i^\dagger\\
&\quad + \widehat{\underline{\mathbf{H}}}_i\widetilde{\boldsymbol{\Lambda}}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger + \sigma_z^2\mathbb{I}_{M_R}. \qquad (55)
\end{aligned}$$

Thus, we have to evaluate the outer expectations as follows:

$$\begin{aligned}
\overline{\mathbf{R}}_{s_k^i \underline{\boldsymbol{y}}_k^i} &= \mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}[\mathbf{R}_{s_k^i \underline{\boldsymbol{y}}_k^i}]\\
\overline{\mathbf{R}}_{\underline{\boldsymbol{y}}_k^i} &= \mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}[\mathbf{R}_{\underline{\boldsymbol{y}}_k^i}].
\end{aligned}$$

To compute for the aforementioned expectations, we use the following lemma [35].

Lemma 1: For a circularly symmetric complex Gaussian random rowwise vector $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and a Hermitian matrix $\mathbf{A}$, we have

$$\mathbb{E}_\mathbf{x}[\mathbf{x}\mathbf{A}\mathbf{x}^\dagger] = \mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}\mathbf{A}\boldsymbol{\mu}^\dagger. \qquad (56)$$

By applying Lemma 1 and using the *a posteriori* channel pdf (12), it is straightforward to find $\overline{\mathbf{R}}_{s_k^i \underline{\boldsymbol{y}}_k^i}$ in (39) as

$$\overline{\mathbf{R}}_{s_k^i \underline{\boldsymbol{y}}_k^i} = \mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left[\mathbf{R}_{s_k^i \underline{\boldsymbol{y}}_k^i}\right] = \delta\sigma_{s_k^i}^2\widehat{\mathbf{h}}_i^\dagger + (\delta - 1)\mathbf{m}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger. \qquad (57)$$

The evaluation of $\overline{\mathbf{R}}_{\underline{\boldsymbol{y}}_k^i}$ in (40) involves the following equalities:

$$\begin{aligned}
\mathbb{E}_{\mathbf{h}_i|\widehat{\mathbf{h}}_i}\left[\mathbf{h}_i\mathbf{h}_i^\dagger\right] &= \delta^2\widehat{\mathbf{h}}_i\widehat{\mathbf{h}}_i^\dagger + (1-\delta)\mathbb{I}_{M_R}\\
\mathbb{E}_{\mathbf{h}_i|\widehat{\mathbf{h}}_i}\left[\mathbf{h}_i\mathbf{m}_{k,i}\underline{\mathbf{H}}_i^\dagger\right] &= \delta^2\widehat{\mathbf{h}}_i\mathbf{m}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger\\
\mathbb{E}_{\underline{\mathbf{H}}_i|\widehat{\underline{\mathbf{H}}}_i}\left[\underline{\mathbf{H}}_i\widetilde{\boldsymbol{\Lambda}}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger\right] &= \delta\widehat{\underline{\mathbf{H}}}_i\widetilde{\boldsymbol{\Lambda}}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger\\
\mathbb{E}_{\underline{\mathbf{H}}_i|\widehat{\underline{\mathbf{H}}}_i}\left[\underline{\mathbf{H}}_i\boldsymbol{\Lambda}_{k,i}\underline{\mathbf{H}}_i^\dagger\right] &= \delta^2\widehat{\underline{\mathbf{H}}}_i\boldsymbol{\Lambda}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger + (1-\delta)\mathrm{tr}(\boldsymbol{\Lambda}_{k,i})\mathbb{I}_{M_R}.
\end{aligned}$$

