A Framework for Iterative Frequency Domain EP-based Receiver Design
Serdar Şahin, Antonio Maria Cipriano, Charly Poulliat and Marie-Laure Boucheret

Abstract—An original expectation propagation (EP) based message passing framework is introduced, wherein transmitted symbols are considered to belong to the multivariate white Gaussian distribution family. This approach allows deriving a novel class of single-tap frequency domain (FD) receivers with a quasi-linear computational complexity in block length, thanks to Fast-Fourier transform (FFT) based implementation. This framework is exposed in detail, through the design of a novel double-loop single-carrier frequency domain equalizer (SC-FDE), where self-iterations of the equalizer with the demapper, and turbo iterations with the decoder, provide numerous combinations for the performance and complexity trade-off. Furthermore, the flexibility of this framework is illustrated with the derivation of an overlap FDE, used for time-varying channel equalization, among others, and with the design of a FD multiple-input multiple-output (MIMO) detector, used for spatial multiplexing. Through these different receiver design problems, this framework is shown to improve the mitigation of inter-symbol, inter-block and multi-antenna interferences, compared to alternative single-tap FD structures of previous works. Thanks to finite-length and asymptotic analysis, supported by numerical results, the improvement brought by the proposed structures is assessed, and then completed by also accounting for computational costs.

Index Terms—Interference cancellation, expectation propagation, frequency domain equalization, turbo equalization.

I. INTRODUCTION

From a communication theory perspective, the design of receivers with affordable complexity to reach the optimum maximum likelihood joint detection and decoding performance is of interest. A major milestone, in this regard, is the discovery of turbo-codes, which paved the way for research on iterative processing techniques built around soft-input soft-output (SISO) receivers [3]. In particular, a turbo receiver using a maximum a posteriori (MAP) detector and a MAP decoder is able to operate at channel symmetric information rate (SIR), thanks to the BCJR algorithm, at the expense of an exponentially scaling computational complexity [4], [5].

In equalization, where MAP detectors are limited to applications with low modulation orders and very short delay spreads, a vast literature exists on extending conventional minimum mean square error (MMSE) linear equalizer (LE) or decision feedback equalizer (DFE) to turbo processing with interference cancellation (IC) [6]–[9]. Among those, block receivers, offer best performance with a computational cost scaling at best quadratically in block length, and approximate finite-impulse response receivers have quadratic complexity in channel spread.

When the statistics of the prior symbol feedback from the decoder is white (i.e. the reliability of prior estimates is static over the block), block linear equalizers (BLE) can be efficiently implemented via FFTs as FD LE, with the so-called “one-tap” filters, where each frequency bin (also called sub-carrier) is independently processed in parallel. Hence, in general, by whitening the estimates used for IC, iterative receivers can be built using one-tap FDEs, with a computational complexity scaling quasi-linearly with the block length [9], [10].

Despite the improvements brought by the turbo-iterations, there is a significant gap between FDE achievable rates and the channel SIR, especially in moderately or highly selective channels. Consequently, non-linear extensions have been explored to improve FDE performance [11]–[17].

Recently, new ideas on Bayesian inference, used in the field of artificial intelligence for solving classification or probability density functions (PDF) estimation problems arouse the interest of the communication theory and signal processing communities. Expectation propagation [23] is a technique for approximate Bayesian inference, which can be used as a message passing algorithm that extends the loopy belief propagation (BP) [24], conventionally used for turbo receiver design. Indeed, EP is used with variables having PDF from the exponential family, which allows for the computation of the posterior.
of symbol-wise extrinsic information, in the context of soft
demapping, that was lost when using BP [25]. There are
various recent receiver proposals, that observed remarkable
performance improvements by exploiting EP [25–50].

This paper introduces a novel category of frequency domain
receivers, obtained by a specific framework of expectation
propagation based message passing algorithm. This approach
is exposed through the design of an elementary FDE frame-
work, and then the impact of this methodology on more
advanced receivers is shown through equalization of time-
varying channels with overlap FDE, and with spatial multi-
plexing with FD multi-antenna detectors. Results on the use
of this approach for SC-FDE design are partially exposed in
[51], and extension of results herein to SC-FDMA multi-user
detectors is exposed in [52].

A. Related Work

There is a significant amount of work on iterative equal-
ization which would deserve a survey paper on its own, here
we restrict ourselves to the significant developments in prior
work related to single-tap FDEs, and to EP-based receivers.

1) Iterative Single-Tap FD Receivers: There is a long
research track on frequency domain equalizers, starting from
very low-complexity linear FDE up to non-linear FD turbo
equalizers. Table 1 lists chronological milestones on develop-
ments regarding how interference cancellation with either
decoder or decision (demapper) feedback is used. The posi-
ton of the FDE derived in Section II of this paper is also
shown. The “schedule” column indicates in which manner the
coder/demapping feedback is used by the equalizer.

First FD turbo linear equalizer - interference canceller (FD
LE-IC) was derived using conventional turbo formalism [7],
[8] to yield the extrinsic (EXT) feedback based FD LE-EXTIC
[10]. However, as [33] noted, using a posteriori probability
(APP) based feedback from the decoder yields significant
improvement in turbo detection. Turbo FDE was extended
to FD LE-APPIC [20]. Nevertheless, APP feedback violates
the independence principle of turbo iterative systems [4], so
theoretical background for such structures was absent.

Independently of the emerging turbo equalization literature,
given that time domain (TD) block DFE structures outper-
fomed block LE [34], derivation of non-linear FDE was
of interest. In particular, a hybrid implementation of block
DFE was carried out in [11, 12]. This structure uses a FD
feedforward filter and a TD feedback filterbank, which carried
out symbol-wise, i.e. serial, interference cancellation with hard
decisions. The use of noise prediction in [13], simplified the
computation of hybrid DFE, by forcing the feedback filter
to be the same as the FD LE filter, while the overall structure
remained equivalent to block DFE.

In [14], the frequency domain feedback concept was intro-
duced, and denoted iterative block DFE (IBDIFE). This
structure uses decision feedback in a blocksize-wise, parallel sched-
ule, allowing the use of FFTs over feedback symbol block,
and significantly reducing complexity. Despite its name, this
structure is a LE-IC, with the decision feedback being used for
interference cancellation, and it is not related to the TD block
DFE in [34]. Indeed, the TD block DFE of [34] uses serial
symbol-wise hard decision feedback via a fairly complicated
feedback filterbank, and thus it is unrelated to the linear IC
scheme of [14]. In [15], variations of IBDIFE were evaluated
with hard or soft APP, and TD or FD feedback. It is noted in
[15, 16] that when used with forward error correction, this
structure is equivalent to FD LE-EXTIC.

In [16], probabilistic data association is used to derive a non-
linear FDE for BPSK, through a self-iterated MMSE LE-IC
using APP feedback from previous detections, before comput-
ing extrinsic LLRs for decoding. This structure, and IBDIFE
[14, 15] were later extended to generalized constellations in
[17]. In the latter work, non-linear block FDE (similar

| Reference | Contribution on single-tap FDE | Decoder Feedback | Decision Feedback | Schedule |
|-----------|-------------------------------|-----------------|------------------|----------|
| 1973 [19] | Initially proposed frequency domain equalization | - | - | - |
| 1994 [19] | Revised FDE through the comparison of SC-FDE relative to multicharacter signalling | - | - | - |
| 2001 [10] | Derived the turbo iterated FD LE-IC. Denoted as FD LE-EXTIC. | Extrinsic | - | Parallel/|
| 2002 [11] | A hybrid DFE with a frequency domain feedforward filter and a | | | |
|            | time domain feedback filter is proposed. | - | Hard | -/Serial |
| 2003 [13] | Simplifies hybrid DFE design by using noise prediction. | - | Hard | /-Serial |
| 2005 [14] | Proposed a non-linear receiver with FD feedforward and feedback | | | |
|            | filters, called iterative block DFE (IBDIFE). | - | Hard/APP | -/Parallel |
| 2006 [20] | Proposed a turbo FD MIMO receiver. It used APP estimates from | APP | - | Parallel/|
|            | the decoder instead of extrinsic. Denoted as FD LE-APPIC. | | | |
| 2007 [15] | Compared FDE with FD feedforward and TD/FD feedback filters. | Extrinsic | Hard/APP | Parallel/|
|            | TD/FD are equivalent (parallel schedule). Soft better than hard. | | -/Parallel | |
| 2008 [16] | Proposed a BPSK receiver that is self-iterated with APP estimates | Extrinsic | APP | Par/Par. |
|            | before decoding at each turbo iteration. | | | |
| 2013 [21] | Derived a self-iterated turbo receiver based on GAMP, that exploits | Extrinsic | APP | Par/Par. |
|            | APP estimates at each turbo iteration. | | | |
| 2015 [22] | Equivalence of coded IBDFE to FD LE-EXTIC is shown. | - | - | - |
| This Paper | Proposes a self-iterated FD LE-IC with EP-based feedback. | Extrinsic | Extrinsic | Par/Par. |
complexity to block LE) were evaluated, using APP decision feedback with serial and parallel schedules. These results are then used to derive a single-tap FD self-iterated LE-IC with an initial IC carried out withEXIT feedback from the decoder, followed by a second round of IC carried out with APP feedback from the detector. Here, this structure is denoted as FD SILE-APPIC. Another APP-based iterative FDE is derived via generalized approximate message passing (GAMP) [21].

