Iterative Equalization with Decision Feedback based on Expectation Propagation

Serdar Sahin, Antonio Cipriano, Charly Poulliat, Marie-Laure Boucheret

Iterative Equalization With Decision Feedback Based on Expectation Propagation

Serdar Şahin, Antonio Maria Cipriano, Charly Poulliat, and Marie-Laure Boucheret

***Abstract*— This paper investigates the design and analysis of minimum mean square error (MMSE) turbo decision feedback equalization (DFE), with expectation propagation (EP), for single carrier modulations. Classical non iterative DFE structures have substantial advantages at high-data rates, even compared with *turbo* linear equalizers-interference cancellers (LE-IC), hence making turbo DFE-IC schemes an attractive solution. In this paper, we derive an iterative DFE-IC, capitalizing on the use of soft feedback based on expectation propagation, along with the use of prior information for improved filtering and interference cancellation. This turbo iterative DFE-IC significantly outperforms turbo LE-IC, especially at high-spectral efficiency and also exhibits performance improvements over existing DFE-IC variants. The proposed scheme can also be self-iterated, as done in the recent trend on EP-based equalizers, and it is shown to be an attractive alternative to linear self-iterated receivers. For time-varying (TV) filter equalizers, an efficient matrix inversion scheme is also proposed, considerably reducing the computational complexity relative to existing methods. Using finite-length and asymptotic analysis on a severely selective channel, the proposed DFE-IC is shown to achieve higher rates than known alternatives, with better waterfall thresholds and faster convergence, while keeping a similar computational complexity.**

***Index Terms*— Interference cancellation, turbo equalization, decision feedback equalizers, expectation propagation.**

I. Introduction

COMMUNICATION systems operating on wide-band channels suffer from inter-symbol interference (ISI), which can be mitigated with an appropriate transceiver design. In particular, for wireless systems where the throughput requirements increase at each new generation, more effective receivers are needed in order to maintain robust data links.

With the discovery of turbo-codes, iterative processing principles were extended to joint detection and decoding techniques via soft-input soft-output (SISO) receivers which use prior information provided by the channel decoder, to further reduce detection errors. Although early turbo equalization techniques, such as maximum a posteriori (MAP) detector using BCJR estimation [1]–[3], can operate near the channel capacity with properly designed coding schemes, their operational complexity significantly increases for large channel delay spread or with high modulation orders. Consequently, finite impulse response (FIR) filter-based turbo equalizers with lowered computational complexity have been proposed. These structures can be categorized into three groups with regards to its filter updates depending on prior information. Other kinds of adaptive FIR receivers are out of this paper's scope. Time-invariant (TI) structures update their filters only once at each packet reception, using the available channel state. Iteration-variant (IV) equalizers are updated at each turbo iteration by additionally using the overall prior information. Time-varying (TV) structures update their filters at each symbol, using both symbol-wise prior information and channel states, making them particularly suitable for doubly selective channels, where the impulse response varies over time.

The first FIR turbo structure, proposed by Glavieux *et al.* [4], uses a time-invariant interference canceller [5], and an application to IV filtering appeared in [6] and [7]. Further extension to TV equalization is provided in [8] and a formal framework presented in [9] derive these receivers from the MAP criterion.

An alternative approach formalized by Tüchler *et. al* [10] consists in designing a TV adaptive LE, by using statistics conditioned on prior information, while solving the MMSE criterion. This structure has been applied to high-order modulations, time-varying channels and to IV, TI, frequency domain structures for lower complexity, and also to multi-user detection for multiple input-multiple output systems [11]–[14].

Equivalence of these approaches was shown in [15], making the TV MMSE LE-IC the most widespread reference. Although turbo LE-IC brings significant improvements over classical filtering, it falls far behind classical DFE [16], [17] with channel coding, at high spectral efficiency operating points. Oppositely, at lower rates, turbo LE-IC is near capacity-achieving while DFE performs poorly.[1]

This paper addresses the design of iterative time-domain TV DFE-IC equalizers, i.e. FIR receivers where prior information and a symbol-wise decision feedback is respectively used on anti-causal and causal symbols, to improve equalization. These

Corresponding author: Serdar Şahin.

S. Şahin is with Thales Communications and Security, 92230 Gennevilliers, France, and also with IRIT/INP Toulouse-ENSEEIHT, 31000 Toulouse, France (e-mail: serdar.sahin@thalesgroup.com).

A. M. Cipriano is with Thales Communications and Security, 92230 Gennevilliers, France (e-mail: antonio.cipriano@thalesgroup.com).

C. Poulliat and M.-L. Boucheret are with IRIT/INP Toulouse-ENSEEIHT, 31000 Toulouse, France (e-mail: charly.pouillat@enseeiht.fr; marie-laure.boucheret@enseeiht.fr).

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/TCOMM.2018.2843760

[1]These facts are also shown in subsection V-C, in Fig. 7.

TABLE I

CLASSIFICATION OF CONSTITUENT FIR TURBO EQUALIZERS VS. THE USAGE OF PRIOR INFORMATION

	Linear Structure			Decision Feedback Structures			
Update Type	TI	IV	TV	Dec. Type	TI	IV	TV
References	[4], [19]	[6], [7], [9], [10], [12], [14], [15]	[8]-[11], [13], [19]	Hard	[20], [21]	[20], [21]	[19]-[21]
				Soft APP	[26]	[23], [24], [27], [29]	**Proposed**
				Soft EP			**Proposed**

receivers are of interest for applications where doubly-selective channels are involved, such as HF communications [18].

A. Related Work

There exist several prior works on DFE-IC. Proposals mainly differ with the nature of decision feedback, and with the filter updating method. Besides, recent complex receivers use DFEs as constituent elements for concatenated equalizers. Hence, for clarity, we propose to classify related works in three sub-categories.

1) Iterative Hard DFE-IC: Among hard feedback structures, DFE-IC in [19] is a classical DFE that uses prior information for IC on anti-causal symbols. This structure is known for its error propagation issues which makes its TV form even less efficient than TI LE, and its extrinsic information transfer (EXIT) analysis yields contradictory results [19, Fig. 14]. In [20], the previous structure is enhanced with a powerful soft demapper that uses the distribution of residual ISI sequences for symbol detection. This modified structure outperforms turbo LE-IC, but this residual ISI distribution is very difficult to derive even in the simple BPSK case. A more practical solution, proposed in [21], consists in approximating the residual ISI at the DFE-IC output to an additive white Gaussian noise (AWGN), which simplifies the demapper. While this solution challenges TV LE-IC on BPSK, its extension to multilevel modulations has not been explored so far. To the authors' knowledge, this is the only DFE-IC outperforming exact TV LE-IC in the reference scenario of Proakis-C channel with BPSK symbols. DFE-ICs in [20] and [21] were later used as constituent elements for more advanced receivers such as bi-directional DFE, or structures obtained by parallel concatenation of FIRs [22].

2) Iterative Soft DFE-IC: Literature on turbo soft DFE-IC is more diverse; although feedback is mostly based on the posterior distribution, there is no common strategy for evaluating its variance [23]–[25]. Such iterative structure is first presented in [26], where various TI DFE with soft feedback are evaluated with a perfect decision hypothesis, within a sub-optimal receiver using hard decoding. In particular, it is seen that soft feedback mitigates to some extent error propagation, despite ignoring decision errors in filter computation. Another notable structure is the IV soft interference canceller in [23]; using both prior and posterior LLRs for filtering and for interference cancellation with BPSK, this scheme significantly outperforms IV LE-IC, but it requires stochastic methods for estimating the correlation properties of posterior LLRs. Several other IV soft feedback structures exist [25], [27], with alternative heuristics for feedback quality assessment. Structural comparison of IV schemes using posterior feedback is given in [24], extending [23] and [27] to higher order modulations, but requiring new heuristics with LE-IC pre-equalization for filter computation. These approaches have drawbacks due to their limitations in usable constellations [23], [25], [27], or due to the sub-optimality of heuristics used in filter computation [23], [24], [27]. Indeed, IV structures need static statistics of its soft feedback for computing its filters, which requires approximations.

Time-varying soft posterior feedback structures do not have such issues; they can update their filters after each symbol is detected, as it had been done for MIMO receivers in [28]. In equalization, the structure closest to [28] is a block-feedback turbo DFE in [29], which updates its filters every P symbols. A classification of the references above is given in Table I.

3) Receivers Based on Expectation Propagation: There is a recent renewal of interest in iterative equalization, brought by the use of an approximate statistical inference method, namely expectation propagation (EP) [30]. This technique can be used as a message passing algorithm, which extends the loopy belief propagation (BP) by using exchange of expectations. When EP is used with probability density functions (PDF) belonging to the exponential family, it is possible to compute an extrinsic message passed from the demapper to the equalizer.

EP has already been used in channel decoding [31], and in receiver design with MIMO receivers [32], block linear equalizers [33] and Kalman smoothers [34], [35]. In particular, a concomitant work has recently extended these schemes to FIR with a self-iterated LE-IC [36]. In [37], EP was applied on multivariate white Gaussian distributions to derive a low-complexity self-iterated frequency domain equalizer. The receivers above use EP in a parallel interference cancellation scheduling through self-iterations, i.e. the whole data block is detected, and then detection process is repeated using EP feedback from the demapper. These structures are not decision feedback structures as in [16], which are natural successive interference cancellers.

Hence, in this paper, we propose to derive a *DFE-IC EP* exploiting the successive interference cancellation schedule of DFE-IC to operate on an EP-based soft feedback. Moreover, we combine this serial detection framework with an outer loop, as in prior work on EP, to obtain a *self-iterated DFE-IC EP*. A low complexity matrix inversion strategy for TV FIR structures is also derived, significantly reducing the computational complexity difference between DFE-IC and LE-IC.

B. Contributions and Paper Outline

The main contributions of this paper are as follows:

- A novel time-varying DFE-IC algorithm, using EP to update its filters, and to cancel residual ISI, is proposed.

It outperforms other constituent FIR receivers known to the authors, while providing an overall efficient complexity-performance trade-off.
- DFE-IC EP is extended to a self-iterated structure, and compared to prior work on self-iterated EP receivers.
- Well-known hard [19], [21] or sub-optimal [24], [29] DFE-IC proposals are extended to TV structures with soft posterior feedback, by using MMSE Bayesian estimators.
- Analytical and asymptotic analysis of DFE-IC is carried out on a highly selective deterministic channel. Performance and computational complexity comparison between LE-IC and different DFE-IC structures is provided.
- A new recursive matrix inversion strategy for TV equalizers is exposed. Compared to the iterative algorithm in [10], it brings between 30% (for long data blocks) and 75% (for shorter blocks) complexity reduction for LE-IC.

The remainder of this paper is organized as follows. The considered BICM communication scheme and the generic FIR receiver model are described in section II. Section III proposes a factor graph model for the system and applies the expectation propagation framework to derive the proposed equalizer in subsection III-D. A novel matrix inversion strategy is detailed in section IV for reducing TV equalization complexity. Section V extends prior work on DFE-IC to the state-of-the-art and compares with the proposed DFE-IC EP. In section VI, DFE-IC EP is self-iterated, and compared with several existing self-iterated EP receivers.

C. Notations

Bold lowercase letters are used for vectors: let $\mathbf{u}$ be a $N \times 1$ vector, then $u_n, n = 0, \dots, N-1$ are its entries. Capital bold letters denote matrices: for a given $N \times M$ matrix $\mathbf{A}$, $[\mathbf{A}]_{n,:}$ and $[\mathbf{A}]_{:,m}$ respectively denote its n^{th} row and m^{th} column, and $a_{n,m} = [\mathbf{A}]_{n,m}$ is the entry (n, m).

$\mathbf{I}_N$ is the $N \times N$ identity matrix, $\mathbf{0}_{N,M}$ and $\mathbf{1}_{N,M}$ are respectively all zeros and all ones $N \times M$ matrices. $\mathbf{e}_n$ is the $N \times 1$ indicator whose only non-zero entry is $e_n = 1$. Operator $\mathbf{Diag}(\mathbf{u})$ denotes the diagonal matrix whose diagonal is defined by $\mathbf{u}$. $\mathbb{R}, \mathbb{C}$, and $\mathbb{F}_k$ are respectively the real field, the complex field and a Galois field of order k. Let x and y be two random variables, then $\mu_x = \mathbb{E}[x]$ is the expected value, $\sigma_x^2 = \text{Var}[x]$ is the variance and $\sigma_{x,y} = \text{Cov}[x, y]$ is the covariance. The probability of x taking a value α is $\mathbb{P}[x = \alpha]$, and probability density functions (PDF) are denoted as $p(\cdot)$. If $\mathbf{x}$ and $\mathbf{y}$ are random vectors, then we define vectors $\boldsymbol{\mu}_\mathbf{x} = \mathbb{E}[\mathbf{x}]$ and $\boldsymbol{\sigma}_\mathbf{x}^2 = \text{Var}[\mathbf{x}]$, the covariance matrix $\boldsymbol{\Sigma}_{\mathbf{x},\mathbf{y}} = \mathbf{Cov}[\mathbf{x}, \mathbf{y}]$ and we note $\boldsymbol{\Sigma}_\mathbf{x} = \mathbf{Cov}[\mathbf{x}, \mathbf{x}]$. $\mathcal{CN}(\mu_x, \sigma_x^2)$ denotes the circularly-symmetric complex Gaussian distribution of mean μ_x and variance σ_x^2, and $\mathcal{B}(p)$ denotes the Bernoulli distribution with a success probability of $0 \leq p \leq 1$.