Now, by using the aforementioned equalities, we obtain

$$\begin{aligned}
\overline{\mathbf{R}}_{\underline{\boldsymbol{y}}_k^i} &= \delta^2\sigma_{s_k^i}^2\widehat{\mathbf{h}}_i\widehat{\mathbf{h}}_i^\dagger + (1-\delta)\sigma_{s_k^i}^2\mathbb{I}_{M_R} + \sigma_z^2\mathbb{I}_{M_R} + \delta^2\widehat{\underline{\mathbf{H}}}_i\boldsymbol{\Lambda}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger\\
&\quad + (1-\delta)\mathrm{tr}(\boldsymbol{\Lambda}_{k,i})\mathbb{I}_{M_R} + \delta^2\widehat{\mathbf{h}}_i\mathbf{m}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger + \delta^2\widehat{\underline{\mathbf{H}}}_i\mathbf{m}_{k,i}^\dagger\widehat{\mathbf{h}}_i^\dagger\\
&\quad - \delta\widehat{\mathbf{h}}_i\mathbf{m}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger - \delta\widehat{\underline{\mathbf{H}}}_i\mathbf{m}_{k,i}^\dagger\widehat{\mathbf{h}}_i^\dagger - \delta\widehat{\underline{\mathbf{H}}}_i\widetilde{\boldsymbol{\Lambda}}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger\\
&\quad - \delta\widehat{\underline{\mathbf{H}}}_i\widetilde{\boldsymbol{\Lambda}}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger + \widehat{\underline{\mathbf{H}}}_i\widetilde{\boldsymbol{\Lambda}}_{k,i}\widehat{\underline{\mathbf{H}}}_i^\dagger. \qquad (58)
\end{aligned}$$

Equation (40) directly follows after rearranging the terms in (58). As a result, based on (57) and (58), we obtain the improved MMSE filter (38).

B. Derivation of the Variance $\sigma_{\eta_{k,i}}^2$ in (44)

Based on (27), we can evaluate the mean-square error (MSE) at the output of the turbo-PIC detector as

$$\begin{aligned}
\sigma_{\mathrm{MSE}}^2 &= \mathbb{E}\left[\left|s_k^i - \mathbf{w}_k^i\underline{\boldsymbol{y}}_k^i\right|^2\right]\\
&= \sigma_{s_k^i}^2 - \mathbf{R}_{s_k^i\underline{\boldsymbol{y}}_k^i}{\mathbf{w}_k^i}^\dagger - \mathbf{w}_k^i\mathbf{R}_{\underline{\boldsymbol{y}}_k^i s_k^i} + \mathbf{w}_k^i\mathbf{R}_{\underline{\boldsymbol{y}}_k^i}{\mathbf{w}_k^i}^\dagger. \qquad (59)
\end{aligned}$$

$\mathbf{w}_k^i = \mathbf{R}_{s_k^i\underline{\boldsymbol{y}}_k^i}\mathbf{R}_{\underline{\boldsymbol{y}}_k^i}^{-1}$; thus, we have $\mathbf{R}_{s_k^i\underline{\boldsymbol{y}}_k^i}\mathbf{w}_k^{i\dagger} = \mathbf{w}_k^i\mathbf{R}_{\underline{\boldsymbol{y}}_k^i}\mathbf{w}_k^{i\dagger}$. Consequently, (59) reduces to

$$\sigma_{\mathrm{MSE}}^2 = \sigma_{s_k^i}^2\mathbf{w}_k^i\mathbf{R}_{\underline{\boldsymbol{y}}_k^i s_k^i}. \qquad (60)$$

For the case of the improved detector, after using $\widetilde{\mathbf{w}}_k^i$ and $\overline{\mathbf{R}}_{s_k^i\underline{\boldsymbol{y}}_k^i}$ from (38) and (39) instead of $\mathbf{w}_k^i$ and $\mathbf{R}_{s_k^i\underline{\boldsymbol{y}}_k^i}$ in (60), we obtain

$$\sigma_{\text{MSE-IM}}^2 = \sigma_{s_k^i}^2(1 - \mu_{k,i}) - (\delta - 1)\widetilde{\mathbf{w}}_k^i\widehat{\underline{\mathbf{H}}}_i\mathbf{m}_{k,i}^\dagger \qquad (61)$$

where $\mu_{k,i} = \delta\widetilde{\mathbf{w}}_k^i\widehat{\mathbf{h}}_i$.Alternatively, based on (43), we have

$$\begin{aligned}
\sigma_{\text{MSE-IM}}^2 &= \mathbb{E}\left[\left|s_k^i - \left(\mu_{k,i}s_k^i + \eta_{k,i}\right)\right|^2\right]\\
&= (1-\mu_{k,i})^2\sigma_{s_k^i}^2 + \sigma_{\eta_k^i}^2 - 2\mathbb{R}e\\
&\quad \times \left(\left(\delta - 1\right)\left(1 - \mu_{k,i}^*\right)\widetilde{\mathbf{w}}_k^i\widehat{\underline{\mathbf{H}}}_i\mathbf{m}_{k,i}^\dagger\right). \qquad (62)
\end{aligned}$$