2) Receivers based on Expectation Propagation: As stated in the introductory paragraphs, EP reignited interest in digital receiver design thanks to a novel type of soft symbol estimates, computed at the demapper, which respects the independence principle of turbo iterative systems, unlike soft APP estimates. EP paradigm has been used for iterative multiple-input multiple-output (MIMO) detection for time domain iterative receiver design thanks to a novel type of soft symbol estimates, and approximate message passing (AMP) algorithms that EP-based extrinsic feedback, and they are shown to outperform turbo-iterations.

B. Contributions and Paper Outline

This paper’s contributions are novel receivers which ensue from an EP-based message passing framework, where transmitted symbols are assumed to belong to the multivariate white Gaussian distributions. These structures use single-tap FD MMSE linear filters, with interference cancellation using the EP-based extrinsic feedback, and they are shown to outperform alternatives from the previous works. Moreover the complexity of these structures have quasi-linear dependence on the block length, unlike cubic or quadratic dependencies of EP-based receivers in prior work. Proposed approach is also compared with approximate message passing (AMP) algorithms that are used in various estimation problems such as compressed sensing.

This approach, which has not been previously used for digital receiver design to the authors’ knowledge, is exposed with the design and analysis of a FDE for quasi-static frequency-selective channels. Then to give a glimpse of the full potential of this framework, more advanced receivers such as an overlap FDE, or a FD MIMO detector are derived and evaluated.

The remainder of this paper is organized as follows. Section II presents SC-FDE receiver design with the proposed message passing framework, and the resulting receiver is analyzed in section III. In section IV the application of this framework is considered for time-varying channel equalization via an overlap FDE, and in section V a multi-antenna spatial multiplexing application is considered. Conclusions are drawn in the end.

C. Notations

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

\( \mathbf{I}_N \) 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. \( e_n \) is the \( N \times 1 \) indicator whose only non-zero entry is \( e_n = 1 \). Operator \( \text{Diag}(\mathbf{u}) \) 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 \( \sigma_{xy} = \text{Cov}[x,y] \) is the covariance. The probability of \( x \) taking a value \( \alpha \) is \( \mathbb{P}[x = \alpha] \), and probability density functions (PDF) are denoted as \( p(.) \). If \( x \) and \( y \) are random vectors, then we define vectors \( \mu_x = \mathbb{E}[x] \) and \( \sigma_x^2 = \text{Var}[x] \), the covariance matrix \( \Sigma_{xy} = \text{Cov}[x,y] \) and we note \( \Sigma_x = \text{Cov}[x,x], \mathcal{CN}(\mu_x, \sigma_x^2) \) denotes the circularly-symmetric complex Gaussian distribution of mean \( \mu_x \) and covariance \( \sigma_x^2 \), and \( \mathcal{B}(p) \) denotes the Bernoulli distribution with a success probability of \( 0 \leq p \leq 1 \).

II. PROPOSED EP-BASED DESIGN FRAMEWORK FOR A BICM SC-FDE SYSTEM

A. System Model

Single-carrier transmission of a block of \( K \) symbols using a bit-interleaved coded modulation (BICM) scheme is considered. In detail, the information block \( \mathbf{b} \in \mathbb{F}_2^{Kd} \) is encoded and then interleaved into a codeword \( \mathbf{d} \in \mathbb{F}_2^K \), with a code rate \( R_c = K_b/K_d \). A memoryless modulator maps \( \mathbf{d} \) into \( \mathbf{x} \in \mathcal{X}^K \), where the constellation \( \mathcal{X} \) has \( M \) elements, and where \( K = K_d/q \), with \( q = \log_2 M \). This constellation is assumed to have a zero mean, and average power \( \sigma_x^2 = 1 \), with equiprobable symbols. This operation associates the \( q \)-word \( \mathbf{d}_k \triangleq [d_{kq}, \ldots, d_{(q+1)k-1}] \) to the symbol \( x_k \), and \( \varphi_j^{-1}(x_k) \) and \( d_{kq} \) are used to refer to \( d_{kq+j} \).

An equivalent baseband circular channel model is considered, including the effects of transceiver modules and the channel propagation. The receiver is assumed to be ideally synchronized in time and frequency and it has perfect channel state information. The received samples are given by

\[
\mathbf{y} = \mathbf{Hx} + \mathbf{w},
\]

where \( \mathbf{H} \in \mathbb{C}^{K \times K} \) is a circulant matrix, generated by \( \mathbf{h} = [h_0, \ldots, h_{L-1}, 0_{K-L-1}] \), the impulse response extended with \( K - L \) zeros, \( L < K \) being the channel spread. Unlike the additive white Gaussian noise (AWGN) model in...
here, a coloured and correlated noise $w$ is considered to capture the impact of eventual interfering signals, with $w \sim \mathcal{CN}(0_K, \Sigma_w)$. This channel model is applicable to different SC-FDE implementations such as the cyclic prefix (CP) SC-FDE, or the zero-padded (ZP) SC-FDE among others.

The normalized $K$-DFT matrix is given by its elements $[F_K]_{m,n} = \exp(-2j\pi mn/K)/\sqrt{K}$, such that $F_K F_K^H = I_K$. Then the equivalent frequency domain transmission model is

$$y = H x + w,$$

with $x = F_K x$, $y = F_K y$, $w = F_K w$, and $H = F_K H F_K^H = \text{Diag}(h_k)$ with the channel frequency response being

$$h_k = \sum_{l=0}^{L-1} h_l \exp(-2j\pi kl/K), \quad k = 1, \ldots, K,$$

and $w \sim \mathcal{CN}(0_K, \Sigma_w)$, with $\Sigma_w = F_K \Sigma_w F_K^H$ is the noise covariance matrix in the FD. To keep the receiver complexity low, the non-diagonal elements of $\Sigma_w$ are ignored, hence use-cases involving interference with non-negligible inter-carrier correlations are out of scope.

The remainder of this section covers the approximation of the posterior probability density function of transmitted symbols, by using an EP-based message passing algorithm. In particular, symbol variables $x$ are assumed to belong to a multivariate white Gaussian distribution, of the form $\mathcal{CN}(\bar{x}, \sigma I_K)$, where the reliability of symbol estimates $\bar{x}$ is given by a scalar $\bar{v}$. The resulting approximate distribution is shown to yield a novel iterative single-tap FD LE-IC.

### B. Factor Graph Model

The joint posterior probability density function (PDF) of data bits, given FD observations, is $p(b, d, x|y)$. The optimal joint MAP receiver operating on FD observations resolves the criterion $\hat{b} = \max_b p(b|y)$. Assuming i.i.d. information bits, the posterior PDF is factorized as

$$p(b, d, x|y) \propto p(y|x) p(x|d) p(d|b).$$

This density is further factorized by using the memoryless mapping $p(x|d) = \prod_{k=0}^{K-1} p(x_k|d_k)$, and the independence assumption in BICM encoding $p(d|b) = \prod_{k=0}^{K-1} \prod_{j=0}^{L-1} p(d_{kj})$, where the probability mass function (PMF) $p(d_{kj}) \triangleq p(d_{kj}|b)$ is seen as a Bernoulli-distributed prior constraint provided by the decoder, from the receiver’s point of view.