II. System Model

A. Transmission Over a Multipath Channel

We consider a single carrier transmission using a bit-interleaved coded modulation (BICM) scheme. Let $\mathbf{b} \in \mathbb{F}_2^{K_b}$ be a binary information packet of length K_b bits. A channel encoder maps $\mathbf{b}$ into a codeword $\mathbf{c} \in \mathbb{F}_2^{K_c}$, with a code rate $R_c = K_b/K_c$, which is then interleaved to give a data block $\mathbf{d} \in \mathbb{F}_2^{K_c}$. A memoryless mapping φ associates $\mathbf{d}$ to the symbol block of length K, denoted $\mathbf{x} \in \mathcal{X}^K$, where the constellation $\mathcal{X} \subset \mathbb{C}$ has M elements. The q-word associated to a symbol is denoted $\mathbf{d}_k = [\mathbf{d}]_{qk:q(k+1)-1}$, and $\varphi_j^{-1}(x_k)$ and $d_{k,j}$ denote the value of the j^{th} bit labelling the k^{th} symbol x_k, i.e. d_{kq+j}. We assume the constellation has zero mean, and has an average symbol power of σ_x^2, with equiprobable symbols.

For the sake of clarity, only the single user, single input-single output T-spaced (symbol spaced) equalization problem is considered. The channel is modelled at the base-band as an equivalent L-tap linear time-varying filter $\mathbf{h}[k] = [h_{k,L-1}, h_{k,L-2} \dots h_{k,0}]$, k being the time index, and where pulse shaping and transceiver filters are accounted for.

The signal going through the channel is then affected by thermal noise w_k at the receiver side, and assuming a perfect channel state information, ideal time and frequency synchronization and the absence of inter-block interference (IBI), the base-band received samples are given by:

$$y_k = \sum\nolimits_{l=0}^{L-1} h_{k,l} x_{k-l} + w_k, \tag{1}$$

where $k = 0, 1, \dots, K+L-2$, and x_k, $k < 0$ and $k > K$ are set to 0. These assumptions can be satisfactorily approached in practice with the use of a unique-word signalling scheme, among other options, to jointly enable channel estimation and the IBI removal. The noise is modelled as $w_k \sim \mathcal{CN}(0, \sigma_w^2)$, i.e. its real and imaginary parts are real independent zero mean Gaussian random processes with $\sigma_w^2/2$ variance each. The transmission can be rewritten as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{2}$$

with $\mathbf{y} = [y_0, \dots, y_{K+L-2}]^T$, $\mathbf{w} = [w_0, \dots, w_{K+L-2}]^T$, $\mathbf{x} = [x_{-L+1}, \dots, x_{K+L-2}]^T$ and $\mathbf{H}$ is the $(K+L-1) \times (K+2L-2)$ matrix whose k^{th} row is $[\mathbf{0}_{1,k-1}, \mathbf{h}[k], \mathbf{0}_{1,K+L-1-k}]$, $k = 1, \dots, K+L-1$.

B. On MMSE FIR Equalization

FIR structures can be modelled by windowed processes; applying a sliding window $[-N_p, N_d]$ on the observation vector $\mathbf{y}$, we define $\mathbf{y}_k = [y_{k-N_p}, \dots, y_{k+N_d}]^T$. N_p and N_d are respectively the number of pre-cursor and post-cursor samples, and we denote $N \triangleq N_p + N_d + 1$, and $N_p' \triangleq N_p + L - 1$ to simplify notations. Then, using the same window on $\mathbf{w}$, and $[-N_p', N_d]$ on $\mathbf{x}$, the channel model becomes

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{w}_k, \tag{3}$$

with $\mathbf{H}_k = [\mathbf{H}]_{k-N_p\,:\,k+N_d,\ k-N_p'\,:\,k+N_d}$, for $k = 0, \dots, K-1$.

Below, a generic structure of an unbiased MMSE FIR receiver is given for comparing different structures and their dynamics in the remainder of the paper. Prior estimates on $\mathbf{x}$ with means $\bar{\mathbf{x}}_k^{\text{fir}} \triangleq [\bar{x}_{k-N_p'}^{\text{fir}}, \dots, \bar{x}_{k+N_d}^{\text{fir}}]$ and variances $\bar{\mathbf{v}}_k^{\text{fir}} \triangleq [\bar{v}_{k-N_p'}^{\text{fir}}, \dots, \bar{v}_{k+N_d}^{\text{fir}}]$ are used for interference cancellation. Then denoting its output estimate on x_k as x_k^e, and the variance

of the residual interference and noise as v_k^e, with

$$\begin{array}{l} x_k^e = \mathbf{f}_k^{\mathbf{fir}H}\mathbf{y}_k + g_k^{\text{fir}} \\ v_k^e = 1/\xi_k^{\text{fir}} - \bar{v}_k^{\text{fir}}, \end{array} \quad \begin{cases} \mathbf{f}_k^{\mathbf{fir}} \triangleq \mathbf{\Sigma}_k^{\mathbf{fir}-1}\mathbf{h}_k/\xi_k^{\text{fir}}, \\ g_k^{\text{fir}} \triangleq \bar{x}_k^{\text{fir}} - \mathbf{f}_k^{\mathbf{fir}H}\mathbf{H}_k\bar{\mathbf{x}}_k^{\mathbf{fir}}, \\ \xi_k^{\text{fir}} \triangleq \mathbf{h}_k^H\mathbf{\Sigma}_k^{\mathbf{fir}-1}\mathbf{h}_k, \end{cases} \tag{4}$$

where $\mathbf{\Sigma}_k^{\mathbf{fir}} \triangleq k_w\sigma_w^2\mathbf{I}_N + \mathbf{H}_k\bar{\mathbf{V}}_k^{\mathbf{fir}}\mathbf{H}_k^H$, $\bar{\mathbf{V}}_k^{\mathbf{fir}} \triangleq \mathbf{diag}(\bar{\mathbf{v}}_k^{\mathbf{fir}})$, $\mathbf{h}_k \triangleq \mathbf{H}_k\mathbf{e}_k$ and $k_w = 1/2$, when signals with one real degree of freedom are used (e.g. $\mathcal{X}$ is BPSK), and otherwise $k_w = 1$ [17]. A proof of these relationships is in Appendix.

Note that $\bar{\mathbf{x}}_k^{\mathbf{fir}}$ and $\bar{\mathbf{v}}_k^{\mathbf{fir}}$ completely characterize such receivers. When $\bar{\mathbf{x}}_{k'}^{\mathbf{fir}}$ and $\bar{\mathbf{v}}_{k'}^{\mathbf{fir}}$ are independent of $x_k^e, v_k^e, \forall k', k$, we call this receiver a LE-IC, and when $\bar{x}_{k'}^{\text{fir}}$ and $\bar{v}_{k'}^{\text{fir}}$ are dependent on x_k^e, v_k^e, $\forall k' < k$, we refer to it as a DFE-IC.

III. Receiver Design With Expectation Propagation

This section focuses on the design of a FIR receiver that approximates the posterior probability distribution on x_k using an EP-based message passing on the system factor graph.

A. Factor Graph Model for FIR Receivers

The optimal joint MAP receiver satisfies the MAP criterion $\hat{\mathbf{b}} = \max_{\mathbf{b}} p(\mathbf{b}|\mathbf{y})$, where, assuming i.i.d. information bits, the posterior PDF can be factorized as follows

$$p(\mathbf{b}|\mathbf{y}) = p(\mathbf{b},\mathbf{d},\mathbf{x}|\mathbf{y}) \propto \underbrace{p(\mathbf{y}|\mathbf{x})}_{\text{channel}}\underbrace{p(\mathbf{x}|\mathbf{d})}_{\text{mapping}}\underbrace{p(\mathbf{d}|\mathbf{b})}_{\text{encoding}}. \tag{5}$$

This density can be further factorized by using:
- the memoryless mapping: $p(\mathbf{x}|\mathbf{d}) = \prod_{k=0}^{K-1} p(x_k|\mathbf{d}_k)$,
- the independence assumption in BICM encoding: $p(\mathbf{d}|\mathbf{b}) = \prod_{k=0}^{K-1}\prod_{j=0}^{q-1} p(d_{k,j})$,

where $p(d_{k,j}) \triangleq p(d_{k,j}|\mathbf{b})$ is a probability mass function (PMF) which is seen as a Bernoulli-distributed prior constraint provided by the decoder, from the receiver's point of view.

The "channel" factor in (5) creates constraints between the whole block of received baseband samples and the transmitted symbols, however to derive a reduced complexity FIR receiver which estimates x_k and $\mathbf{d}_k$, the windowed model in (3) is needed. The FIR approximation posterior is

$$\begin{aligned} &p\left(\bar{\mathbf{d}}_k, \mathbf{x}_k|\mathbf{y}_k\right) \\ &\quad\propto \prod_{k'=k-N_p'}^{k+N_d} p(\mathbf{y}_k|\mathbf{x}_k)p(x_{k'}|\mathbf{d}_{k'})\prod_{j=0}^{q-1} p(d_{k',j}), \end{aligned} \tag{6}$$

where $\bar{\mathbf{d}}_k = \mathbf{d}_{k-N_p-L+1:k+N_d}$. Note that working with $p\left(\bar{\mathbf{d}}_k, \mathbf{x}_k|\mathbf{y}\right) \approx p\left(\bar{\mathbf{d}}_k, \mathbf{x}_k|\mathbf{y}_k\right)$ is not the only option for estimating x_k. Indeed x_k can be estimated through inference on $\mathbf{x}_{k'}$, with $k' = k - N_d, \ldots, k + N_p'$, but by selecting $\mathbf{x}_k$, this option is indirectly translated to the choice of window parameters, which is a common aspect of FIR equalizers.

A message-passing based decoding algorithm iteratively estimates the variable nodes (VN) x_k and $d_{k,j}$ by using constraints imposed by factor nodes (FN). Factor nodes are non proper PDFs for resolving transmission steps. The decoder FN models BICM encoding constraints with

$$f_{\text{DEC}}(d_{k,j}) \triangleq p(d_{k,j}), \tag{7}$$

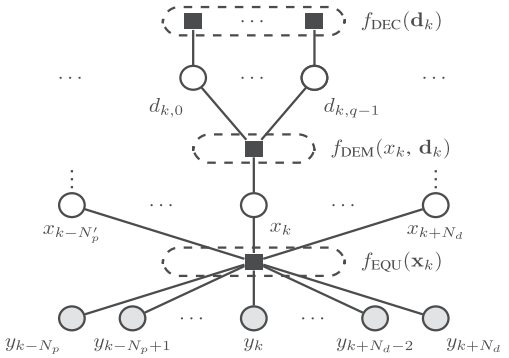

Fig. 1. Factor graph for the posterior PDF (6) on x_k and $\mathbf{d}_k$.

and the demapper FN incorporates mapping constraints with

$$f_{\text{DEM}}(x_k, \mathbf{d}_k) \triangleq p(x_k|\mathbf{d}_k) = \prod_{j=0}^{q-1}\delta(d_{k,j} - \varphi_j^{-1}(x_k)), \tag{8}$$

where δ is the Dirac delta function. The multipath channel constraints are modelled within the equalization factor node

$$f_{\text{EQU}}(\mathbf{x}_k) \triangleq p(\mathbf{y}_k|\mathbf{x}_k) \propto e^{-\mathbf{y}_k^H\mathbf{y}_k/\sigma_w^2 + 2\mathcal{R}(\mathbf{y}_k^H\mathbf{H}_k\mathbf{x}_k)/\sigma_w^2}, \tag{9}$$

where the dependence on $\mathbf{y}_k$ is omitted, as observations are fixed during the message-passing procedure. Using these notations, the posterior (6) gives the factor graph shown in Fig. 1.

B. Expectation Propagation Message Passing Framework

EP-based message passing algorithm is an extension of loopy belief propagation, where VNs are assumed to lie in the exponential distribution family [38]. Consequently, the exchanged messages are depicted by tractable distributions, and they allow iterative computation of a fully-factorized approximation for cumbersome posterior PDFs such as $p(\bar{\mathbf{d}}_k, \mathbf{x}_k|\mathbf{y}_k)$. Updates at a FN F connected to variable nodes $\mathbf{v}$ are as follows. Messages exchanged between a VN v_i, the i^{th} component of $\mathbf{v}$, and factor node F are

$$m_{v\to\text{F}}(v_i) \triangleq \prod_{\text{G}\neq F} m_{\text{G}\to v}(v_i), \tag{10}$$

$$m_{\text{F}\to v}(v_i) \triangleq \text{proj}_{\mathcal{Q}_{v_i}}[q_{\text{F}}(v_i)]/m_{v\to\text{F}}(v_i), \tag{11}$$

where $\text{proj}_{\mathcal{Q}_{v_i}}$ is the Kullback-Leibler projection towards the probability distribution $\mathcal{Q}_{v_i}$ of VN v_i. The posterior $q_{\text{F}}(v_i)$ is an approximation of the marginal of the true posterior $p(\mathbf{v})$ on v_i, obtained by combining the true factor on FN F with messages from the neighbouring VNs

$$q_{\text{F}}(v_i) \triangleq \int_{\mathbf{v}^{\setminus i}} f_{\text{F}}(\mathbf{v})\prod_{v_j} m_{v\to\text{F}}(v_j)\mathbf{d}\mathbf{v}^{\setminus i}, \tag{12}$$

where $\mathbf{v}^{\setminus i}$ are VNs without v_i [38]. The projection operation for exponential families is equivalent to *moment matching*, which simplifies the computation of messages [30], [38].