$$\begin{aligned}
\widetilde{W}(\boldsymbol{y}_k|\widehat{\mathbf{H}}, \boldsymbol{s}_k) &= \mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left[\frac{\exp\left\{-(\boldsymbol{y}_k - \mathbf{H}\boldsymbol{s}_k)^\dagger\boldsymbol{\Sigma}_z^{-1}(\boldsymbol{y}_k - \mathbf{H}\boldsymbol{s}_k)\right\}}{\det\{\pi\boldsymbol{\Sigma}_z\}}\right]\\
&= \frac{\exp\left\{-(\boldsymbol{y}_k - \boldsymbol{\Sigma}_\Delta\widehat{\mathbf{H}}\boldsymbol{s})^\dagger\boldsymbol{\Sigma}_z^{-1}\left(\mathbb{I} + \boldsymbol{\Sigma}_\Delta\boldsymbol{\Sigma}_\mathcal{E}\|\boldsymbol{s}_k\|^2\boldsymbol{\Sigma}_z^{-1}\right)^{-1}(\boldsymbol{y}_k - \boldsymbol{\Sigma}_\Delta\widehat{\mathbf{H}}\boldsymbol{s}_k)\right\}}{\det\left(\pi\boldsymbol{\Sigma}_z\left(\mathbb{I} + \boldsymbol{\Sigma}_\Delta\boldsymbol{\Sigma}_\mathcal{E}\|\boldsymbol{s}_k\|^2\boldsymbol{\Sigma}_z^{-1}\right)\right)} \qquad (52)
\end{aligned}$$

Comparing (61) and (62) leads to

$$\sigma^2_{\eta^i_k} = \left(\mu_{k,i} - \mu^2_{k,i}\right)\sigma^2_{s^i_k} + (\delta - 1) \times \left[2\mathbb{R}e\left(\left(1-\mu^*_{k,i}\right)\widetilde{\mathbf{w}}^i_k \underline{\widehat{\mathbf{H}}}_i \mathbf{m}^\dagger_{k,i}\right) - \widetilde{\mathbf{w}}^i_k \underline{\widehat{\mathbf{H}}}_i \mathbf{m}^\dagger_{k,i}\right]. \quad (63)$$

Notice that, by setting δ to be equal to one (which corresponds to perfect CSIR) in (63), we retrieve the classical expression derived in the literature, i.e., $\sigma^2_{\eta^i_k} = (\mu_{k,i} - \mu^2_{k,i})\sigma^2_{s^i_k}$ (for example, see [16] and [29]).

REFERENCES

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[2] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[3] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 686–693, Nov. 1991.

[4] P. Garg, R. K. Mallik, and H. M. Gupta, "Performance analysis of space-time coding with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 257–265, Jan. 2005.

[5] G. Taricco and E. Biglieri, "Space-time decoding with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1874–1888, Jul. 2005.

[6] K. Ahmed, C. Tepedelenhoglu, and A. Spanias, "Effect of channel estimation on pair-wise error probability in OFDM," in *Proc. IEEE ICASSP*, May 2004, pp. 745–748.

[7] A. Leke and J. M. Cioffi, "Impact of imperfect channel knowledge on the performance of multicarrier systems," in *Proc. IEEE GLOBECOM*, Nov. 1998, pp. 951–955.

[8] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criteria in the presence of channel estimation errors, mobility, and multiple paths," *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 199–207, Feb. 1999.

[9] S. M. S. Sadough, P. Piantanida, and P. Duhamel, "MIMO-OFDM optimal decoding and achievable information rates under imperfect channel estimation," in *Proc. SPAWC*, Jun. 2007, pp. 1–5.

[10] S. M. S. Sadough and P. Duhamel, "Improved iterative detection and achieved throughputs of OFDM systems under imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5039–5050, Dec. 2008.