As we focus on iterative detection and decoding for a given transmission, we will focus on the posterior for the estimation of variables $d_{kj}$ and $x_k$, and remove $b$ from notations. Hence (4) can be factorized as

$$p(d, x|y) \propto p(y|x) \prod_{k=0}^{K-1} p(x_k|d_k) \prod_{j=0}^{L-1} p(d_{kj}).$$

This process is iteratively carried out by a message-passing based detection and decoding algorithm operating on the variables nodes (VN) $x_k$ and $d_{kj}$ by using constraints imposed by factor nodes (FN) corresponding to the factorization of the posterior PDF in (5). The equalization (EQU) FN resolves the multipath channel constraints with

$$f_{\text{EQU}}(x) \triangleq p(y|x) \propto e^{-\frac{1}{2} \sum_{k,j} y_{kj}^2 + 2Re(\sum_{k} y_{k}^H H_{kj}^F x)},$$

where the dependence on $x$ is omitted, as the FD observations are unchanged during iterative detection. Demapper (DEM) FN handles the mapping constraints with

$$f_{\text{DEM}}(x_k, d_k) \triangleq p(x_k|d_k) \propto \prod_{j=0}^{L-1} \delta(d_{kj} - \varphi^{-1}(x_k)),$$

where $\delta$ is the Dirac delta function, and finally, channel coding constraints are handled by

$$f_{\text{DEC}}(d_k) \triangleq \prod_{j=0}^{L-1} p(d_{kj}).$$

The considered BICM SC-FDE system factor graph is given by Fig. 1. Note that, unlike the finite-impulse response receiver factor graph in [30], EQU FN impacts all transmitted symbols.

### C. Proposed EP-based Message Passing Framework with White Gaussian Distributions

Expectation propagation, extends belief propagation as a message passing algorithm by assuming the variable nodes to have PDFs belonging to the exponential family [24]. This results in exchanged messages to be depicted by tractable distributions, which allows for the iterative computing of a fully-factorized approximation of challenging PDFs such as $p(d, x|y)$. Resulting approximation can then be marginalized on variables of interest, to yield the desired estimates.

Updates involving a FN $F$, connected to variable nodes $V$ are as follows. Messages exchanged between VN $v_i$, the $i$th component of $v$, and $F$ are given by

$$m_{v\rightarrow F}(v_i) \triangleq \prod_{v \neq v_i} m_{v\rightarrow v_i}(v_i),$$

$$m_{F\rightarrow v_i}(v_i) \triangleq \text{proj}_{Q_{v_i}} [q_{F}(v_i)]/m_{v\rightarrow F}(v_i),$$

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

$$q_{F}(v_i) \triangleq \int_{v \setminus v_i} f_F(v) \prod_{v \setminus v} m_{v\rightarrow F}(v_j) dv^{\setminus_i},$$

where $v^{\setminus_i}$ are VNs without $v_i$. [24]. The projection operation for exponential families is equivalent to moment matching, which simplifies the computation of messages [24].

For the proposed framework, our simplifying assumption is that VNs $x$ lie in multivariate white Gaussian distributions. Hence, a message involving these VNs is fully characterized by a vector mean and a scalar variance. On the other hand
$d_{k,j}$ follow a Bernoulli distribution (which is included in the exponential family), whose messages are characterized by binary log-likelihood ratios (LLRs), as in the conventional belief propagation algorithm.

D. Derivation of Exchanged Messages

In this subsection, the framework above is applied to the considered factor graph, by first, defining exchanged messages, and then computing their characterizing parameters.

The messages arriving on the VN $x_k$ are Gaussians with

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

where means are dependent on $k$ and variances are static. Oppositely, the messages arriving on the VN $d_{k,j}$ are Bernoullis

$$m_{\text{DEM} \to d}(d_{k,j}) \propto B\left(p^d_d\right), m_{\text{EQU} \to d}(d_{k,j}) \propto B\left(p^e_d\right).$$

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

$$L(d_j^d) \equiv \ln \frac{p_d}{1 - p_d} = \ln \frac{1 - p_d}{p_d}.\quad (15)$$

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 SISO receiver’s perspective, i.e., $L_a(d_{k,j})$, $L_e(d_{k,j})$ respectively characterize $m_{\text{DEM} \to d}(d_{k,j})$ and $m_{\text{EQU} \to d}(d_{k,j})$. Fig. 2 illustrates a conventional view of the receiver with the quantities above.

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. (9), $m_{e \to f}(v_i) = m_{g \to v}(v_i)$, for all VN $v_i$, and FN $F, G, F \neq G$ they are connected to.

1) Messages from DEC to DEM: DEC FN is assumed to be a SISO channel decoder, that generates prior information $L_a(d)$ to DEM, when extrinsic LLRs $L_e(d)$ is given to it by DEM.

Using these prior LLRs with the mapping constraints in (7), the prior PMF on $x_k = \alpha$, is

$$P_k(\alpha) \propto \prod_{j=0}^{J-1} e^{-\varphi^{-1}(\alpha)L_a(d_{k,j})}, \forall \alpha \in \mathcal{X}.\quad (16)$$

This is a categorical PMF corresponding to the marginal of $f_{\text{DEM}}(x_k, d_k)m_{\text{DEM} \to \text{DEC}}(d_k)$ on $x_k$ [25], used hereafter to compute approximate marginals $q_{\text{DEM}}(x_k)$ and $q_{\text{DEM}}(d_{k,j})$.

2) Messages from DEM to EQU: An approximate posterior on the variable node $x_k$ is computed at the demapper, using eq. (11), with

$$q_{\text{DEM}}(x_k) = \sum_{d_k} f_{\text{DEM}}(x_k, d_k)m_{x \to \text{DEM}}(x_k)$$
$$\prod_{j=0}^{J-1} m_{d \to \text{DEM}}(d_{k,j}).\quad (17)$$

This is a posterior categorical PMF on $x_k = \alpha$, given by eqs. (12) and (16), denoted as

$$D_k(\alpha) \propto \exp \left(-|\alpha - \bar{x}_k^d|/\nu^e_\alpha\right)P_k(\alpha), \forall \alpha \in \mathcal{X}.\quad (18)$$

For computing messages towards EQU via eq. (10), the posterior PMF is projected into a Gaussian distribution through moment matching. The mean and the variance of $D_k$ are

$$\mu^d_k \equiv \mathbb{E}_{D_k}[x_k] = \sum_{\alpha \in \mathcal{X}} \alpha D_k(\alpha),$$
$$\gamma^d_k \equiv \text{Var}_{D_k}[x_k] = \sum_{\alpha \in \mathcal{X}} \alpha^2 D_k(\alpha) - |\mu^d_k|^2.\quad (19)$$

The result of the projection on $x_k$ is $\mathcal{CN}(\mu^d_k, \gamma^d_k)$, using moment matching [24, eq. (19)], where means are matched, but the variance needs to satisfy, $\forall k$, $\gamma^d_k = \gamma^d$. Working with white Gaussians creates an overdetermined constraint on the reliability of estimates without any exact solutions. An approximate solution, given by the ordinary least-squares, coincides with the sample average

$$\gamma^d_k \triangleq K^{-1}\sum_{k=0}^{K-1}\gamma^d_k.\quad (20)$$

Then $m_{\text{DEM} \to x}(x_k)$ is computed as in (10), by using a Gaussian division [23], which yields

$$x_k^e = \frac{\mu^d_k v^e - x_k^e \gamma_k^d}{v^e - \gamma_k^d},$$
$$x_k^d = \frac{\nu^d_k v^e - x_k^d \gamma_k^d}{v^e - \gamma_k^d}.\quad (21)$$

The major novelty in using EP lies in this expression; the computation of an extrinsic feedback to the equalizer from the demapper. Attempting this with categorical distributions, as in BP, would completely remove $m_{\text{DEM} \to x}(x_k)$, and the extrinsic “feedback” to EQU would simply become prior PMF $P_k$ [25], which results in a receiver equivalent to FD LE-EXITIC [10].