In this paper symbol VNs are assumed to lie in the family of multivariate circularly symmetric Gaussians with diagonal covariance matrices, making the approximate distributions fully factorized to independent Gaussians. Hence, a message on x_k will be defined by a mean and a variance. The VNs $d_{k,j}$ are considered to follow Bernoulli distributions (which is

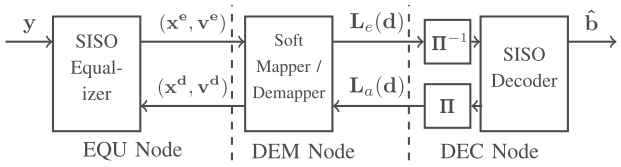

Fig. 2. Factor nodes shown as an iterative BICM receiver.

included in the exponential family), and their messages can be described by bit log-likelihood ratios (LLR).

This formalism is very generic and allows the derivation of many receiver structures. It has been used to derive a MIMO detector in [32], and a Kalman smoother in [34]. However EP receivers can also be derived without a message passing formalism, as recently shown for the block [33] or FIR [36] equalizers. To the authors' knowledge, message-passing formalism was not previously used for FIR design, and it is favoured in this paper because of the available scheduling options it allows to clearly identify.

C. Derivation of Exchanged Messages

This section details the EP-based message passing algorithm's application to the considered factor graph. First, exchanged messages are defined, and then their characterizing parameters are explicitly computed. See Fig. 2 for a conventional view of the receiver with these quantities.

The messages arriving on the VN x_k are Gaussians with

$$m_{\text{EQU}\to x}(x_k) \propto \mathcal{CN}\left(x_k^e, v_k^e\right), \tag{13}$$
$$m_{\text{DEM}\to x}(x_k) \propto \mathcal{CN}\left(x_k^d, v_k^d\right), \tag{14}$$

whereas messages arriving on the VN $d_{k,j}$ are Bernoullis

$$m_{\text{DEC}\to d}(d_{k,j}) \propto \mathcal{B}\left(p_d^a\right), \quad m_{\text{DEM}\to d}(d_{k,j}) \propto \mathcal{B}\left(p_d^e\right). \tag{15}$$

During the message passing procedure, the characteristic parameters of these distributions are updated following a selected schedule. For Bernoulli distributions, it is rather preferable to work with bit LLRs, rather than the success probability p_d:

$$L(d_j) \triangleq \ln \frac{\mathbb{P}[d_j = 0]}{\mathbb{P}[d_j = 1]} = \ln \frac{1 - p_d}{p_d}. \tag{16}$$

We use $L_a(\cdot)$, $L_e(\cdot)$ and $L(\cdot)$ operators to denote respectively a priori, extrinsic and a posteriori LLRs. When applied to $d_{k,j}$, this vocabulary represents the receiver's perspective, i.e. $L_a(d_{k,j})$, $L_e(d_{k,j})$ respectively characterize $m_{\text{DEC}\to d}(d_{k,j})$ and $m_{\text{DEM}\to d}(d_{k,j})$.

Finally, considering the factor graph shown on Fig. 1, all variable nodes are only connected to a pair of distinct factor nodes. Consequently, using eq. (10), $m_{v\to \text{F}}(v_i) = m_{\text{G}\to v}(v_i)$, for all VN v_i, and FN $\text{F}, \text{G}, \text{F} \neq \text{G}$ they are connected to.

1) Messages From DEC to DEM: In this paper, we assume DEC is a SISO decoder providing prior information $L_a(\mathbf{d})$ to DEM, whenever it receives extrinsic information $L_e(\mathbf{d})$ by DEM.

The demapper uses these prior LLRs, along with the DEM FN (8) to compute a prior PMF on $x_k = \alpha$, $\forall \alpha \in \mathcal{X}$ with

$$\mathcal{P}_k(\alpha) \propto \prod_{j=0}^{q-1} e^{-\varphi_j^{-1}(\alpha) L_a(d_{k,j})}. \tag{17}$$

This is a categorical PMF corresponding to the marginal of $f_{\text{DEM}}(x_k, \mathbf{d}_k) m_{d\to\text{DEC}}(\mathbf{d}_k)$ on x_k [32], used hereafter to compute approximate marginals $q_{\text{DEM}}(x_k)$ and $q_{\text{DEM}}(d_{k,j})$.

2) Messages From DEM to EQU: The demapper computes an approximate posterior on the VN x_k using eq. (12) with

$$q_{\text{DEM}}(x_k) = \sum\nolimits_{\mathbf{d}_k} f_{\text{DEM}}(x_k, \mathbf{d}_k) m_{x\to\text{DEM}}(x_k) \times \prod\nolimits_{j=0}^{q-1} m_{d\to\text{DEM}}(d_{k,j}). \tag{18}$$

This is a posterior categorical PMF on the elements x_k of $\mathcal{X}$, which can be computed using eqs. (13) and (17), which will be denoted as

$$\mathcal{D}_k(\alpha) \propto \exp\left(-k_w |\alpha - x_k^e|^2 / v_k^e\right) \mathcal{P}_k(\alpha), \quad \forall \alpha \in \mathcal{X}. \tag{19}$$

For computing messages towards EQU, the posterior PMF is projected into $\mathcal{CN}$ through moment matching. The mean and the variance of $\mathcal{D}_k$ are

$$\mu_k^d \triangleq \mathbb{E}_{\mathcal{D}_k}[x_k] = \sum\nolimits_{\alpha\in\mathcal{X}} \alpha \mathcal{D}_k(\alpha),$$
$$\gamma_k^d \triangleq \text{Var}_{\mathcal{D}_k}[x_k] = \sum\nolimits_{\alpha\in\mathcal{X}} |\alpha|^2 \mathcal{D}_k(\alpha) - |\mu_k^d|^2. \tag{20}$$

When $m_{x\to\text{DEM}}(x_k) \propto 1$, i.e. when there is no information from the EQU node (equivalent to $x_k^e = 0$ and $v_k^e = +\infty$), $\mathcal{D}_k = \mathcal{P}_k$, and we denote the prior mean and variances as

$$x_k^p \triangleq \mathbb{E}_{\mathcal{P}_k}[x_k], \quad v_k^p \triangleq \text{Var}_{\mathcal{P}_k}[x_k]. \tag{21}$$

Note that these values are used as soft feedback in conventional turbo equalization.

Then in order to calculate $m_{\text{DEM}\to x}(x_k)$ as in (11), a Gaussian division [30] is implemented

$$x_k^* = \frac{\mu_k^d v_k^e - x_k^e \gamma_k^d}{v_k^e - \gamma_k^d}, \quad \text{and,} \quad v_k^* = \frac{v_k^e \gamma_k^d}{v_k^e - \gamma_k^d}. \tag{22}$$

This is the major novelty in using EP: the computation of an extrinsic feedback from the demapper to the equalizer. Attempting this with categorical distributions, as in BP, would completely remove $m_{x\to\text{DEM}}(x_k)$, and the extrinsic "feedback" to EQU would simply be the prior PMF $\mathcal{P}_k$ [32], which would yield a receiver equivalent to LE-IC [19].

EP message passing algorithm consists in minimizing global divergence through iterative minimization of simpler local divergences. Thus, it might lock on undesirable fixed points, and a damping heuristic, as recommended in [38, eq. (17)], is used to improve accuracy

$$v_k^{d(\text{next})} = \left[(1-\beta)/v_k^* + \beta/\bar{v}_k^{d(\text{prev})}\right]^{-1},$$
$$x_k^{d(\text{next})} = v_k^{d(\text{next})} \left[(1-\beta)\frac{x_k^*}{v_k^*} + \beta \frac{x_k^{d(\text{prev})}}{v_k^{d(\text{prev})}}\right], \tag{23}$$

where $0 \leq \beta \leq 1$ configures the damping, and its effectiveness has been verified in [36].

3) Messages From EQU to DEM: The equalizer computes an approximate posterior on the VN x_k using eq. (12) with

$$q_{\text{EQU}}(x_k) = \int_{\mathbf{x}_k^{\setminus k}} f_{\text{EQU}}(\mathbf{x}_k) \prod\nolimits_{k'=k-N_p'}^{k+N_d} m_{x\to\text{EQU}}(x_{k'}) \mathbf{d}\mathbf{x}_k^{\setminus k}. \tag{24}$$

The integrand of the equation above is a multivariate Gaussian distribution $\mathcal{CN}(\boldsymbol{\mu}^{\mathbf{e}}, \boldsymbol{\Gamma}^{\mathbf{e}})$, hence, using eq. (9), its covariance and mean satisfy

$$\begin{aligned}\boldsymbol{\Gamma}_k^{\mathbf{e}} &= (\mathbf{V}_k^{\mathbf{d}\,-1} + \sigma_w^{-2}\mathbf{H}_k^H\mathbf{H}_k)^{-1},\\ \boldsymbol{\mu}_k^{\mathbf{e}} &= \boldsymbol{\Gamma}^{\mathbf{e}}(\mathbf{V}_k^{\mathbf{d}\,-1}\mathbf{x}_k^{\mathbf{d}} + \sigma_w^{-2}\mathbf{H}_k^H\mathbf{y}_k),\end{aligned} \tag{25}$$

where $\mathbf{V}_k^{\mathbf{d}} = \mathbf{diag}(\mathbf{v}_k^{\mathbf{d}})$, with $\mathbf{v}_k^{\mathbf{d}} = [v_{k-N'_p}^d, \ldots, v_{k+N_d}^d]$, and $\mathbf{x}_k^{\mathbf{d}} = [x_{k-N'_p}^d, \ldots, x_{k+N_d}^d]$. Using some matrix algebra, and Woodbury's identity on $\boldsymbol{\Gamma}^{\mathbf{e}}$, the mean μ_k^e and the variance γ_k^e of the marginalized PDF $q_{\text{EQU}}(x_k)$ are given by

$$\begin{aligned}\gamma_k^e &= \mathbf{e}_k^H\boldsymbol{\Gamma}_k^{\mathbf{e}}\mathbf{e}_k = v_k^d(1 - v_k^d\mathbf{h}_k^H\boldsymbol{\Sigma}_k^{\mathbf{d}\,-1}\mathbf{h}_k),\\ \mu_k^e &= \mathbf{e}_k^H\boldsymbol{\mu}_k^{\mathbf{e}} = x_k^d + v_k^d\mathbf{h}_k^H\boldsymbol{\Sigma}_k^{\mathbf{d}\,-1}(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k^{\mathbf{d}}),\end{aligned} \tag{26}$$

with $\boldsymbol{\Sigma}_k^{\mathbf{d}} = k_w\sigma_w^2\mathbf{I}_N + \mathbf{H}_k\mathbf{V}_k^{\mathbf{d}}\mathbf{H}_k^H$. Message to the demapper is then extracted with the Gaussian density division in eq. (11)

$$v_k^e = \frac{\gamma_k^e v_k^d}{v_k^d - \gamma_k^e}, \quad \text{and,} \quad x_k^e = \frac{v_k^d\mu_k^e - \gamma_k^e x_k^d}{v_k^d - \gamma_k^e}. \tag{27}$$

Developing these yields a FIR expression as in (4) with $\bar{\mathbf{x}}_k^{\mathbf{ep}} \triangleq [x_{k-N'_p}^d, \ldots, x_{k+N_d}^d]$ and $\bar{\mathbf{v}}_k^{\mathbf{ep}} \triangleq [v_{k-N'_p}^d, \ldots, v_{k+N_d}^d]$ for IC.

4) Messages From DEM to DEC: The demapper computes an approximate posterior on the VN $d_{k,j}$ using eq. (12) with

$$\begin{aligned}q_{\text{DEM}}(\mathbf{d}_k) = \sum\nolimits_{x_k\in\mathcal{X}} f_{\text{DEM}}(x_k, \mathbf{d}_k)m_{x\to\text{DEM}}(x_k)\\ \times \prod\nolimits_{j=0}^{q-1} m_{d\to\text{DEM}}(d_{k,j}).\end{aligned} \tag{28}$$

As bit LLRs are used to represent messages to DEC, this distribution is marginalized on $d_{k,0}, \ldots, d_{k,q-1}$ [32], and the division in eq. (11) is directly carried out with LLRs

$$L_e(d_{k,j}) = \ln\sum_{\alpha\in\mathcal{X}_j^0}\mathcal{D}_k(\alpha) - \ln\sum_{\alpha\in\mathcal{X}_j^1}\mathcal{D}_k(\alpha) - L_a(d_{k,j}), \tag{29}$$

with $\mathcal{X}_j^p = \{\alpha\in\mathcal{X} : \varphi_j^{-1}(x) = p\}$ where $p\in\mathbb{F}_2$.

D. Proposed Self-Iterated DFE-IC EP Receiver

A factor graph (sec. III-A) and messages exchanged over it (sec. III-C) are necessary to derive a receiver algorithm, but may be insufficient when considering a graph with cycles. Indeed, specifying a scheduling for the update of VNs and FNs is also required.

In this paper, a serial scheduling across variable nodes x_k is considered. In detail, when EQU updates a VN x_k, factor node DEM is immediately activated in order to provide its own extrinsic estimation of x_k, jointly using prior information from the decoder and the equalizer's extrinsic output. This results in a DFE-IC structure, using a novel kind of soft feedback, unlike any hard or soft APP feedback previously used in the literature [19]–[21], [23]–[27]. Moreover, when detection across the whole block is completed, this serial scheduling can be repeated by keeping the previously updated DEM messages, yielding a *self-iterated DFE-IC EP* structure.

To clarify the dynamics of the proposed receiver, $\tau = 0, \ldots, \mathcal{T}$ denotes turbo iterations (TI), i.e. exchanges between the DEM and DEC factor nodes. Each TI consists of $s = 0, \ldots, \mathcal{S}_\tau$ self-iterations (SI) (may vary with τ), i.e. exchanges between EQU and DEM factor nodes, which *sequentially* updates the whole block $\mathbf{x}$. In the following, EQU $\leftrightarrow$ DEM messages derived previously are appended a superscript (s), but τ is omitted for readability.