[11] S. M. S. Sadough, P. Piantanida, and P. Duhamel, "Achievable outage rates with improved decoding of BICM multiband OFDM under channel estimation errors," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Oct. 2006, pp. 1806–1810.

[12] S. M.-S. Sadough, M. M. Ichir, P. Duhamel, and E. Jaffrot, "Wavelet based semi-blind channel estimation for ultra wideband OFDM systems," *IEEE Trans. Veh. Technol.*, to be published.

[13] A. M. Tonello, "Space-time bit-interleaved coded modulation with an iterative decoding strategy," in *Proc. Veh. Technol. Conf.*, Sep. 2000, pp. 473–478.

[14] J. J. Boutros, F. Boixadera, and C. Lamy, "Bit-interleaved coded modulations for multiple-input multiple-output channels," in *Proc. Int. Symp. Spread Spectrum Tech. Appl.*, Sep. 2000, pp. 123–126.

[15] M. A. Khalighi and J. J. Boutros, "Semi-blind channel estimation using EM algorithm in iterative MIMO APP detectors," *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 3165–3173, Nov. 2006.

[16] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046–1061, Jul. 1999.

[17] M. Sellathurai and S. Haykin, "Turbo-BLAST for wireless communications: Theory and experiments," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2538–2546, Oct. 2002.

[18] H. Lee, B. Lee, and I. Lee, "Iterative detection and decoding with improved V-BLAST for MIMO-OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 504–513, Mar. 2006.

[19] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–945, May 1998.

[20] J. Balakrishnan, M. Rupp, and H. Viswanathan, "Optimal channel training for multiple antenna systems," in *Proc. Conf. Multiaccess, Mobility Teletraffic Wireless Commun.*, Miami, FL, Dec. 2000, pp. 25–36.

[21] P. Robertson, P. Hoeher, and E. Villebrun, "Optimal and suboptimal maximum *a posteriori* algorithms suitable for turbo decoding," *Eur. Trans. Telecommun.*, vol. 8, no. 2, pp. 119–125, Mar./Apr. 1997.

[22] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 284–287, Mar. 1974.

[23] M. A. Khalighi, J. Boutros, and J.-F. Hélard, "Data-aided channel estimation for turbo-PIC MIMO detectors," *IEEE Commun. Lett.*, vol. 10, no. 5, pp. 350–352, May 2006.

[24] L. Scharf, *Statistical Signal Processing*. Reading, MA: Addison-Wesley, 1991.

[25] T. Oomori, T. Asai, and T. Matsumoto, "A matched filter approximation for SC/MMSE turbo equalizers," *IEEE Commun. Lett.*, vol. 5, no. 7, pp. 310–312, Jul. 2001.

[26] M. A. Khalighi, J.-F. Hélard, and S. Bourennane, "Choice of appropriate space-time coding scheme for MIMO systems employing channel coding under BICM," in *Proc. IEEE SPAWC*, Cannes, France, Jul. 2006, pp. 1–5.

[27] M. A. Khalighi, J.-F. Hélard, and S. Bourennane, "Contrasting orthogonal and nonorthogonal space-time schemes for perfectly known and estimated MIMO channels," in *Proc. ICCS*, Singapore, Oct./Nov. 2006, pp. 1–5.

[28] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1804–1824, Jul. 2002.

[29] A. Dejonghe and L. Vandendorpe, "Turbo-equalization for multilevel modulation: An efficient low-complexity scheme," in *Proc. ICC*, 2002, pp. 1863–1867.

[30] J.-C. Belfiore, G. Rekaya, and E. Viterbo, "The golden code: A 2×2 full-rate space-time code with nonvanishing determinants," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1432–1436, Apr. 2005.

[31] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.

[32] M. Tuchler, R. Koetter, and A. C. Singer, "Turbo equalization: Principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754–767, May 2002.

[33] M. Bilodeau and D. Brenner, *Theory of Multivariate Statistics*. New York: Springer-Verlag, 1999.

[34] M. Schwartz, W. R. Bennett, and S. Stein, *Communications Systems and Techniques*. New York: McGraw-Hill, 1966.

[35] G. Seber and A. Lee, *Linear Regression Analysis*. Hoboken, NJ: Wiley, 2002.