The feedback produced by EP is erroneous, if the denominator in eq. (21) is negative, which may be caused by conflicts among the equalizer’s output and the mapping constraints. In this case [23] replaces the concerned $(x_k^e, v_k^e)$ with their values from a previous iteration, and [25] uses posteriors $(\mu^d_k, \gamma^d_k)$ instead. From experimentation not exposed here, the latter case is found to be more advantageous. However, unlike these references, the use of static variances greatly reduces the occurrence of $v^e \leq \gamma^d$, and if it occurs, we use $\mu_k$ and $\gamma^d$ instead.

EP message passing minimizes local divergences (on marginal posteriors) in order to minimize a global divergence (full posterior). Thus, it does not guarantee convergence and it might lock on undesirable fixed points. As in [24, eq. (17)], a feature-based damping heuristic is used

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

Fig. 2. Factor nodes shown as an iterative BICM receiver.
with tuning parameter $0 \leq \beta \leq 1$. \cite{41} uses a linear smoother between DEM’s and DEC’s extrinsic estimates, which is inefficient in a self-iterated EQU-DEM schedule. Hence it is here extended to linearly smooth messages across self-iterations

$$v^{\text{d(next)}} = (1 - \beta) v^* + \beta v^{\text{d(prev)}},$$  
$$x^{\text{d(next)}}_k = (1 - \beta) x^*_k + \beta x_k^{\text{d(prev)}}. \tag{23}$$

Although both approaches are expected to asymptotically lead to similar limits, with numerical experimentation, the feature-based approach often converges faster for the same $\beta$. However, inversions used in this approach cause numerical issues in some configurations, which makes the linear damping more preferable.

3) Messages from EQU to DEM: Approximate posterior on the VN $x_k$ (eq. \ref{11}) is given by

$$q_{\text{EQU}}(x_k) = \int_{x_k} q_{\text{EQU}}(x) \prod_{k' = 0}^{K-1} m_{x \rightarrow \text{EQU}}(x_{k'}) \, dx_k \tag{24}$$

Denoting the integrand above as $\mathcal{C}N(\mu^e, \Gamma^e)$, and using eq. \ref{6}, we have

$$\Gamma^e = (\Gamma_{1k}/v^d + \mathcal{F}_H H \Sigma_{1} \mathcal{F}_K)^{-1},$$  
$$\mu^e = \Gamma^e (\mu_1^d/v^d + \mathcal{F}_H H \Sigma_{1} \mathcal{F}_K) \xi,$$  
$$x^d = [x^d_0, \ldots, x^d_{K-1}],$$

Using some matrix algebra, and Woodbury’s identity on $\Gamma^e$, the variance $\gamma^e_k$ and the mean $\mu^e_k$ of the marginalized PDF $q_{\text{EQU}}(x_k)$ are given by

$$\gamma^e_k = \xi^{-1} h_k/(\gamma^2_k + v^d |h_k|^2),$$  
$$\mu^e_k = \xi^{-1} h_k + v^d \xi f_\kappa^e(y_k - h_k x^d_k), \tag{26}$$

where $\Gamma^e$ and $x^d_k$ are the FD spectrum of respectively $\mu^e_k$ and $x^d_k$. Parameters $f_\kappa^e$ and $\xi$ follow

$$f_\kappa = \xi^{-1} h_k/(\gamma^2_k + v^d |h_k|^2),$$  
$$\xi = \xi^{-1} \sum_{k=0}^{K-1} \gamma_k^{-2} + \xi^{-2} |h_k|^2. \tag{27}$$

Noting that $\gamma^e_k$ does not depend on $k$, $q_{\text{EQU}}(x_k)$ already belongs to the family of white multivariate Gaussians, the projection operation in \ref{10} has no effect, $\gamma^e_k = \gamma^e_k$, $\forall k$. Hence the Gaussian division of $q_{\text{EQU}}(x_k)$ by $m_{x \rightarrow \text{EQU}}(x_k)$ is readily carried out to compute parameters of $m_{\text{EQU} \rightarrow x}(x_k)$

$$x^e_k = x^d_k + f_\kappa^e (y_k - h_k x^d_k),$$  
$$v^e = \xi^{-1} - v^d. \tag{29}$$

Note that these expressions result in the conventional MMSE FD LE-IC structure, with interference cancellation being carried out using extrinsic EP feedback $x^d_k$.

4) Messages from DEM to DEC: The demapper computes an approximate posterior on the VN $d_{k,j}$ using eq. \ref{11}

$$q_{\text{DEM}}(d_k) = \sum_{x_k \in \mathcal{X}} f_{\text{DEM}}(x_k, d_k) m_{x \rightarrow \text{DEM}}(x_k) \prod_{j=0}^{q-1} m_{d \rightarrow \text{DEM}}(d_{k,j}). \tag{31}$$

The marginalization of this posterior on $d_{k,0}, \ldots, d_{k,q-1}$ \cite{25}, and the division in \ref{10} is directly carried out with bit LLRs

$$L_e(d_{k,j}) = \ln \frac{\sum_{\alpha \in \mathcal{X}} \mathcal{P}_k(\alpha)}{\sum_{\alpha \in \mathcal{X}} \mathcal{P}_k(\alpha)} - L_a(d_{k,j}), \tag{32}$$

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

---

**Algorithm 1 Proposed FD SILE-EPIC Receiver**

**Input** $\mathbf{y}, H, \sigma_w^2$

1. Initialize decoder with $L^0_a(d_k) = 0, \forall k$.
2. for $\tau = 0$ to $T$
   3. Initialize eq. with $\hat{x}_k^{\tau(0)} = 0, \forall k$ and $\sigma_v^{\tau(0)} = +\infty$.
   4. Use decoder’s $L^\tau_a(d_k)$ to compute $P_k^\tau$ via \ref{10}, $\forall k$.
   5. for $s = 0$ to $S$
      6. Use \ref{18} to update demapper posteriors, $\forall k$.
      7. Generate soft feedback using \ref{21}, \ref{23}.
      8. Compute $\xi^{(\tau,s)}$ using \ref{28}, and $\sigma_v^{(\tau,s)}$ using \ref{30}.
      9. Equalize using \ref{27} and \ref{29}, for $k = 0, \ldots, K-1$.
   10. end for
11. Provide extrinsic outputs $L^\tau_e(d_k)$ to the decoder using \ref{12}, in order to obtain next priors $L^\tau+1_a(d_k), \forall k$.
12. end for

---

E. Proposed FD Self-Iterated LE-EPIC Receiver

As the considered factor graph has cycles, it is not possible to derive a receiver algorithm with only the messages exchanged over it; a schedule for coordinating the update of variable and factor nodes is needed.

To keep the equalization complexity reasonable, a parallel scheduling across variables nodes $x_k$ is considered, in line with conventional FD LE or block LE receivers. Note that the use of a serial schedule would yield a DFE-like structure \cite{30}.

To fully exploit the benefits of the feedback computed by the demapper, a flexible double-loop FDE structure is proposed. The first loop refers to the exchange of extrinsic information between the decoder and the demapper in a turbo-iteration (TI), while the second loop refers to the message exchange in a self-iteration (SI) between the demapper and the equalizer.

Each TI $\tau = 0, \ldots, T$ consists of $S_{\tau}$ SIs (may depend on $\tau$), where EQU and DEM factor nodes are updated in parallel schedule, for $s = 0, \ldots, S_{\tau}$, and then the DEC factor nodes are updated with a selected SISO decoder. To clarify this, Algorithm 1 below explicitly describes the proposed scheduling, where involved quantities are indexed by $(\tau,s)$ in the superscript.

The iterative FDE derived in this section, by applying the EP framework in the FD, with the family of white Gaussian distributions, yields the low-complexity single-tap receiver structure shown in Fig. 3. In the next section, the behaviour of this receiver will be assessed with achievable rate analysis and comparisons with structures from the prior work.

---

III. EP-based SC-FDE Performance Evaluation

Here, the receiver derived in the previous section is used with a fixed number of SI $S_{\tau} = S$ per TI, and it is referred to...
as the $S$-self-iterated FD LE-IC with EP (FD $S$-SILE-EPIC).

A. Asymptotic Analysis

In order to evaluate asymptotic behaviour ($\tau \to \infty$) of the proposed receiver, extrinsic information transfer (EXIT) analysis is used [42]. This consists in characterizing the behaviour of iterative SISO components with single-parameter transfer functions, by tracking the extrinsic mutual information (MI) exchanges. The behaviour of a SISO receiver is represented by the transfer function $I_E = T_{R}(I_A, h, \Sigma_w)$ which depends on the channel parameters, with $I_A$ and $I_E$ being the MI between coded bits and respectively the a priori and extrinsic LLRs of the module.