Algorithm 1 Proposed *Self-Iterated DFE-IC EP* Receiver

Input $\mathbf{y}$, $\mathbf{H}$, σ_w^2

Initialize decoder with $L_a^{(0)}(\mathbf{d}_k) = 0, \forall k$.
for $\tau = 0$ to $\mathcal{T}$ **do**
  $\forall k = 0, \ldots, K-1$, use $L_a^{(\tau)}(\mathbf{d})$ to compute $\mathcal{P}_k^{(\tau)}$ with (17), and set $(x_k^{d(0)}, v_k^{d(0)}) \leftarrow (x_k^p, v_k^p)$ using (21).
  for $s = 0$ to $\mathcal{S}_\tau$ **do**
    for $k = 0$ to $K-1$ **do**
      Equalize using (27) and get $(x_k^{e(s)}, v_k^{e(s)})$.
      Use (19)-(20) to update $\mathcal{D}_k^{(s+1)}$, and generate EP feedback $(x_k^{d(s+1)}, v_k^{d(s+1)})$ with (22)-(23).
      If $v_k^{d(s+1)} \le 0$, then $(x_k^{d(s+1)}, v_k^{d(s+1)}) \leftarrow (\mu_k^d, \gamma_k^d)$ and store k in the set $\mathcal{I}_{\text{err}}^{(s)}$.
    end for
    $\forall k \in \mathcal{I}_{\text{err}}^{(s)}$, $(x_k^{d(s+1)}, v_k^{d(s+1)}) \leftarrow (x_k^{d(s)}, v_k^{d(s)})$.
  end for
  Compute $L_e^{(\tau)}(\mathbf{d}_k)$ using $\mathcal{D}_k^{(\tau,\mathcal{S}_\tau)}$ with (29), $\forall k$, and provide them to the decoder, to obtain $L_a^{(\tau+1)}(\mathbf{d}_k), \forall k$.
end for

The proposed scheduling, given in Algorithm 1, generates an EP FIR receiver which uses the following means and variances for interference cancellation

$$\begin{aligned}\bar{\mathbf{x}}_k^{\textbf{dfe-ep}(s)} &\triangleq [x_{k-N'_p}^{d(s+1)}, \ldots, x_{k-1}^{d(s+1)}, x_k^{d(s)}, \ldots, x_{k+N_d}^{d(s)}]^T,\\ \bar{\mathbf{v}}_k^{\textbf{dfe-ep}(s)} &\triangleq [v_{k-N'_p}^{d(s+1)}, \ldots, v_{k-1}^{d(s+1)}, v_k^{d(s)}, \ldots, v_{k+N_d}^{d(s)}]^T,\end{aligned} \tag{30}$$

for $k = 0, \ldots, K-1$. This layout shows that this structure indeed follows a time-varying DFE-IC evolution, with anti-causal symbols using demapper's output from the previous self-iteration, and causal symbols using current EP feedback from the demapper. The extrinsic feedback from the demapper is obtained by using jointly the prior information from the previous TI, and the past equalizer outputs of the current and previous self iterations (see (19)-(23)). The Algorithm 1 also incorporates a mechanism to deal with EP-based feedback's infamous negative variances [32], [33], with the set $\mathcal{I}_{\text{err}}^{(s)}$ which stores their indexes. These values are replaced with APP-based variances in the current SI, and then replaced again with their previous values for the next SI.

Although equation (4) is useful for FIR analysis, causal and anti-causal feedback of DFE-IC should be separated in practice. Using

$$\mathbf{E}^{\mathbf{c}} = [\mathbf{I}_{N'_p}, \mathbf{0}_{N'_p, N_d+1}], \quad \mathbf{E}^{\mathbf{a}} = [\mathbf{0}_{N_d+1, N'_p}, \mathbf{I}_{N_d+1}], \tag{31}$$

we define $\mathbf{H}_k^{\mathbf{c}} = \mathbf{H}_k\mathbf{E}^{\mathbf{c}T}$ and $\mathbf{H}_k^{\mathbf{a}} = \mathbf{H}_k\mathbf{E}^{\mathbf{a}T}$, to respectively operate on $\bar{\mathbf{x}}_k^{\mathbf{c}(s)} = \mathbf{E}^{\mathbf{c}}\bar{\mathbf{x}}_k^{\textbf{dfe-ep}(s)}$, and $\bar{\mathbf{x}}_k^{\mathbf{a}(s)} = \mathbf{E}^{\mathbf{a}}\bar{\mathbf{x}}_k^{\textbf{dfe-ep}(s)}$, as a generalized interference cancellation scheme. The *SI DFE-IC*

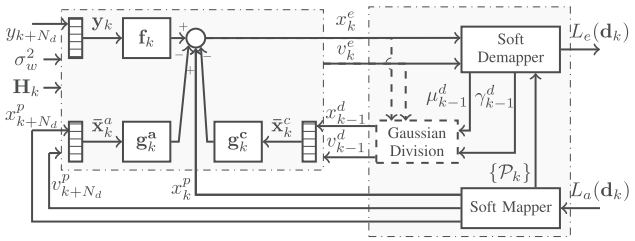

Fig. 3. TV DFE-IC EP (dashed) / APP (no dashed) structure.

EP of (30), is rewritten as:

$$x_k^{e(s)} = \bar{x}_k^{a(s)} + \mathbf{f}_k^{(s)H}\mathbf{y}_k - \mathbf{g}_k^{\mathbf{c}(s)H}\bar{\mathbf{x}}_k^{\mathbf{c}(s)} - \mathbf{g}_k^{\mathbf{a}(s)H}\bar{\mathbf{x}}_k^{\mathbf{a}(s)},$$
$$v_k^{e(s)} = 1/\xi_k^{\text{dfe-ep}(s)} - \bar{v}_k^{a(s)}, \tag{32}$$

with $\mathbf{f}_k^{(s)} = \boldsymbol{\Sigma}_k^{\mathbf{dfe\text{-}ep}(s)-1}\mathbf{h}_k/\xi_k^{\text{dfe-ep}(s)}$, $\mathbf{g}_k^{\mathbf{c}(s)} = \mathbf{H}_k^{\mathbf{c}\,H}\mathbf{f}_k$, and $\mathbf{g}_k^{\mathbf{a}(s)} = \mathbf{H}_k^{\mathbf{a}\,H}\mathbf{f}_k^{(s)}$. When $\mathcal{S}_\tau = 0$, the proposed receiver is a strict TV DFE-IC EP, with $\cdot_k^{d(s+1)} = \cdot_k^d$ and $\cdot_k^{d(s)} = \cdot_k^p$, this case is shown on Fig. 3 with the dashed module.

In conclusion, we have applied message passing framework of EP for equalization, using *sliding window* observations. This results in a novel message computation given by (22)-(23) and (27), unlike blockwise messages in [32] and [33]. Moreover, by using an hybrid serial/parallel schedule, our structure operates as a self-iterated DFE-IC, unlike the self-iterated LE-IC scheme concurrently developed in [36]. In the following, a matrix inversion strategy reducing the computational complexity difference between DFE-IC and LE-IC is introduced.

IV. Matrix Inversion for Time-Varying Sliding Window Turbo Equalizers

A. Shortcomings of Existing Approaches

Time-varying FIR as in (4) have excessive computational costs due to symbol-wise filter updates, requiring recursive matrix inversion methods. This section overviews the problem of computing $\mathbf{f}_k = \boldsymbol{\Sigma}_k^{-1}\mathbf{h}_k$, for $k = 0, \ldots, K-1$ efficiently.

In [10], Tüchler *et al.* propose for LE-IC, a recursive matrix inversion algorithm, based on common submatrices between successive inverses. The procedure requires computing an initial inverse (Gauss-Jordan inversion) with a complexity order[2] of $4N^3/3$, but further recursions' complexity is $2N^2$.

Practical implementations avoid inversion by solving the system $\boldsymbol{\Sigma}_k\mathbf{f}_k = \mathbf{h}_k$ for $\mathbf{f}_k$ with triangular factorizations [39], using forward/backward substitutions. This approach is even more advantageous in equalization where the system is sparse.

In this paper, we propose a novel recursive inversion strategy for LE-IC and DFE-IC, based on an initial Cholesky decomposition, and followed by sparse rank-1 updates/downdates of the factors for following inversions. Unlike [39], our algorithm is able to deal with channel matrices evolving in time, making it more efficient for turbo TV FIR. For LE-IC the complexity order is of N^2, hence roughly 50% less complex than [10].

[2]"Order" means asymptotic expansion as $N \to +\infty$, assuming $N \propto 3L$, i.e. sliding window operating on $4L$ symbols.

B. Cholesky Factor Update for MMSE LE-IC

We consider a LE-IC with priors variances $\bar{v}_k$, let $\mathbf{L}_{k-1}$ be the lower triangular Cholesky decomposition of the covariance matrix $\boldsymbol{\Sigma}_{k-1}$, i.e. $\mathbf{L}_{k-1}\mathbf{L}_{k-1}^H = \boldsymbol{\Sigma}_{k-1}$. Algorithm 2 uses $\mathbf{L}_{k-1}$ and the latest values $\bar{v}_{k+N_d}$ and $\mathbf{e}_{k+N_d}^H\mathbf{H}_k$ (new row with $\mathbf{h}[k+N_d]$) appended to the sliding window, to compute $\mathbf{L}_k$. Impact of latest generated value is appended to the decomposition using $\mathbf{l}_{\mathbf{12}}^H$ and l_{22}, then past data is removed. The resulting updated Cholesky decomposition is a rank-1 update [40] of $\mathbf{L}_{\mathbf{22}}$, defined within algorithm 2.

Algorithm 2 Cholesky Update Algorithm for LE-IC.

Input $\mathbf{L}_{k-1}, \sigma_w^2, \bar{v}_{k+N_d}, \mathbf{H}_{k-1}, \mathbf{H}_k, \bar{\mathbf{V}}_{k-1}$
Output $\mathbf{L}_k$
{*Add a row and a column*}
$[\mathbf{h}_{\mathbf{1}k}, h_{2k}] \leftarrow [0, \mathbf{e}_{k+N_d}^H\mathbf{H}_k]$
$\mathbf{w} \leftarrow \mathbf{H}_{k-1}\bar{\mathbf{V}}_{k-1}\mathbf{h}_{\mathbf{1}k}$
$\mathbf{l}_{\mathbf{12}} \leftarrow \mathbf{L}_{k-1}^{-1}\mathbf{w}$
$l_{22} \leftarrow \sqrt{\mathbf{h}_{\mathbf{1}k}^H\bar{\mathbf{V}}_{k-1}\mathbf{h}_{\mathbf{1}k} + \bar{v}_{k+N_d}|h_{2k}|^2 - \mathbf{l}_{\mathbf{12}}{}^H\mathbf{l}_{\mathbf{12}} + \sigma_w^2}$
{*Build augmented matrix and remove row & column*}
$\begin{bmatrix} \times & \mathbf{0}_{1,N} \\ \mathbf{l}_{\mathbf{21}} & \mathbf{L}_{\mathbf{22}} \end{bmatrix} \leftarrow \begin{bmatrix} \mathbf{L}_{k-1} & \mathbf{0}_{N,1} \\ \mathbf{l}_{\mathbf{12}}{}^H & l_{22} \end{bmatrix}$
{*Rank-1 update* $\mathbf{L}_k\mathbf{L}_k^H = \mathbf{L}_{\mathbf{22}}\mathbf{L}_{\mathbf{22}}{}^H + \mathbf{l}_{\mathbf{21}}\mathbf{l}_{\mathbf{21}}{}^H$}
for $l = 1$ to N **do**
  $r \leftarrow \sqrt{[\mathbf{L}_{\mathbf{22}}]_{l,l}^2 + |[\mathbf{l}_{21}]_l|^2}, c \leftarrow \frac{[\mathbf{L}_{\mathbf{22}}]_{l,l}}{r}, s \leftarrow \frac{[\mathbf{l}_{21}]_l^*}{r}$
  $[\mathbf{L}_{\mathbf{22}}]_{l:N,l} \leftarrow c[\mathbf{L}_{\mathbf{22}}]_{l:N,l} + s[\mathbf{l}_{21}]_{l:N}$
  $[\mathbf{l}_{21}]_{l:N} \leftarrow c[\mathbf{l}_{21}]_{l:N} - s^*[\mathbf{L}_{\mathbf{22}}]_{l:N,l}$
end for
$\mathbf{L}_k \leftarrow \mathbf{L}_{\mathbf{22}}$

These steps, followed by forward/backward substitutions $\mathbf{f}_k = \mathbf{L}_k^{-H}\mathbf{L}_k^{-1}\mathbf{h}_k$, allow low complexity filter computation.

C. Cholesky Factor Update for MMSE DFE-IC

In the case of DFE-IC, the diagonal of the covariance matrix $\bar{\mathbf{V}}^{\text{tdfe}}$ is composed of two independently sliding parts: one for causal symbols $\bar{v}_k^c$, between symbols $k-N_p'$ and $k-1$, the other for anti-causal $\bar{v}_k^a$, between symbols k and $k+N_d$. The LE-IC update procedure above handles the addition of $\bar{v}_{k+N_d}^a$ and the removal of $\bar{v}_{k-N_p'-1}^c$, but the change in $(k-1)^{\text{th}}$ symbol remains to be updated.

Algorithm 3 gives a such update procedure for DFE-IC, by applying either a rank-1 update or downdate on $\tilde{\mathbf{L}}_k$, the Cholesky factor who has already been updated by algorithm 2, depending on the sign of $\bar{v}_{k-1}^c - \bar{v}_{k-1}^a$. Such updates are carried out using Givens plane rotations [40].

D. Computational Complexity Analysis

The computational complexity of the proposed algorithm is evaluated with the number of required multiply and accumulate units, estimated by the number of real additions and multiplications, amounting to half a floating point operation (0.5 FLOPs) each.