Seyed Mohammad-Sajad Sadough (S'04–M'08) was born in Paris, France, in 1979. He received the B.Sc. degree in electrical engineering (electronics) from the Shahid Beheshti University, Tehran, Iran, in 2002 and the M.Sc. and Ph.D. degrees in electrical engineering (telecommunications) from Paris Sud 11 University, Orsay, France, in 2004 and 2008, respectively.

From 2004 to 2007, he was with the National Engineering School in Advanced Techniques (ENSTA), Paris, and the Laboratory of Signals and Systems, Supélec—National Scientific Research Center (CNRS), Gif sur Yvette, France. He was a Lecturer with the Department of Electronics and Computer Engineering, ENSTA, where his research activities were focused on improved reception schemes for ultrawideband communication systems. From December 2007 to September 2008, he was a Postdoctoral Researcher with the laboratoire des Signaux et Systèmes (LSS), Supélec-CNRS, where he was involved in research projects with Alcatel-Lucent on satellite mobile communication systems. Since October 2008, he has been with the Faculty of Electrical and Computer Engineering, Shahid Beheshti University, where he is currently an Assistant Professor with the Department of Telecommunications. His research interests include signal processing for wireless communications, with particular emphasis on multicarrier and multiple-input–multiple-output systems, joint channel estimation and decoding, iterative reception schemes, and interference cancellation under partial channel-state information.

Mohammad-Ali Khalighi (SM'07) received the Ph.D. degree in electrical engineering from the Institut National Polytechnique de Grenoble, Grenoble, France, in 2002.

From 2002 to 2005, he was with Grenoble Images Parole Signal Automatique (GIPSA)-Lab, the Ecole Nationale Superieure des Télécommunications—Telecom ParisTech, Paris, France, and the Institut d'Électronique et de Télécommunications de Rennes Laboratory, Rennes, France, as a Postdoctoral Research Fellow. Since 2005, he has been with the École Centrale Marseille and the Institut Fresnel, Marseille, France, where he is currently an Assistant Professor. His research interests include coding, signal detection, and channel estimation for high-data-rate communication systems.

Pierre Duhamel (M'87–SM'87–F'98) was born in France in 1953. He received the B.Eng. degree in electrical engineering from the National Institute for Applied Sciences, Rennes, France, in 1975 and the Dr.Eng. and Doctorat ès Sciences degrees from Orsay University, Orsay, France, in 1978 and 1986, respectively.

From 1975 to 1980, he was with Thomson-CSF, Paris, France, working on circuit theory and signal processing, including digital filtering and analog fault diagnosis. In 1980, he joined the National Research Center in Telecommunications, Issy les Moulineaux, France, where his research activities were first concerned with the design of recursive charged-coupled device filters. Later, he worked on fast algorithms for computing Fourier transforms and convolutions and applied similar techniques to adaptive filtering, spectral analysis, and wavelet transforms. From 1993 to September 2000, he was a Professor with the National School of Engineering in Telecommunications, Paris, where his research activities focused on signal processing for communications, and where he was the Head of the Signal and Image Processing Department from 1997 to 2000. He is currently with the National Scientific Research Centre/Laboratory of Signals and Systems—Supélec, Gif sur Yvette, France, where he is developing studies in signal processing for communications (including equalization, iterative decoding, multicarrier systems, and cooperation) and signal/image processing for multimedia applications, including source coding, joint source/channel coding, watermarking, and audio-processing. He is currently investigating the application of recent information theory results to communication theory.

Dr. Duhamel was the Chair of the Digital Signal Processing Committee from 1996 to 1998 and a member of the signal processing for Com Committee until 2001. He was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1989 to 1991, an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS, and a Guest Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING Special Issue on Wavelets. He was a Distinguished Lecturer of the IEEE in 1999 and was a General Cochair of the 2001 International Workshop on Multimedia Signal Processing, Cannes, France. He was also a Technical Cochair of the 2006 IEEE International Conference on Acoustics, Speech, and Signal Processing, Toulouse, France. He is a coauthor of the paper on subspace-based methods for blind equalization, for which he received the Best Paper Award from the IEEE TRANSACTIONS ON SIGNAL PROCESSING in 1998. He received the Grand Prix France Telecom from the French Science Academy in 2000.