One primordial use of EXIT analysis is performance prediction through evaluation of MI evolution. However, this entails strong assumption on the distribution of the prior LLRs of the SISO module, which cannot be met for most receivers other than MAP detectors. This issue can cause these transfer functions to be too optimistic in some cases, providing only an upper-bound on asymptotic performance. In this regard, the accuracy of transfer functions is assessed by comparing it with actual MI trajectories, obtained with finite-length simulations.

In Fig. 4, EXIT charts of the proposed receiver, for $S = 0, 1, 3, 5, 10$, using a fixed linear damping (see eq. (23)), with $\beta = 0.75$, is provided in solid curves, within the highly selective Proakis C channel, $h = [1, 2, 3, 2, 1]/\sqrt{19}$, using the Gray-mapped 8-PSK constellation. Self-iterations are seen to significantly improve the MI for high $I_A$, which indicates a boosted convergence speed and an improved achievable rate. However, improvements for $I_A = 0$ is relatively small, thus, the finite-length performance improvement will strongly depend on the EXIT chart of the decoder. In particular, for non-optimized but powerful turbo-like codes which have near-flat EXIT curves, improvement on the decoding threshold will be limited. However with a properly designed code, significant improvement would be possible.

This figure also shows the reverse transfer curve of the recursive systematic convolutional (RSC), code $[1, 5/7]_8$. More-over, in dashed curves, the finite-length MI trajectories of this receiver with data blocks of length $K_s = 768$ bits, using this channel decoder is plotted. The trajectories of the proposed EP-based receiver appears to follow the predicted transfer function fairly well, despite the short packet length, unlike the APP-based receivers as observed in [40]. This suggests that this receiver’s EXIT analysis reflects its practical behaviour.

Another significant use of EXIT analysis is the evaluation of achievable rates (i.e. spectral efficiency) of the receiver. These values are numerically obtained using the area theorem of EXIT charts [43], and when considering the MAP detector, these rates constitute an accurate approximation of the channel symmetric information rate [44], the highest possible transmission rate for practical constellations, without channel knowledge at the transmitter. Achievable rates of the FD LE-EXTIC and the proposed receiver are given in Fig. 5 for the Proakis C channel with 8-PSK and 64-QAM constellations. The Gaussian capacity of this channel, without transmit power optimization, is also plotted in dashed lines, it is computed using eq. (1) with the vector-input AWGN channel capacity. Channel SIR with 8-PSK is given by the MAP detector curve in 8-PSK, but it is not plotted with 64-QAM due to the excessive computational resources it requires [44]. An exponential feature-based damping (see eq. (22)) with $\beta = 0.7 \times 0.9^s$ is used for 8-PSK, whereas a fixed linear damping (see eq. (23)) with $\beta = 0.8$ is used for 64-QAM.

For 8-PSK, while the conventional FD LE-EXTIC [10] follows the SIR limit within 0.5 dB up to 0.75 bits/s/Hz, proposed EP-based self-iterations increase this range up to 2 bits/s/Hz. In the 64-QAM case, FD LE-EXTIC follows the channel capacity within 1 dB up to 1 bit/s/Hz and 3.33 bits/s/Hz becomes achievable with 20 SI. For a rate-1/2 coded usage, the proposed receiver with $s \to +\infty$ brings over 3.9 dB and 10.7 dB improvement, over the conventional turbo FD LE in this channel, for respectively 8-PSK and 64-QAM constellations. These rates are achievable with properly designed coding schemes.

B. Comparison with Single Tap FDE in Prior Work

In this paragraph, observations in the previous section are completed with finite-length results within the same channel with a RSC code with soft MAP decoder. Block error rate (BLER) is obtained by Monte-Carlo simulations, with 30000 sent packets per point. Unlike in asymptotic analysis, here we use dynamic damping that also depends on turbo-iterations, $\tau$, and accelerates convergence. A feature-based damping with $\beta_{\tau,s} = 0.7 \times 0.9^{s+\tau}$ is used for 8-PSK, and a hybrid damping, consisting of a linear smoothing in the first TI, and feature-based damping afterwards, is applied with $\beta_{\tau,s} = 0.85^{1+\tau}$, for 64-QAM. Several single-tap FD equalizers are compared to our proposal in Fig. 6 the conventional linear equalizer (LE-EXTIC), the LE-IC with APP feedback [20], [33] (LE-APPIC), and the self-iterated LE-IC of [12] (SILE-APPIC). The equalization complexity of these receivers is of the same order of computational complexity of $K \log_2 K$ at a given SI and TI, with slight differences underlying in the feedback computation.

Fig. 4. EXIT curves and finite-length average MI trajectories of the proposed equalizer with 8-PSK in Proakis C channel at $E_b/N_0 = 15$ dB.
Results show in Fig. 6 show that our proposal brings significant improvement on the decoding threshold, that grows with the number of SIs, at all TIs. On the contrary, multiple SIs with APP feedback degrades this threshold (not shown here due to lack of space). Without TI, 3 SIs bring respectively 9 dB and 6 dB gains for 8-PSK and 64-QAM, compared to LE-EXTIC, at BLER = 10^-1. Performance in 64-QAM is limited at low PER without TIs, but our proposal with a single TI and 3 SIs reaches PER the prior work reach with 6 TIs, e.g. with six times lower decoding complexity. Besides, asymptotically (6 TIs), SIs with EP bring over 8 dB gain with respect to SILE-APPIC, and about 5 dB gain over LE-APPIC, for 64-QAM, at BLER = 10^-2. Compared to FD LE-EXTIC, 3 SIs bring around 4 dB and 11.5 dB gain, respectively for 8-PSK and 64-QAM, which is close to the 1/2-rate gains observed in the asymptotic analysis above.

These results encourage replacing TIs with SIs as demapping complexity is often insignificant relative to decoding.

C. Comparison with EP-based Receivers in Prior Work

There are numerous emerging EP-based receivers in the literature, as stressed in the introduction, and in this section the proposed FDE is compared with self-iterated time-domain block (SIBLE-EPIC, denoted nuBEP in [29]) and filter (SIFLE-EPIC, denoted EP-F in [29]) receivers and to the single-tap FD receiver, D-EP, in [37]. The proposed receiver is not compared to the exact FD receiver, J-EP in [36], [37], as it is equivalent to the SIBLE-EPIC with a single SI, without damping, making it sub-optimal compared to the SIBLE-EPIC. The block receiver in [28], is neither included in the comparison, as it is a sub-optimal block receiver which ignores prior information from the decoder at each SI (but a comparison is available in [31]).

In Fig. 7 the bit error rate (BER) of the proposed receiver is compared with alternatives listed above. We consider 8-PSK constellation, and the low-density parity check (LDPC) coded Proakis C scenario from [29]. The regular (3, 6) LDPC code is obtained by Progressive-Edge Growth (PEG) algorithm, and the decoder uses BP algorithm up to 100 iterations. The FD receiver, D-EP, cannot decode in Proakis-C channel, up to very high signal to noise ratios due to its sensitivity to channel nulls [37, eq. (48)]. Our FD proposal is seen to perform nearly as good as the TD EP-based receivers, with an order of computational complexity of (S + 1)K log_2 K instead of...
Fig. 7. BER comparison in Proakis C with 8-PSK, $K_d = 4096$ and rate-1/2 regular $(3, 6)$ LDPC code.

Fig. 8. SI LE-IC and FD SILE-IC in Proakis C with LDPC coded 16-QAM, with 5 turbo iterations.

Fig. 9. Performance complexity trade-off for self-iterations in LDPC coded Proakis C.

3$LK^2$ (SIBLE-EPIC, 2 SIs) and or $27KL^2$ (SIFLE-EPIC, 2 SIs). For $\tau = 5$, block and filter TD receivers have around 0.2 dB gain over FD 3-SILE EPIC, but they are respectively around 500 and 16 times more complex.