FLOP count ratios between different FIR implementations are plotted in Fig. 4, depending on the channel spread, with

Algorithm 3 Cholesky Update Algorithm for DFE-IC.

Input $\tilde{\mathbf{L}}_k, \bar{v}^a_{k-1}, \bar{v}^c_{k-1}, [\mathbf{H}_k]_{:,-1}$
Output $\mathbf{L}_k$

$\mathbf{w} \leftarrow \sqrt{|\bar{v}^a_{k-1} - \bar{v}^c_{k-1}|}[\mathbf{H}_k]_{:,-1}$
for $l = N_p$ to N **do**
  if $\bar{v}^c_{k-1} < \bar{v}^a_{k-1}$ **then**
    {*Rank-1 downdate* $\mathbf{L}_k\mathbf{L}_k^H = \tilde{\mathbf{L}}_k\tilde{\mathbf{L}}_k^H - \mathbf{w}\mathbf{w}^H$}
    $r \leftarrow \sqrt{[\tilde{\mathbf{L}}_k]^2_{l,l} - |[\mathbf{w}]_l|^2}$, $c \leftarrow \frac{[\tilde{\mathbf{L}}_k]_{l,l}}{r}$, $s \leftarrow \frac{[\mathbf{w}]^*_l}{r}$
    $[\tilde{\mathbf{L}}_k]_{l:N,l} \leftarrow c[\tilde{\mathbf{L}}_k]_{l:N,l} - s[\mathbf{w}]_{l:N}$
  else if $\bar{v}^c_{k-1} > \bar{v}^a_{k-1}$ **then**
    {*Rank-1 update* $\mathbf{L}_k\mathbf{L}_k^H = \tilde{\mathbf{L}}_k\tilde{\mathbf{L}}_k^H + \mathbf{w}\mathbf{w}^H$}
    $r \leftarrow \sqrt{[\tilde{\mathbf{L}}_k]^2_{l,l} + |[\mathbf{w}]_l|^2}$, $c \leftarrow \frac{[\tilde{\mathbf{L}}_k]_{l,l}}{r}$, $s \leftarrow \frac{[\mathbf{w}]^*_l}{r}$
    $[\tilde{\mathbf{L}}_k]_{l:N,l} \leftarrow c[\tilde{\mathbf{L}}_k]_{l:N,l} + s[\mathbf{w}]_{l:N}$
  end if
  $[\mathbf{w}]_{l:N} \leftarrow c[\mathbf{w}]_{l:N} - s^*[\tilde{\mathbf{L}}_k]_{l:N,l}$
end for
$\mathbf{L}_k \leftarrow \tilde{\mathbf{L}}_k$

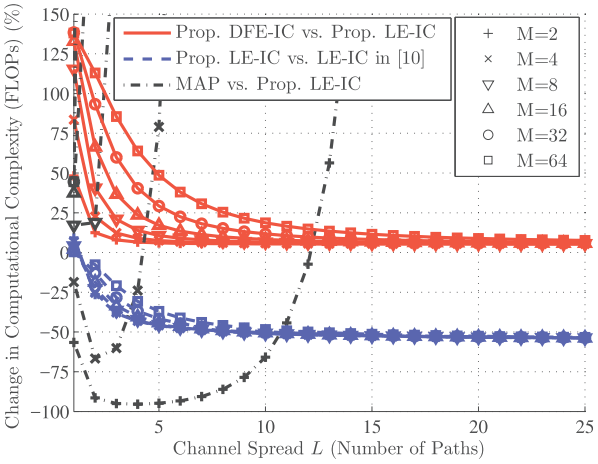

Fig. 4. Complexity comparison of LE-IC and DFE-IC with proposed matrix inversion algorithm.

a block length $K = 2048$ and a FIR window given by $N = 3L + 2$, $N_d = 2L$. The blue dashed curves show the FLOP count ratio of a LE-IC using our strategy relative to using the algorithm in [10], for different constellation orders. Up to 50% saving is observed as channel spread increases.

DFE-IC FLOP count is compared to LE-IC, both using the proposed inversion strategies, with red solid lines. This ratio is high for a low number channel taps, but decreases to 7% as L increases, more or less quickly depending on the modulation order M. Finally, MAP detector is seen to be an interesting alternative to FIR receivers for BPSK/QPSK signalling, in channels with very short channel spreads.

V. Comparison With the Prior Work on Time-Varying DFE-IC Structures

In this section, the DFE-IC based on EP feedback, proposed in section III-D, in its canonical form without self-iterations ($\mathcal{S}_\tau = 0$) and without damping is compared to alternative state-of-the-art TV DFE-IC structures.

First, to provide a fair performance comparison with alternatives, existing suboptimal DFE-IC schemes [19], [21], [24] are extended to time-varying structures using soft posterior feedback. Next, analytical and asymptotic analysis, and Monte Carlo simulations show the superiority of DFE-IC based on EP relative to LE-IC, classical DFE and concurrent DFE-IC structures.

A. On the TV DFE-IC Based on Bayesian Estimators

References on time-varying DFE-IC with soft feedback are limited. Hence, here existing methods are generalized and improved before comparison, thanks to our framework, in order to provide a fair comparison. Until EP, soft posterior feedback was the only imperfect feedback with a reasonable complexity in the literature, applicable to any constellation. Nevertheless, it is not possible to derive a structure using such feedback within the conventional BP formalism, but here its usage is justified with Bayesian inference.

One can consider the equalization problem within a Bayesian framework, where a particular realization of a random data symbol is estimated. For instance, the conventional MMSE linear turbo receiver [10] is also the MAP estimator, if priors are forced to lie in the family of Gaussian distributions [9]. Hence this equalizer is the unbiased Bayesian estimator $\mathbb{E}_{\mathcal{L}_k}[x_k|\mathbf{y}_k, \mathbf{H}_k]$, where the joint prior distribution $\mathcal{L}_k(\mathbf{x}_k) \propto \prod_{l=k-N'_p}^{k+N_d} \mathcal{CN}(x^p_l, v^p_l)$ is used. However, in Bayesian estimation theory, the mean square error can be further reduced, using a *sequential* MMSE estimator, which improves its posterior with previously estimated data ([41, Sect. 12.6]). Following this idea, we propose the improved estimator $\mathbb{E}_{\mathcal{A}_k}[x_k|\mathbf{y}_k, \mathbf{H}_k]$, based on the joint posterior $\mathcal{A}_k(\mathbf{x}_k) \propto \prod_{l=k}^{k+N_d} \mathcal{CN}(x^p_l, v^p_l) \prod_{l=k-N'_p}^{k-1} \mathcal{CN}(\mu^d_l, \gamma^d_l)$, where μ^d_l and γ^d_l are given by (20). In the following, we derive a posterior feedback based DFE-IC using this estimator for IC, with model (4).

1) Exact TV DFE-IC With APP Feedback: This equalizer is a generalization of invariant schemes in [23] and [24] to TV structures. It is derived by using the joint posterior $\mathcal{A}_k(\mathbf{x}_k)$ with the model (4), derived in the Appendix. The resulting APP FIR structure, is fully defined by

$$\begin{aligned} \bar{\mathbf{x}}^{\text{app}}_k &= [\mu^d_{k-N'_p}, \ldots, \mu^d_{k-1}, v^p_k, \ldots, x^p_{k+N_d}]^T, \\ \bar{\mathbf{v}}^{\text{app}}_k &= [\gamma^d_{k-N'_p}, \ldots, \gamma^d_{k-1}, v^p_k, \ldots, v^p_{k+N_d}]^T. \end{aligned} \tag{33}$$

This structure will be referred as *DFE-IC APP* in the remainder of this paper, and illustrated in Fig. 3 without the dashed module.

2) TV DFE-IC With Perfect APP Feedback: Here we propose to generalize [19], [26] to APP feedback, with perfect decision hypothesis. This imposes decision covariances to 0, focusing the MMSE filter design to only mitigate anti-causal symbol interference. However, its use of hard feedback, i.e. $\arg\max_\alpha \mathcal{D}_k(\alpha)$, was shown to be seriously prone to error propagation [19]. While [26] showed improvements with soft posterior feedback on non-turbo, invariant structures, here, we extend this case to time-varying turbo structures.

This case named *DFE-IC PAPP*, differs from the DFE-IC APP with the variance estimates:

$$\begin{aligned}\bar{\mathbf{x}}_k^{\mathbf{papp}} &= \bar{\mathbf{x}}_k^{\mathbf{app}},\\ \bar{\mathbf{v}}_k^{\mathbf{papp}} &= [\mathbf{0}_{N_p'}^T, v_k^p, \ldots, v_{k+N_d}^p]^T.\end{aligned} \tag{34}$$

3) Hybrid TV DFE-IC With APP Feedback: This structure is an extension of the TV structure from [21] to APP feedback. In [21], the DFE-IC with perfect hard decisions from [19] is improved by adding an estimate of the decision error to the equalizer output variance v_k^e. This quantity is given by $\mathrm{Var}_{\mathcal{D}_k}[\mathbf{g}_k^{cH}([\mathbf{x}-\boldsymbol{\mu}^{\mathbf{d}}]_{k-N_p':k-1})]$, using (19). Moreover, this structure checks whether this variance causes sign changes in extrinsic LLRs, and sets ambiguous LLRs to zero.

This receiver is extended to use APP soft feedback, instead of hard decisions, and denoted *DFE-IC HAPP*.

B. Analytic Comparison of DFE-IC vs. LE-IC

This paragraph semi-analytically assesses the behaviour of a DFE-IC relative to a LE-IC to underline the interest in jointly using decision feedback and prior information for IC.

In fact, LE-IC operating with priors $(\bar{x}_k, \bar{v}_k)$ provides a lower bound for the achievable information rate of a DFE-IC structure using the same prior information for its anti-causal symbols $(\bar{x}_k^{\mathrm{a}}, \bar{v}_k^{\mathrm{a}}) = (\bar{x}_k, \bar{v}_k)$, alongside decision feedback estimates $(\bar{x}_k^{\mathrm{c}}, \bar{v}_k^{\mathrm{c}})$ (see (32)). By exploiting the structural similarities between DFE-IC and LE-IC, the causal feedback's impact is reflected on a ratio of post-equalization SNR[3]

$$G = \frac{\mathrm{SNR}_{\mathrm{out}}^{\mathrm{dfe}}}{\mathrm{SNR}_{\mathrm{out}}^{\mathrm{le}}} = \frac{\sigma_x^2}{\mathbb{E}[v_k^{e(\mathrm{dfe})}]}\frac{\mathbb{E}[v_k^{e(\mathrm{le})}]}{\sigma_x^2} = \frac{\xi^{\mathrm{dfe}}}{\xi^{\mathrm{le}}}\frac{1-\bar{v}\xi^{\mathrm{le}}}{1-\bar{v}\xi^{\mathrm{dfe}}} \tag{35}$$

where $\bar{v} = \mathbb{E}[\bar{v}_k]$ and $\xi^{\mathrm{XX}} = \mathbb{E}[\xi_k^{\mathrm{XX}}]$, where XX is "le" or "dfe". This gain is greater than unity iff $\xi^{\mathrm{dfe}} \geq \xi^{\mathrm{le}}$, or equivalently iff $\mathbb{E}[\bar{\mathbf{V}}_k^{\mathbf{le}} - \bar{\mathbf{V}}_k^{\mathbf{dfe}}]$ is positive semi-definite. Hence having $\bar{v} > \bar{v}^c$, $\bar{v}^c = \mathbb{E}[\bar{v}_k^c]$ for DFE-IC is required for achieving improvements. Based on empirical and experimental evidence not presented here, the conjecture $\mathbb{P}[\bar{v}_k^c > \bar{v}_k] < 0.5$ has been verified over a wide range of input SNRs, and for random constellations, for $\bar{v}_k^c = v_k^d$ (DFE-IC EP) and for $\bar{v}_k^c = \gamma_k^d$ (DFE-IC APP). This ensures $\bar{v} > \bar{v}^c$ and thus, LE-IC output SNR is a lower bound on DFE-IC EP/APP, as possible detection degradations are small.

G is plotted in Fig. 5, with $N = 17, N_d = 10$ and $\sigma_x^2 = 1$ for the static Proakis-C channel, $\mathbf{h} = [1,2,3,2,1]/\sqrt{19}$; when decisions are more reliable than priors, G increases, otherwise DFE-IC brings small improvements. When $\bar{v}^a \to 1$, there is no prior information, and decisions bring a significant gain. Oppositely, when $\bar{v}^a \to 0$, prior information is already close to the ideal, and DFE-IC cannot improve further. This indicates boosted performance at initial turbo-iterations.

C. Asymptotic Analysis and Performance Prediction

To assess the full potential of DFE-IC, asymptotic analysis is used to evaluate its achievable rates. Extrinsic information

[3] $\mathrm{SNR}_{\mathrm{out}}^{\mathrm{XX}} = \sigma_x^2/\mathbb{E}[v_k^{e(\mathrm{XX})}]$ is the post-equalization SNR, where XX is "dfe" or "le", (see (4) for v_k^e). Superscript "le" refers to the use of $(\bar{x}_k, \bar{v}_k)$ for IC, and "dfe" refers to the use of $(\bar{x}_k^{\mathrm{a}}, \bar{v}_k^{\mathrm{a}})$ and $(\bar{x}_k^{\mathrm{c}}, \bar{v}_k^{\mathrm{c}})$ for IC, as in (32).

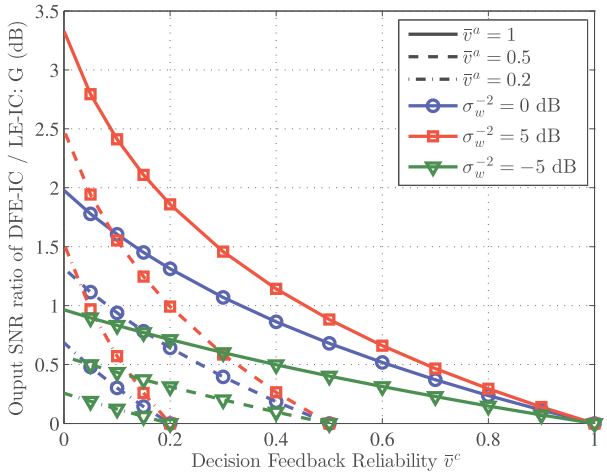

Fig. 5. Post-equalization SNR ratio G depending on channel SNR σ_w^{-2}, prior reliability $\bar{v}^a$ and "decision" reliability $\bar{v}^c$.

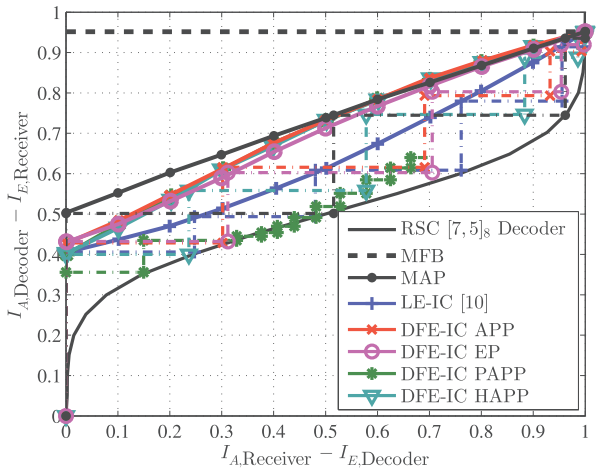

Fig. 6. EXIT curves and average MI trajectories of FIR equalizers with BPSK in Proakis C channel at $E_b/N_0 = 7$dB.

transfer (EXIT) analysis [42] of a SISO module is used as a tool for characterizing its asymptotic limits, by tracking extrinsic mutual information (MI) exchanges between the iterative components. Essentially, a SISO receiver can be characterized by a simple transfer function $I_E = \mathcal{T}_R(I_A, \mathbf{H}, \sigma_w^2)$, where I_A and I_E are the MI between coded bits and respectively its input prior LLRs and output extrinsic LLRs, and σ_w^2 and $\mathbf{H}$ show its dependence on the channel and the received SNR.

In Fig. 6, transfer curves $\mathcal{T}_R$ are plotted in solid lines for considered receivers along with the reverse transfer $\mathcal{T}_D^{-1}$ of the BCJR decoder of a recursive systematic convolutional (RSC) code. DFE-IC APP yields a higher I_R than LE-IC for all I_A, unsurprisingly given the posterior feedback, and there is little difference with DFE-IC EP, which has slightly lower rates at low prior information. In particular, the improvement at $I_A = 0$ lets us conjecture a lower waterfall threshold in BPSK, and the higher slope of the $\mathcal{T}_R$ curve at low I_A hints an improved convergence speed across turbo iterations.

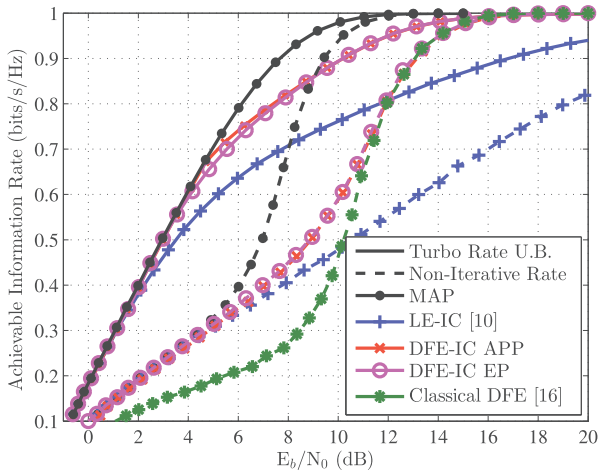

Fig. 7. Achievable spectral efficiency on deterministic Proakis C channel with BPSK.

Another use of EXIT analysis is performance prediction, however this involves strong assumptions on prior inputs that often cannot be met for FIR turbo equalizers in practice. Hence, EXIT curves only provide an upper-bound on information rate for receivers other than MAP. In this respect, it is then interesting to compare transfer curves, with actual MI trajectories (in dashed lines in Fig. 6).

It had been noted in [19], that trajectories of DFE-IC with hard, "perfect" decision assumption do not follow EXIT curves; this issue remains with DFE-IC PAPP, although less severely, indicating that the "perfect decisions" assumption causes a severe information loss. Other FIRs' trajectories overall follow receiver and decoder curves and reach MFB, but after a few iterations, they no longer make contact with transfer curves, losing convergence speed. This is a common disadvantage of FIR equalizers, attributed to short cycles caused by neighbouring symbol correlations, as shown in [19, Fig. 16]. However note that among DFE-IC receiver, EP feedback yields trajectories that remains closest to EXIT curves, making it easier to predict.

The achievable spectral efficiency for a given receiver can be measured with the help of the area theorem for EXIT charts [43]. In Fig. 7, achievable rates for BPSK constellation are plotted. Note that for MAP receivers, this rate is an accurate approximation of the channel symmetric information rate (SIR) [44]. As non-iterative FIR do not depend on prior inputs, their achievable rates are also accurately computed. For turbo FIR, upper bounds are obtained by combining results of area theorem with the channel SIR. Tightness of this bound depend on the closeness of true MI trajectories to EXIT charts in Fig. 6, so APP feedback's asymptotic performance is likely to be overestimated compared to EP feedback.

D. Finite-Length Comparison With Existing Schemes

Monte Carlo integration remains the most reliable analysis approach joint detection of BPSK symbols is considered with parameters in section V-B, and $K_u = 2048$, coded with a terminated $[7, 5]_8$ RSC code. Bit error rate (BER) of various receivers are plotted in Fig. 8. For the reported iterations, the DFE-IC APP outperforms other APP feedback DFE structures, and their convergence speeds are compared on the right side of the figure, at a block error rate (BLER) of 10^{-2}. EP-based feedback provides further improvement of the threshold by 0.5 dB relative to APP, and it is shown to reach MFB limit within 7 iterations, earlier than DFE-IC APP.

Assessing DFE-IC performance at low spectral efficiency conditions, as above, is of interest, to remedy the poor behaviour of classical DFE at those operating points (see Fig. 7). Indeed, turbo processing helps DFE structures to outperform LE at all rates. A higher spectral efficiency case is plotted on the left side of the Fig. 9, with 8-PSK constellation in the same configuration; DFE-IC APP is shown to improve LE-IC waterfall threshold by 2dB, DFE-IC EP asymptotically provides an additional 1.2dB. On the right side of the Fig. 9, 16-QAM is considered; showing that DFE-IC EP provides further performance enhancements for one or more iterations.

Finally, the coded performance of DFE-IC is balanced with complexity considerations. In Fig. 10, the receiver computational complexity (FLOPs per symbol) required to decode with a BLER of 10^{-2} is plotted as a function of E_b/N_0. These values are computed, assuming the use of the proposed matrix inversion algorithm in section IV, and by accounting for the equalization, the demapping and the decoding costs. A curve represents the evolution of BLER and the computational costs of a receiver accross turbo iterations.

DFE-IC provides a better trade-off than LE-IC; at any given complexity, it is more efficient, especially at initial iterations, and the asymptotic E_b/N_0 gap between LE-IC and DFE-IC increases with the modulation order M. The use of EP feedback is more advantageous at higher iterations, for higher order constellations, while APP is more efficient for non-iterative receivers.

In conclusion, DFE-IC outperforms LE-IC in various aspects: it converges faster towards MFB, has a lower decoding threshold than LE-IC, especially at higher spectral efficiencies. Among DFE-IC with APP feedback, exact derivation DFE-IC APP is superior according to both finite-length and asymptotic analysis. Although EXIT charts show little difference between DFE-IC EP and APP, in practical simulations EP feedback tends to outperform APP. This is justified by the tightness of EP MI trajectories to EXIT curves; APP is overestimated. Although it DFE-IC EP appears to be able to reach channel SIR at low to medium spectral efficiencies, there is still a gap to MAP performance.

In the following, the use of self-iterations will be assessed to further improve performances.

VI. Comparison With the Prior Work on Self-Iterated EP Structures

Some recent EP-based receivers [32], [33], [35]–[37] have observed remarkable performance improvements in repeating the detection process in a parallel schedule through self-iterations. As the demapping process is computationally

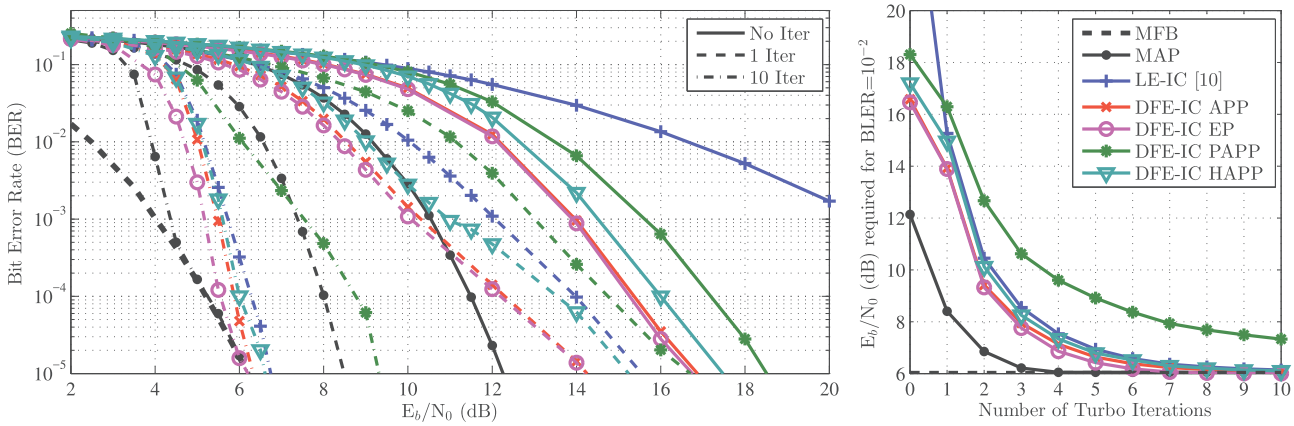

Fig. 8. BER and convergence performance of the proposed DFE-IC in Proakis C channel with BPSK constellation.

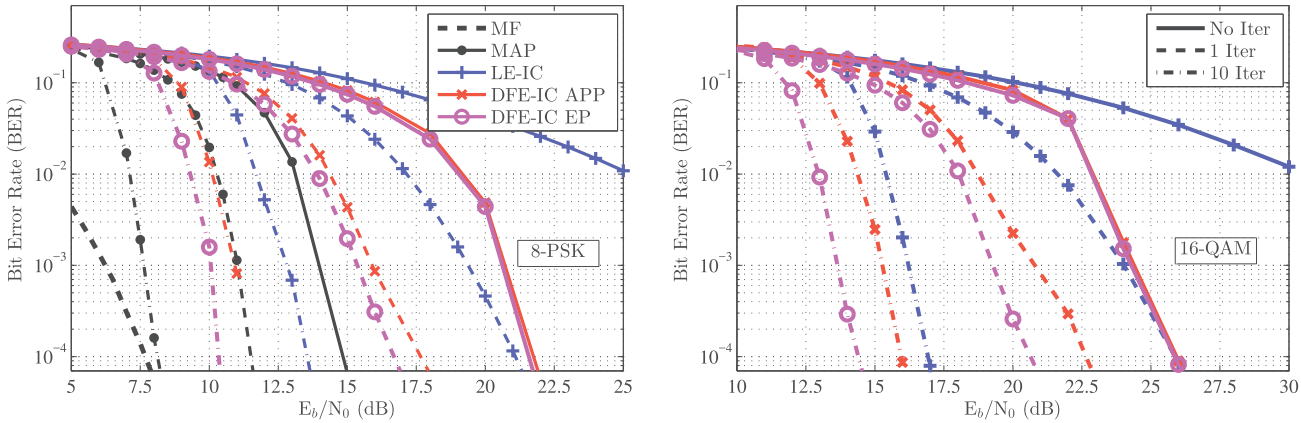

Fig. 9. BER performance of the proposed DFE-IC in Proakis-C with 8-PSK and 16-QAM constellations.

less intensive than channel decoding, such structures are of practical interest. In this section, the benefits in using a self-iterated DFE-IC EP compared to structures in prior work is investigated.