Another LDPC-coded scenario in the Proakis C, with 16-QAM and with rate 1/2 and 3/4 encoding over $K_b = 2048$ bits is reported in Fig. 8. All receiver use feature-based damping with the optimized parameter in [29], i.e. $\beta = \min(0.3, 1 - e^{\tau/1.5}/10)$. The regular $(3, 12)$ LDPC code is also obtained by the PEG algorithm. In the rate-1/2 case, the proposed FDE is lower-bounded in BER by the block receiver, and following one SI, the difference between FD SILE-EPIC and SIFLE-EPIC is negligible. For the high rate case, at the right side of the figure, filter receiver’s performance is over 1 dB worse for BER $< 10^{-3}$, and although SIBLE-EPIC still has a better decoding threshold, it recovers less diversity than the proposed FD SILE-EPIC. This phenomenon should not be surprising, as exact receivers can be more prone to error propagation when decoder provides erroneous feedback, as also observed in filter receivers [45].

These error rate results are completed with detailed computational complexity estimations in Fig. 9. This is evaluated with the number of multiply and accumulate units required to implement the receiver, estimated by the number of real additions and multiplications, amounting to half a floating point operation (0.5 FLOPs) each. Complexity is plotted versus the required bit SNR to decode transmitted blocks with BLER $= 10^{-2}$, for $\tau = 0, \ldots, 5$. These FLOP-counts also include the decoder complexity, which is considerably higher than equalizer complexity. The proposed receiver performs overall efficient, both complexity and energy-wise, compared to the SIBLE-EPIC, with respectively 2.5, 4 and 5.4 times lower complexity for $S = 0, 1$ and 2 in the rate 3/4 case, and
with respectively 2, 3.1 and 4.1 times lower complexity for the rate 1/2 case. This ratio is around ten times bigger, if the decoding complexity is not accounted for.

D. On the Impact of Imperfect Channel Estimation

In this section, the performance of the proposed FD SILE-EPIC with imperfect channel estimates is evaluated. A mismatched receiver is considered to operate on an channel estimate \( \hat{h} \), corrupted by Gaussian noise \( \nu \) whose variance \( \sigma_v^2 \) was selected using the model \( \sigma_v^2 = \sigma_w^2 / (K_P \sigma_\nu^2) \), where \( K_P \geq L \) is the number of pilot symbols that would have been used for channel estimation in a complete receiver. We assume transmission of 8-PSK blocks with \( K = 256 \) symbols and the quality of channel estimate is assessed via the overhead, defined by the ratio \( K_P / K \).

Figure 10 illustrates the behaviour of BER versus overhead, and the proposed EP-based receiver is shown to be robust to estimation errors. Indeed, for a target BER of \( 10^{-2} \), a significant reduction of overhead is achieved with EP-based self-iterations; while baseline FDE with 2 TIs requires around 19% overhead for channel estimation, using our proposal, one turbo and one self iteration requires only 8% overhead and one self and 2 TIs requires 4%. Thus PHY data frames with shorter number of pilot symbols could be designed to increase spectral efficiency.

E. Comparison with work on Approximate Message Passing

AMP is a commonly used technique in signal processing fields such as compressed sensing or data classification, which is based on belief propagation, often with Gaussian approximation for tractable MMSE estimation. In particular, Generalized AMP (GAMP) is adapted for linear probabilistic models as in Eq. (2) [38], however it is designed for fixed priors and cannot be directly applied in the context of turbo detection. In [21], Guo et al. reworked GAMP for turbo-equalization by generating extrinsic outputs (GAMP-ext), and derived a self-iterated FDE.

Considering that GAMP algorithm is derived using loopy belief propagation, a recent improved extension based on EP is Vector AMP (VAMP) [39]. In order to draw parallel’s between AMP-based algorithms and the proposed EP-based framework, similarly to [21], VAMP algorithm can be tweaked to operate with extrinsic outputs, and extended to circularly complex Gaussian distributions to be applied on frequency domain observations (VAMP-ext).

Derivation details of VAMP-ext are not given due to lack of space, and it results in a similar algorithm to FD SILE-EPIC, but with notable differences in damping and convergence heuristics. In Figure 11 FDEs based on AMP are compared with the proposed FD SILE-EPIC with both linear (“-lin”, see eq. (23)) and feature-based damping (“-feat”, see eq. (22)). Numerical results indicate that our original proposal FD SILE-EPIC converges to further lower error rates than AMP-based alternative, with over 1 dB gain on GAMP-ext, and over 0.6 dB on VAMP-ext. These results show that AMP-based algorithms themselves are not well-adapted to turbo-detection use-case, and that it is preferable to address such systems using the founding theory of EP.

F. Conclusion on EP-based SC-FDE

Finite-length error rate performance and the asymptotic analysis show that the proposed SC-FDE receiver, obtained by the considered EP-based message passing framework in section II outperforms similar alternative receivers (single-tap FDE), and performs almost identically to the exact TD receivers while having a significantly lower computational complexity.

EP with multivariate white Gaussian distributions is exposed in this elementary SC-FDE system, to improve readability and to simplify performance analysis. In the following, we give an overview of application of this framework to more complex communication systems.

IV. Application to Time-Varying Channel Equalization: Overlap FDE

A. System Model

A notable issue of FDE is its inability to mitigate time-varying channels whose coherence time is shorter than the processing block duration. In this case the FD channel matrix is no longer diagonal, and inter-carrier interference is generated. Overlap FDE is a possible approach for mitigating problems above, without significantly increasing the receiver complexity.

This technique consists in using \( N \)-point FFTs, with \( N < K \), to carry out baseband processing, on virtual overlapping sub-blocks of received samples [46], [47]. This strategy inherently generates IBI between sub-blocks, which is mitigated either by selecting an appropriate sub-block length \( N \), or by using additional signal processing. Some recent usage examples include its usage with faster-than-Nyquist signalling [48], and with doubly selective channels [49]. In this section, various EP-based overlap FDE receivers are derived and evaluated.
where $H$ and $N$ are overlapping between two successive sub-blocks, as shown in Fig. 12 and by extracting the remaining $N_d = N - N_l - N_r$ symbols, this procedure is repeated for $N_r = \lfloor K/N_d \rfloor$ sub-blocks in parallel. For extending this scheme to use the proposed EP-based framework from the previous sections; one could implement each $N_c$ equalizer of length $N$ using FD SILE-EPIC. Hence each sub-equalizer would have its own self-iteration loops, and independently evolving estimate variances. But as BICM is used across all sub-blocks, differences of estimate variances between sub-blocks is small, hence, for simplicity, all the sub-equalizers (FD SILE-EPIC) are assumed to use a common SI loop, with the common output variance denoted $\sigma^2 \equiv N_b \sum_{n=1}^{v} v_n$, where $v_n$ is the $n$th sub-equalizer’s output variance, and the common feedback variance is $\sigma^2$.

We denote this overlap FDE scheme, no-interference (NI), its performance at a given signal-to-noise ratio (SNR) can be close to that of SC-FDE with $K$-point FFT, if $N$, $N_l$ and $N_r$ are sufficient to remove all residual IBI over $N_d$ extracted symbols [50]. However, for moderately or highly selective channels, the IBI spread can be very large, requiring $N \gg N_d$. Otherwise, residual IBI is present, and causes detection errors, whose occurrence increase with the SNR, due to interference enhancement caused by mismatched filter weights.

**B. Conventional Overlap FDE with “No-Interference”**

Overlap FDE, also called FDE with overlap-and-save or overlap-and-cut, carries out a linear deconvolution with multiple circular convolutions. Given a signal block $v \in \mathbb{C}^K$, its $N$-point sub-blocks are denoted $v_k = [v_k, \ldots, v_{k+N-1}]^T$, with $v_k = 0$, for all $k < 0$ or $k \geq K$. SC-FDE model with sub-blocks is written as

$$\tilde{y}_k = H_k \tilde{x}_k + G_k (\tilde{x}_{k-N} - \tilde{x}_k) + \tilde{w}_k, \quad (33)$$

where $H_k$ is a $N \times N$ circular channel matrix as in eq. (2), and $G_k$ is an $N \times N$ matrix, whose $L-1$ upper diagonals are equal to those of $H_k$, and other elements are zeros. Unlike the channel model in eq. (2), here the channel may quasi-statically vary between sub-blocks. Hence with a small enough $N$, a time-varying frequency selective channel can be approximated by this model.