Independently of our work, an EP-based FIR structure is derived in the concomitant work [36]. Unlike the message passing formalism used in section III, structure in [36] is obtained by approximating a self-iterated block receiver, derived by EP-based approximation of the posterior PDF (5). The resulting FIR structure uses a *parallel* schedule, and corresponds to a LE-IC within each SI. Using our formalism, it is equivalent to updating all VNs x_k with messages from EQU sequentially, and only then activating DEM to update posterior approximations. This process is iterated with DEM sending back an extrinsic message to EQU, and finally DEM computes messages towards DEC. In the following, the structure denoted as "EP-F" in [36], is refered as a self-iterated LE-IC (SI LE-IC), with following mean and variances used for IC

$$\bar{\mathbf{x}}_k^{\text{le-ep}(s)} = [x_{k-N_p'}^{d(s)}, \ldots, x_{k+N_d}^{d(s)}]^T,$$
$$\bar{\mathbf{v}}_k^{\text{le-ep}(s)} = [v_{k-N_p'}^{d(s)}, \ldots, v_{k+N_d}^{d(s)}]^T. \quad (36)$$

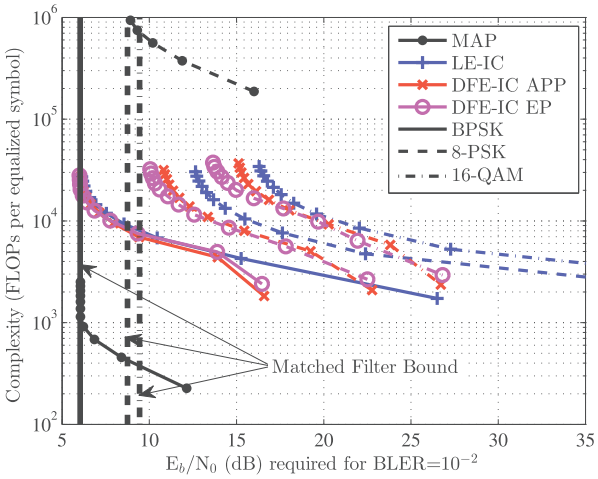

Fig. 10. Performance complexity trade-off in Proakis C.

If the computations of messages on EQU is carried out only once ($\mathcal{S}_\tau = 0$), this receiver yields the same result as the conventional turbo LE-IC [10].

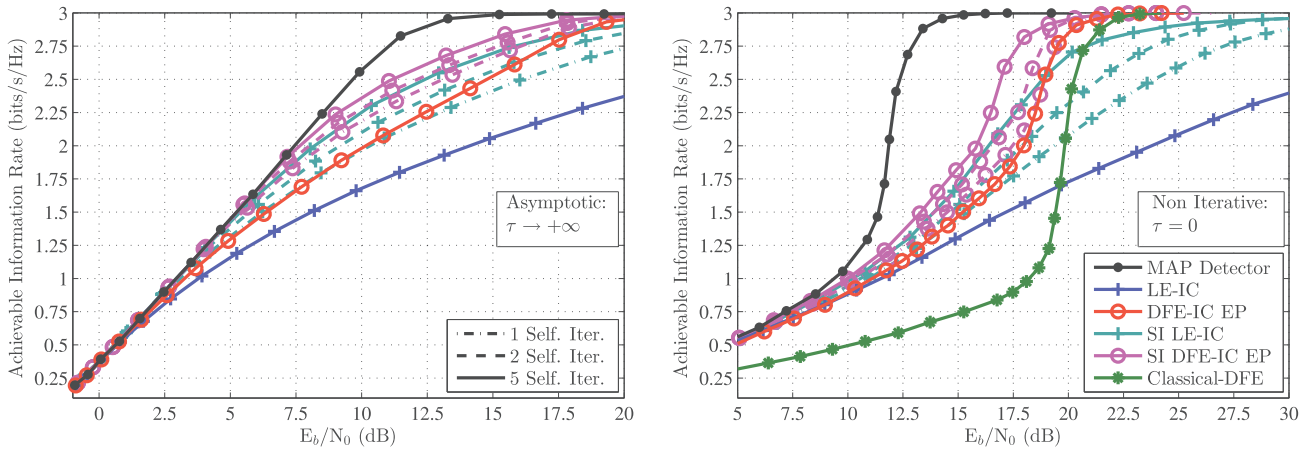

Fig. 11. Achievable Rates of Self-iterated LE-IC and DFE-IC in Proakis-C with 8-PSK constellation.

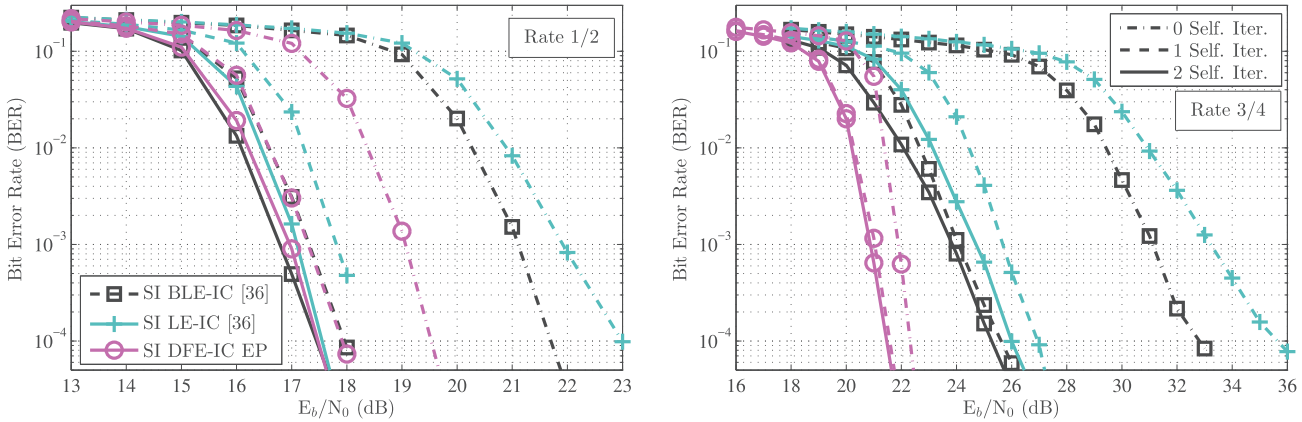

Fig. 12. SI LE-IC and DFE-IC in Proakis C with LDPC coded 16-QAM, with 5 turbo iterations.

A. Asymptotic Comparison

First, we look into the achievable rates of SI LE-IC and DFE-IC EP to identify operating points where self-iterations have an advantage.

We consider 8-PSK signalling on the Proakis-C channel, and use the area theorem to obtain an upper bound on asymptotic achievable rates (i.e. $\tau \to \infty$), plotted on the left side of Fig. 11. Information rates of the optimal MAP detector, LE-IC and DFE-IC EP without SI, and SI LE-IC and SI DFE-IC are considered. For self-iterated receivers, a static damping with $\beta = 0.6$ is used. Numerical results show that SI is not required for LE-IC up to 0.75 bits/s/Hz (i.e. using a code rate less than 1/4), as LE-IC is close to the SIR, whereas DFE-IC EP continues to follow MAP rates up to 1 bit/s/Hz (up to a code rate of 1/3). On the other hand, when using 5 self-iterations, DFE-IC EP follows MAP rates within 0.5 dB up to 2.25 bits/s/Hz, while LE-IC follows up to 1.85 bits/s/Hz. It is also interesting to note that DFE-IC EP with 2 SI outperforms LE-IC with 5 SI, at all rates, indicating at faster convergence of DFE-IC EP towards asymptotic limits.

At the right side of Fig. 11, non-turbo iterative achievable rates of these receivers, and those of the classical DFE [16], are compared. These rates are accurate, and not an upper bound, unlike asymptotic rates, and note that MAP detector is a mere maximum likelihood (ML) detector in this case. Although self-iterations significantly improve LE-IC performance, at rates above 2.75 bits/s/Hz, classical DFE still outperforms these receivers. DFE-IC EP on the other hand outperforms alternative FIRs at any given self iteration.

Note that the gap to capacity still remains significant for non turbo iterative rates, and to some extent, for asymptotic rates. Hence with the objective of deriving capacity achieving practical receivers in mind, future work should explore the usage of the proposed DFE-IC EP as a constituent element for bidirectional DFE [20] or for concatenated FIR [22] receivers.

B. Finite-Length Comparison

In this section, numerical finite-length results complete the previous analysis. In addition to receivers above, the self-iterated block linear receiver (SI BLE-IC), denoted nuBEP in [36], is considered. Without self-iterations, this receiver

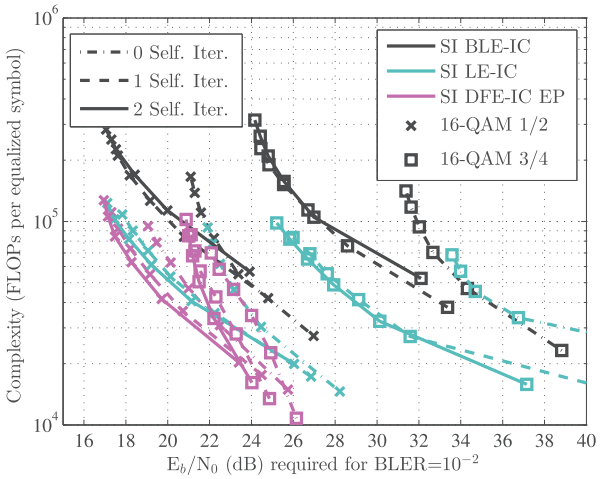

Fig. 13. Performance complexity trade-off for self-iterations in LDPC coded Proakis C.

is equivalent to turbo block LE-IC [45], and it outperforms the self-iterated block receiver and Kalman smoother in [33] and [35]. SI BLE-IC provides a lower bound to the BER performance of SI LE-IC.

A low density parity check (LDPC) coded 16-QAM transmissions over the Proakis C channel, with rate 1/2 and 3/4 encoding of $K_b = 2048$ bits (Fig. 12). The proposed SI DFE-IC EP uses respectively $\beta = \min(0.5, 1\text{-e}^{\tau/2.5}/10)$ and $\beta = \min(0.1, 1 - e^{\tau/1.5}/10)$ for damping, in these two cases, whereas the optimized damping reported in [36] is kept for SI BLE-IC and SI LE-IC. The LDPC codes are obtained by path edge growth method, and a BP decoder up to a 100 iterations is used. The low rate case, with (3,6) regular LDPC, shows that while all self-iterated receivers reach the same asymptotic performance as $\mathcal{S}_\tau$ increases, DFE-IC converges much faster at intermediary iterations. On the other hand, at the high rate configuration, with (3,12) regular LDPC, DFE-IC is strictly superior to LE-IC, even without self-iterations. Asymptotically even the exact SI BLE-IC is 3.8 dB behind the proposed SI DFE-IC.

These numerical performance results are completed with computational complexity considerations in Fig. 13, where decoding threshold for BLER $= 10^{-2}$ is evaluated for $\tau = 0, \ldots, 5$, for each receiver. In the medium rate (2 bits/s/Hz: 16-QAM with rate 1/2 code) case the three considered receivers converges to the same asymptotic limit near 17 dB, but DFE-IC offers lower complexity at intermediary iterations. At 3 bits/s/Hz configuration (16-QAM with rate 3/4 code), with 5 TI and 3 SI, DFE-IC requires 3 dB less energy, and 3 times less computational resources than BLE-IC. With $\tau = s = 0$, LE-IC is unable to decode, BLE-IC decodes around 39 dB, and DFE-IC decodes with 13 dB less energy.

These numerical results confirms conclusions drawn by the asymptotic analysis; the proposed SI DFE-IC is of a significant interest for high data rate applications where linear structures are less efficient. Using the efficient implementation method of section IV, DFE-IC outperforms prior work in terms of both complexity and performance.

VII. Conclusion

This paper investigates on the use of decision feedback with turbo equalization, for improving the limitations of linear equalizers for high data rate applications.

Turbo DFE structures in the literature consist in either using hard feedback with symbol-wise adaptive filters, or soft posterior feedback with symbol-wise invariant filters. The former perform poorly at low spectral efficiency, and require complex mechanisms to improve this issue, whereas the latter are outperformed even by the conventional TV LE-IC. Both schemes are extended to time-variant soft feedback structures in this paper, with different filter computation hypotheses. We show that an exact approach justified with sequential Bayesian MMSE estimators (DFE-IC APP) outperforms other APP feedback alternatives.

However, due to the use of posterior estimates, this structure does not fit within the turbo principle which requires the exchange of extrinsic information. Consequently, we focus our discourse on the derivation of FIR DFE within the expectation propagation framework, which allows the computation of a novel type of extrinsic feedback from the demapper to the equalizer. Building upon the emerging trend on self-iterated EP-based equalizers, the proposed DFE-IC can be self-iterated to further improve performances.

Thanks to finite-length and asymptotic analysis, DFE-IC EP, with SI or not, is shown to set new upper limits in achievable performance among FIR turbo receivers. At high data rates, even exact self-iterated block linear receivers fall over 3 dB behind the proposal.

Finally, the gap of achievable rates by turbo DFE-IC to the channel capacity remains still significant at very high spectral efficiencies. Bidirectional extension of TV DFE-EP should be explored to try to close this gap.

Appendix
Derivation of MMSE FIR With IC

In this appendix, FIR equalization with interference cancellation is derived by minimizing the Bayesian MMSE criterion $J = \mathbb{E}_{\mathcal{A}_k}[|x_k - x_k^{e\prime}|^2]$, where $x_k^{e\prime} = \mathbf{f}_k'^T \mathbf{y}_k + g_k'$ is the equalized linear estimate, and $\mathcal{A}_k$ is a joint multivariate Gaussian prior distribution on $\mathbf{x}_k$ defined with means $\bar{\mathbf{x}}_k^{\mathbf{fir}}$ and variances $\bar{\mathbf{v}}_k^{\mathbf{fir}}$ (see sec. II-B). $\mathbb{E}_{\mathcal{A}_k}[\cdot]$ and $\mathbf{Cov}_{\mathcal{A}_k}[\cdot]$ respectively denote the expectation and the covariance with respect to distribution $\mathcal{A}_k$. Solution to this is given by $\mathbb{E}_{\mathcal{A}_k}[x_k|\mathbf{y}_k, \mathbf{H}_k]$, i.e. the symbol mean with respect to $p_{\mathcal{A}_k}(x_k|\mathbf{y}_k, \mathbf{H}_k)$. This distribution is the marginalization of the conjugate Gaussian posterior $p_{\mathcal{A}_k}(\mathbf{x}_k|\mathbf{y}_k, \mathbf{H}_k)$, i.e. of likelihood $p(\mathbf{y}_k|\mathbf{x}_k, \mathbf{H}_k)$ and prior $\mathcal{A}_k$. Hence, $x_k^{e\prime}$ is deduced by multiplying the MMSE estimator of $\mathbf{x}_k$ [41] by $\mathbf{e}_k$:

$$\mathbf{f}_k' = \mathbf{e}_k^H \mathbf{Cov}_{\mathcal{A}_k}[\mathbf{y}_k, \mathbf{x}_k](\mathbf{Var}_{\mathcal{A}_k}[\mathbf{y}_k])^{-1}, \tag{37}$$

$$g_k' = \mathbf{e}_k^H \mathbb{E}_{\mathcal{A}_k}[\mathbf{x}_k] - \mathbf{f}_k'^T \mathbb{E}_{\mathcal{A}_k}[\mathbf{y}_k], \tag{38}$$

by developing expectations above with prior statistics, it holds

$$\mathbf{f}_k' = \bar{v}_k^{\text{fir}} \mathbf{h}_k^H (\mathbf{\Sigma}_k^{\text{fir}})^{-1}, \tag{39}$$

$$g_k' = \bar{x}_k^{\text{fir}} - \mathbf{f}_k'^T \bar{\mathbf{x}}_k^{\text{fir}}, \tag{40}$$

with $\mathbf{\Sigma}_k^{\text{fir}} = k_w \sigma_w^2 \mathbf{I}_N + \mathbf{H}_k \mathbf{V}_k^{\mathbf{fir}} \mathbf{H}_k^H$ and $\mathbf{V}_k^{\mathbf{fir}} = \mathbf{diag}(\mathbf{v}_k^{\mathbf{fir}})$. This receiver is biased, as its MMSE estimators' nature:

$$\mathbb{E}_{\mathcal{A}_k}[x_k^{e\prime} | x_k = x] = (1 - \bar{v}_k^{\text{fir}} \xi_k^{\text{fir}}) \bar{x}_k^{\text{fir}} + \bar{v}_k^{\text{fir}} \xi_k^{\text{fir}} x,$$

with $\xi_k^{\text{fir}} = \mathbf{h}_k^H \mathbf{\Sigma}_k^{\text{fir}-1} \mathbf{h}_k$. Removing additive and multiplicative biases with $x_k^e = (x_k^{e\prime} - (1 - \bar{v}_k^{\text{fir}} \xi_k^{\text{fir}}) \bar{x}_k^{\text{fir}}) / (\bar{v}_k^{\text{fir}} \xi_k^{\text{fir}})$ yields the estimator given in (4), which completes the proof.

REFERENCES

[1] C. Douillard *et al.*, "Iterative correction of intersymbol interference: Turbo-equalization," *Trans. Emerg. Telecommun. Technol.*, vol. 6, no. 5, pp. 507–511, Sep. 1995.

[2] 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.

[3] G. Bauch and V. Franz, "A comparison of soft-in/soft-out algorithms for 'turbo detection,'" in *Proc. Int. Conf. Telecommun.*, 1994, pp. 259–263.

[4] A. Glavieux, C. Laot, and J. Labat, "Turbo equalization over a frequency selective channel," in *Proc. 1st Symp. Turbo Codes*, 1997, pp. 96–102.

[5] A. Gersho and T. L. Lim, "Adaptive cancellation of intersymbol interference for data transmission," *Bell Syst. Tech. J.*, vol. 60, no. 9, pp. 1997–2021 Nov. 1981.

[6] I. Fijalkow, A. Roumy, S. Ronger, D. Pirez, and P. Vila, "Improved interference cancellation for turbo-equalization," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 1, Jun. 2000, pp. 416–419.

[7] A. Roumy, "Egalisation et décodage conjoints: Méthodes turbo," Ph.D. dissertation, Dept. Sci. Eng., Univ. Cergy-Pontoise, Cergy, France, Oct. 2000.

[8] 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.

[9] J. Boutros and G. Caire, "Iterative multiuser joint decoding: Unified framework and asymptotic analysis," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1772–1793, Jul. 2002.

[10] M. Tüchler, A. C. Singer, and R. Koetter, "Minimum mean squared error equalization using *a priori* information," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 673–683, Mar. 2002.

[11] A. Dejonghe and L. Vandendorpe, "Turbo-equalization for multilevel modulation: An efficient low-complexity scheme," in *Proc. IEEE Int. Conf. Commun.*, vol. 3, Apr. 2002, pp. 1863–1867.

[12] M. Tüchler and J. Hagenauer, "Linear time and frequency domain turbo equalization," in *Proc. IEEE 54rd Veh. Technol. Conf.*, vol. 4, Nov. 2001, pp. 2773–2777.

[13] R. Otnes and M. Tüchler, "Iterative channel estimation for turbo equalization of time-varying frequency-selective channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1918–1923, Nov. 2004.

[14] R. Visoz, A. O. Berthet, and S. Chtourou, "A new class of iterative equalizers for space-time BICM over MIMO block fading multipath AWGN channel," *IEEE Trans. Commun.*, vol. 53, no. 12, pp. 2076–2091 Dec. 2005.

[15] R. Le Bidan, "Turbo-equalization for bandwith-efficient digital communications over frequency-selective channels," Ph.D. dissertation, Dept. Signal Commun., INSA Rennes, Rennes, France, Jan. 2003.

[16] C. A. Belfiore and J. H. Park, "Decision feedback equalization," *Proc. IEEE*, vol. 67, no. 8, pp. 1143–1156, Aug. 1979.

[17] J. M. Cioffi, *Equalization*. Stanford, CA, USA: Stanford Univ., 2008, ch. 3. [Online]. Available: http://web.stanford.edu/group/cioffi/doc/book/chap3.pdf

[18] M. A. Elgenedy, E. Sourour, and M. Nafie, "Iterative MMSE-DFE equalizer for the high data rates HF waveforms in the HF channel," in *Proc. Asilomar*, Nov. 2013, pp. 1243–1247.

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

[20] S. Jeong and J. Moon, "Soft-in soft-out DFE and bi-directional DFE," *IEEE Trans. Commun.*, vol. 59, no. 10, pp. 2729–2741, Oct. 2011.

[21] S. Jeong and J. Moon, "Turbo equalization based on bi-directional DFE," in *Proc. IEEE Int. Conf. Commun.*, May 2010, pp. 1–6.

[22] S. Jeong and J. Moon, "Self-iterating soft equalizer," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3697–3709, Sep. 2013.

[23] R. R. Lopes and J. R. Barry, "The soft-feedback equalizer for turbo equalization of highly dispersive channels," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 783–788, May 2006.

[24] J. Tao, "On low-complexity soft-input soft-output decision-feedback equalizers," *IEEE Commun. Lett.*, vol. 20, no. 9, pp. 1737–1740, Sep. 2016.

[25] V. D. Trajkovic, P. B. Rapajic, and R. A. Kennedy, "Turbo DFE algorithm with imperfect decision feedback," *IEEE Signal Process. Lett.*, vol. 12, no. 12, pp. 820–823, Dec. 2005.

[26] J. Balakrishnan, "Mitigation of error propagation in decision feedback equalization," M.S. thesis, Cornell Univ., Ithaca, NY, USA, 1999.

[27] H. Lou and C. Xiao, "Soft-decision feedback turbo equalization for multilevel modulations," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 186–195, Jan. 2011.

[28] J. W. Choi, A. C. Singer, J. W. Lee, and N. I. Cho, "An improved soft feedback V-BLAST detection technique for TURBO-MIMO systems," in *Proc. IEEE ICASSP*, Mar. 2008, pp. 3181–3184.

[29] J. Tao, J. Wu, Y. R. Zheng, and C. Xiao, "Enhanced MIMO LMMSE turbo equalization: Algorithm, simulations, and undersea experimental results," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3813–3823, Aug. 2011.

[30] T. P. Minka, "A family of algorithms for approximate Bayesian inference," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., MIT, Cambridge, MA, USA, Jan. 2001.

[31] J. M. Walsh, "Distributed iterative decoding and estimation via expectation propagation: Performance and convergence," Ph.D. dissertation, Faculty Graduate School, Cornell Univ., Ithaca, NY, USA, May 2006.

[32] M. Senst and G. Ascheid, "How the framework of expectation propagation yields an iterative IC-LMMSE MIMO receiver," in *Proc. GLOBECOM*, Dec. 2011, pp. 1–6.

[33] I. Santos, J. J. Murillo-Fuentes, R. Boloix-Tortosa, E. Arias-de-Reyna, and P. M. Olmos, "Expectation propagation as turbo equalizer in ISI channels," *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 360–370, Jan. 2017.

[34] P. Sun, C. Zhang, Z. Wang, C. N. Manchón, and B. H. Fleury, "Iterative receiver design for ISI channels using combined belief- and expectation-propagation," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1733–1737, Oct. 2015.

[35] I. Santos, J. J. Murillo-Fuentes, E. Arias-de-Reyna, and P. M. Olmos, "Probabilistic equalization with a smoothing expectation propagation approach," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2950–2962 May 2017.

[36] I. Santos, J. J. Murillo-Fuentes, E. Arias-de-Reyna, and P. M. Olmos. (2018). "Turbo EP-based equalization: A filter-type implementation." [Online]. Available: https://arxiv.org/abs/1711.08188

[37] S. Şahin, A. M. Cipriano, C. Poulliat, and M.-L. Boucheret, "Iterative equalization based on expectation propagation: A frequency domain approach," in *Proc. IEEE 26th Eur. Signal Process. Conf.*, Sep. 2018.

[38] T. Minka, "Divergence measures and message passing," Microsoft Res., Cambridge, U.K., Tech. Rep. MSR-TR-2005-173, 2005.

[39] C. Studer, S. Fateh, and D. Seethaler, "ASIC implementation of soft-input soft-output MIMO detection using MMSE parallel interference cancellation," *IEEE J. Solid-State Circuits*, vol. 46, no. 7, pp. 1754–1765 Jul. 2011.

[40] G. H. Golub and C. F. Van Loan, *Matrix Computations*, vol. 3. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1996.

[41] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[42] S. ten Brink, "Designing iterative decoding schemes with the extrinsic information transfer chart," *AEU Int. J. Electron. Commun.*, vol. 54, no. 6, pp. 389–398, Nov. 2000.

[43] J. Hagenauer, "The EXIT chart—Introduction to extrinsic information transfer in iterative processing," in *Proc. IEEE 12th Eur. Signal Process. Conf.*, Sep. 2004, pp. 1541–1548.

[44] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavcic, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3498–3508, Aug. 2006.

[45] M. Tüchler and A. C. Singer, "Turbo equalization: An overview," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 920–952, Feb. 2011.

Serdar Şahin was born in Ankara, Turkey, in 1992. He received the M.Sc. Eng. degree in control systems and electronics engineering from the INSA de Toulouse, University of Toulouse, France, in 2015. He is currently pursuing the Ph.D. degree in digital communications with IRIT-ENSEEIHT, Toulouse, and Thales Communications and Security, Gennevilliers. His main research interests include iterative receiver design, practical cooperative transmission schemes, and PHY layer abstraction.

Antonio Maria Cipriano was born in Padua, Italy, in 1976. He received the Laurea degree in telecommunications engineering from the University of Padova, Italy, in 2000, and the Ph.D. degree in digital communications jointly from the University of Padova and the Ecole Nationale Supérieure des Télécommunications (ENST) Paris, France, in 2005. In 2001, he spent eight months working as a Young Engineer at Eutelsat, France. He held post-doctoral position at Télécom ParisTech, Paris, and the France Telecom Research and Development (now Orange Labs), from 2005 to 2007. In 2007, he joined Thales Communications and Security as a Digital Communication Engineer and was involved in several national and international research projects on 4G and 5G communication systems. His main research interests lie in the broad area of digital communication systems. He is currently involved in research about PHY layer abstractions, relaying for ad hoc mobile networks and advanced receiver design.

Charly Poulliat received the E.E. degree from the Ecole Nationale Supérieure de l'Electronique et de ses Applications (ENSEA), Cergy-Pontoise, France, in 2001, and the M.S. degree in image and signal processing, the Ph.D. degree in electrical and computer engineering, and the Habilitation degree (HDR) from the University of Cergy-Pontoise, France, in 2001, 2004, and 2010, respectively. From 2004 to 2005, he was a Post-Doctoral Researcher with the UH Coding Group supervised by Prof. M. Fossorier, University of Hawaii at Manoa, HI, USA. In 2005, he joined the Signal and Telecommunications Department, Engineering School, ENSEA, as an Assistant Professor. Since 2011, he has been a Professor with the National Polytechnic Institute of Toulouse, INP-ENSEEIHT, University of Toulouse. He is also with the Signal and Communications Group, CNRS IRIT Laboratory. His research interests include signal processing for digital communications, waveform design, channel coding, iterative system design, and optimization.

Marie-Laure Boucheret received the Eng. degree in electrical engineering from ENST Bretagne, Brest, France, in 1985, the Ph.D. degree in digital communications from Télécom ParisTech in 1997, and the Habilitation à diriger les recherches degree from the INPT University of Toulouse in 1999. From 1985 to 1986, she has been a Research Engineer with the French Philips Research Laboratory. From 1986 to 1991, she has been an Engineer with Thales Alenia Space, first as a Project Engineer (TELECOM II Program) then as a Study Engineer at the Transmission Laboratory. From 1991 to 2005, she was an Associate Professor then a Professor with Télécom ParisTech. Since 2005, she has been a Professor with the National Polytechnic Institute of Toulouse, ENSEEIHT, University of Toulouse. She is also with the Signal and Communication Group, IRIT Laboratory. Her fields of interest are signal processing for communication and satellite communications.