SC-FDE is used on sub-blocks, by ignoring the IBI term, and $N_l$ symbols from the head and $N_r$ symbols from the tail of the equalized sub-block are thrown away. $N_l + N_r$ symbols are overlapping between two successive sub-blocks, as shown in Fig. 12 and by extracting the remaining $N_d = N - N_l - N_r$ symbols, this procedure is repeated for $N_r = \lfloor K/N_d \rfloor$ sub-blocks in parallel. For extending this scheme to use the proposed EP-based framework from the previous sections; one could implement each $N_c$ equalizer of length $N$ using FD SILE-EPIC. Hence each sub-equalizer would have its own self-iteration loops, and independently evolving estimate variances. But as BICM is used across all sub-blocks, differences of estimate variances between sub-blocks is small, hence, for simplicity, all the sub-equalizers (FD SILE-EPIC) are assumed to use a common SI loop, with the common output variance denoted $\sigma^2 \equiv N_b \sum_{n=1}^{v} v_n$, where $v_n$ is the $n$th sub-equalizer’s output variance, and the common feedback variance is $\sigma^2$.

We denote this overlap FDE scheme, no-interference (NI), its performance at a given signal-to-noise ratio (SNR) can be close to that of SC-FDE with $K$-point FFT, if $N$, $N_l$ and $N_r$ are sufficient to remove all residual IBI over $N_d$ extracted symbols [50]. However, for moderately or highly selective channels, the IBI spread can be very large, requiring $N \gg N_d$. Otherwise, residual IBI is present, and causes detection errors, whose occurrence increase with the SNR, due to interference enhancement caused by mismatched filter weights.

**C. Overlap FDE with Interference Rejection**

Interference enhancement caused by overlap FDE NI causes prohibitive constraints for selecting $N$, $N_l$ and $N_r$, in order to avoid residual IBI. Moreover, for channel with severe spectral nulls, IBI spread can be as large as $N$, making overlap FDE unusable with any parameters. If the channel is also time-varying, coherent time constraints on $N$ are imposed, which may cause overlap FDE NI to yield no viable solution.

In this paragraph, the interference rejection (IR) strategy, which mitigates interference enhancement, is exposed, by designing filters that account for the presence of IBI. The equivalent noise which also includes the IBI is

$$\tilde{w}_k' \triangleq G_k (\tilde{x}_{k-N} - \tilde{x}_k) + \tilde{w}_k. \quad (34)$$

Considering the noise model (see eq. (2)) used in the FDE design, in section 11 one can compute a SC-FDE equalizer [27]-[30] using the equivalent noise covariance

$$\Sigma_{\tilde{w}_k'} = \Sigma_{\tilde{w}_k} + 2\sigma^2 F_N G_k H_N^H H_N^H, \quad (35)$$

assuming i.i.d. transmitted symbols. The equalizer neglects noise correlations between different subcarriers, but accounts for the FD coloured noise with diagonals of matrix (35). IR was applied using the whitened covariance of the IBI in [50], however using a coloured representation, as in this paper, was shown significantly improved performance [51].

---

**Fig. 11.** Comparison of iterative receivers based on AMP with the proposed FD SILE EPIC in Proakis C, with RSC [1, 5/7]8 coded 8-PSK and with static BER-optimized damping for each receiver, at each iteration and SNR.

**Fig. 12.** Overlap FDE processing scheme with sub-blocks.
This strategy does not suffer from error enhancement at high SNR, and produce steady error-floors. Nevertheless IR can perform slightly worse than NI at low SNR, due to pessimistic representation of IBI covariance.

D. Overlap FDE with Interference Cancellation

To completely remove residual IBI in overlap FDE with limited overlap interval, interference cancellation is needed, especially for highly selective channels, where equalization filter has time response of length comparable to FFT, and spreads IBI over all symbols.

There are various approaches to IBI cancellation in overlap FDE, either with serial decision feedback for joint ISI/IBI cancellation [22], with hard decision feedback for successive IBI cancellation [53], or with hybrid turbo and hard successive decision feedback [49]. Unlike these references, which uses decisions on previously processed sub-blocks, here we focus on parallel IBI cancellation, using solely a feedback generated from the previous SI/TI, for ensuring parallel processing of sub-blocks in practical implementations. Moreover, EP-based feedback is used, as its overall superiority compared to EXT or APP feedback was shown in the previous section.

At \( \tau = s = 0 \), IR is used via (35), then IBI is removed before the \( N \)-point FFTs with

\[
\tilde{y}_k \triangleq \tilde{y}_k - G_k \left( x_{k-N}^d(\tau', s') - x_k^d(\tau', s') \right),
\]

where \( \tau' \) and \( s' \) denote the previous TI/SI index.

Moreover, unlike prior work on overlap FDE-IC, we use adaptive IR, by accounting for the residual IBI in filter weight computations with

\[ \sum \omega_k = \sum \omega_k + 2 \nu^d(\tau', s') F_N G_k^H F_N^H. \]

As in overlap FDE NI/IR strategies above, \( N_b \) parallel equalizers are operated concurrently for detecting all sub-blocks. Finally, it is possible, depending on the channel coherence time, to set \( N_l = N_r = 0 \), for \( \tau > 0 \), to reduce \( N_b \), as in [49], to reduce the receiver complexity.

E. Inter-block interference mitigation performance

In this section, \( K \)-block quasi-static channels are considered, to focus on the EP-based overlap FDEs’ IBI mitigation capabilities. The benefits of SI are compared to the conventional FD LE-EXTIC (i.e. \( S = 0 \)), for overlap FDE, possibly equipped with IR and/or IC. The IC strategy of setting \( N_l = N_r = 0 \) for \( \tau > 0 \) for overlap FDE IC is used for these simulations.

First we consider an uncoded scenario, similar to the benchmark [51], with QPSK constellation in a quasi-static Rayleigh fading frequency-selective channel with symbol spaced 16-path uniform power delay profile (EQU16). Transmission parameters are \( K = 2048 \), \( N = 256 \) and \( N_l = N_r = 16, \) and 80000 block transmissions per SNR are used to numerically approximate the BER for \( S = 0 \ldots 3 \) in Fig. 13-(a). The conventional scheme (NI) is unusable, as the overlap interval is insufficient to contain all IBI, and SI (\( \beta = 0.25 \times 0.9^{\tau} \)) do not resist to IBI amplifications. But IR significantly benefits from SI, as it further reduces the error floor. Finally, overlap IC with SIs removes most of the interference, even with a single SI.

A more extreme case, with strong IBI, is considered in Fig. 13-(b) (16-QAM, RSC [1, 5/7]), within a 7-path static AWGN channel, with uniform power delay profile. We consider 50000 block transmissions with \( K = 1024, N = 128, \) and \( N_l = N_r = 7 \), to evaluate the BER. In this case, SI (\( \beta = 0.75 \times 0.9^{\tau} \)) alone cannot remove error floors even with IC and channel coding, but with the help of a single TI, even E IR’s error floor, with EP-based SI, becomes at least two order of magnitudes smaller than traditional FD LE-EXTIC.

F. Performance in a doubly-selective channel

The behaviour of the overlap FDE with the proposed EP-based self-iterations is evaluated within a mobile ad-hoc network (MANET) scenario where mobile-to-mobile communications between two high-speed vehicles is considered in a harsh environment. The mountainous channel model from [54] Tab. 5.10) is used. Vehicles are assumed to move at 130 km/h each, in opposing directions, hence generating a maximum
Doppler shift of 96 Hz, assuming the use of a carrier frequency at 400 MHz. A snapshot of a random channel realization is plotted in Fig. 14(a).

SC transmissions with 1/2-rate-coded 16-QAM constellation is considered, with a baud-rate of 1 Mbaud/s, and a root-raised-cosine pulse-shaping with a roll-off factor of 0.35. In this case, the base-band channel spread is $L = 45$ symbols. $N = 256$ symbol is chosen to ensure that the channel remains static on each sub-block. We consider $K = 1536$ and $N_t = N_r = 18$, an overlap length of 18 symbols is chosen as most significant paths of the mountainous channel, (and other urban, hilly or rural channels in $\mathcal{N}$) are contained within 18 $\mu$s.

In Fig. 14(b), the block error rate (BLER) of overlap FDE IR/IC is plotted. It can be seen that IR cannot get rid of the error floor but using overlap IC and one TI, robust transmissions are possible. In this case, the base-band channel spread is $L = 45$ symbols. $N = 256$ symbol is chosen to ensure that the channel remains static on each sub-block. We consider $K = 1536$ and $N_t = N_r = 18$, an overlap length of 18 symbols is chosen as most significant paths of the mountainous channel, (and other urban, hilly or rural channels in $\mathcal{N}$) are contained within 18 $\mu$s.

V. APPLICATION TO MULTI-ANTENNA SPATIAL MULTIPLEXING: FD MIMO RECEIVER

A. System Model and Overview of Resolution

Here, the extension of the SC-FDE model in section II to incorporate multiple antennas is considered. The transmitter and the receiver have respectively $T$ and $R$ antennas, and space-time bit-interleaved coded modulation (STBCM) is used $[55]$. This ensures the transmitted symbols blocks $x_t$ on each transmit antenna $t$ to be independent of each other, and coded bits $d_{k,t} \in \mathbb{F}_2^T$ associated to each symbol to be bit-wise independent as a generalization of the BICM.

Assuming the use of a CP on each antenna, and using ideal synchronization and ideal channel state knowledge hypotheses at the receiver, received samples on the $r$th antenna are

$$y_r = \sum_{t=1}^{T} h_{r,t} x_t + w_r,$$

where the $K \times K$ matrix $h_{r,t}$ is the circulant channel matrix associated to the $L$-tap impulse response $[h_{1,r,t}, \ldots, h_{L,r,t}]$ of the channel between $t$th TX, and the $r$th RX antennas, and where the noise $w_r \sim \mathcal{CN}(0, \sigma_w^2 I_K)$. The FD channel is $H_{r,t} = F_K h_{r,t} F_K^{-1}$, as in section II. Stacking receiver antennas to form $Y_r = [y_1; \ldots; y_r]$, transmit antennas for $X = [x_1; \ldots; x_T]$, we have

$$Y_r = H_{r,t} F K X + W_r,$$

where $H$ is a $RK \times TK$-partitioned-diagonal matrix, $F_K T = I_F \otimes F_K$ is the $T$-block DFT matrix, and $W_r = [w_1; \ldots; w_r]$ with $w_r \sim \mathcal{CN}(0, \sigma_w^2 I_K)$.

A MIMO detector can be designed by applying the proposed framework on the joint PDF $p(d, x|y)$ of this STBCM system, factorized as

$$p(y|x) \prod_{t=1}^{T} \prod_{k=0}^{K-1} p(x_{k,t} | d_{k,t}) \prod_{j=0}^{K-1} p(d_{k,j,t}).$$

Detailed derivation is not given here due to lack of space, however a multi-user MIMO system for non-orthogonal multiple-access with SC-FDMA waveform, which generalizes this system, is derived in $[32]$. The resulting MIMO detector’s outputs are given by

$$\hat{x}_{r,t}^d = \xi_{r,t}^d + \sum_{r'=1}^{R} \sum_{t'=1}^{T} \lambda^d_{k,r,r'} h_{k,r',t},$$

$$\hat{x}_{r,t}^d = 1/\xi_{t} - v_{r,t},$$

where $\sum_{r'=1}^{R} \sum_{t'=1}^{T} \lambda_{k,r,r'}^{d} h_{k,r',t'}$ with $\lambda_{k,r,r'}^{d}$ being the $k$th diagonal of $\Sigma^d$, $(r,r')$th partition, and $\xi_t = K^{-1} \sum_{r=1}^{R} h_{k,r,t}^H d_{k,r,t}$. The covariance matrix $\Sigma^d$ is given by

$$\Sigma^d = \sigma_w^2 I_{RK}^{d} + \sum_{t=1}^{T} v_{r,t}^d H_t^H H_t,$$

where $H_t \in \mathbb{C}^{RK \times K}$ is given by the partitioning $H_t = [H_{t,1}, \ldots, H_{t,R}, \ldots, H_{t,T,U}]$. This covariance matrix and its inverse have a partitioned-diagonal structure, which allows

![Fig. 14. EP-based overlap FDE performance in the doubly-selective mountainous channel.](image-url)
using $\chi_{k,r,r'}^d$ for computationally-efficient detection. For each antenna $t$, a separate EP-based demapper is used, with their specific input and output variances, i.e. respectively $v_r^d$ and $v_{r'}^d$, for characterizing temporally white soft estimates of transmitted symbols.

B. Performance Comparison

The proposed FD MIMO detector is evaluated in the spatial-multiplexing scenario of [17]: over the generalized AWGN Proakis B channel with $T = R = 2$, $K = 128$. Up to 2 SIs are considered with $\beta = \max(0.3, 0.5 \times (0.8)^{r-t})$, and average BER per antenna is evaluated. In Fig. 15(a), the multi-antenna interference (MAI) is mitigated with a parallel IC (PIC) schedule, as in [17], i.e. with simultaneous detection over antennas in each SI, and simultaneous decoding of all antennas in each TI. Our proposal displays remarkable gains over APP-based prior work, with over 2 dB and 2.5 dB gains at 4 TIs, at BER $= 10^{-5}$, with respectively 1 and 2 SIs.

In Fig. 15(b), the MAI is mitigated with a successive IC (SIC) schedule, as in [32], i.e. with simultaneous detection over antennas in each SI, but with successive decoding of antennas in each TI. This approach is known to converge faster. Our proposal outperforms concurrent structures for all TI, with over 1.5 dB margin for BER $= 10^{-5}$. Moreover, SILE-APPIC with either SIC or PIC, at 4 TIs, is outperformed by either 1-SILE-EPIC with $T = 1$ with PIC or SIC. Asymptotically ($T = 4$), SIC improves our proposal’s BER around 0.5 dB over PIC, but SIC with 1 TIs is shown to significantly outperform alternatives, which provides an attractive compromise of fewer decoder iterations, but increased detector iterations, to provide an attractive complexity-performance options, especially when using powerful decoders.

VI. Conclusion

This paper considers a new framework based on expectation propagation message passing for the design of low-complexity digital receivers. This approach’s particularity lies in constraining transmitted symbol variable nodes to lie in multivariate temporally-white Gaussian distributions. This allows deriving low-complexity EP-based demappers which can provide extrinsic symbol-wise feedback, whose reliability measure is characterized by a scalar variance.

The proposed methodology is exposed through the design and analysis of an elementary SC-FDE receiver, which is shown to be either more energy-efficient or less complex than alternatives of the state of the art. Resulting receiver can be seen as a double-loop, low-complexity single-tap FDE which can achieve remarkable energy savings with conventional forward error correction techniques. In particular, it is shown through asymptotic analysis, that a considerable portion of the channel symmetric information rate region becomes achievable.

Furthermore, the flexibility of the proposed approach for receiver design is shown by applying it to the three categories of overlap FDE receivers, and a MIMO detector for spatial multiplexing. In all cases, significant improvements were observed in terms of performance-complexity trade-off. Other practical applications of this framework will be exposed in future works.

REFERENCES

[1] K. Higuchi and A. Benjebbour, “Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access,” IEICE Trans. on Commun., vol. 98, no. 3, pp. 403–414, 2015.
[2] A. Bazzi, B. M. Masini et al., “On the performance of IEEE 802.11p and LTE-V2V for the cooperative awareness of connected vehicles,” IEEE Trans. Veh. Technol., vol. 66, no. 11, pp. 10419–10432, Nov. 2017.
[3] G. Bauch and V. Franz, “A comparison of soft-in/soft-out algorithms for turbo detection,” in Proc. Int. Conf. on Telecomm., 1994, pp. 259–263.
[4] C. Douillard, M. Jézéquel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, “Iterative correction of intersymbol interference: Turbo-quantization,” Eur. Trans. Telecomm., vol. 6, no. 5, pp. 507–511, Sep. 1995.
[5] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate,” IEEE Trans. Inf. Theory, vol. 20, no. 2, pp. 284–287, Mar. 1974.
[6] A. Glavieux, C. Laot, and J. Labat, “Turbo equalization over a frequency selective channel,” in Proc. 1st Symp. Turbo Codes, 1997, pp. 96–102.
[7] 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.
