

Factor Graph Based Message Passing Algorithms for Joint Phase-Noise Estimation and Decoding in OFDM-IM

Qiaolin Shi, Nan Wu, *Member, IEEE*, Hua Wang, *Member, IEEE*, Xiaoli Ma, *Fellow, IEEE*, Lajos Hanzo, *Fellow, IEEE*

Abstract—In order to glean benefits from orthogonal frequency division multiplexing combined with index modulation (OFDM-IM) in the presence of strong Phase-Noise (PHN), in this paper, low-complexity joint PHN estimation and decoding methods are developed in the framework of message passing on a factor graph. Both the Wiener process and the truncated discrete cosine transform (DCT) expansion model are considered for approximating the PHN variation. Then based on these a factor graph is constructed for explicitly representing the joint estimation and detection problem. Taking full account of the sparse and structured *a priori* information arriving from the soft-in soft-out (SISO) decoder of a turbo receiver, a modified generalized approximate message passing (GAMP) algorithm is invoked for decoupling the frequency-domain symbols. In the decoupling step, mean field (MF) approximation is employed for solving the unknown nonlinear transform matrix problem imposed by PHN. Furthermore, merged belief propagation and MF (BP-MF) methods amalgamated both with sequential and parallel message passing schedules are introduced and compared to the proposed GAMP based algorithms in terms of their bit error ratio (BER) vs. complexity. Our simulation results demonstrate the efficiency of the proposed algorithms in the presence of both perfect and imperfect channel state information.

Index Terms—Orthogonal frequency division multiplexing (OFDM), index modulation (IM), phase noise (PHN), discrete cosine transform (DCT), message passing algorithm.

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is one of the most widespread multi-carrier transmission techniques designed for satisfying the increasing demand for high-rate wireless communications. Compared to single-carrier systems, OFDM is capable of improving the spectral efficiency (SE), despite its reduced detection complexity, in wideband channels. Recently, the emerging concept of index modulation (IM) [1], [2] has been applied to classic OFDM to potentially attain a lower peak-to-average power ratio (PAPR) and improved bit error ratio (BER) performance over its conventional OFDM counterpart [3]–[6]. By conveying information through multiple distinguishable modulation modes using all subcarriers

and the permutations of these distinguishable modes, a novel multiple-mode OFDM-IM scheme was developed in [7]–[9] for improving the SE.

To elaborate a little further, in the OFDM-IM scheme, information is conveyed not only by conventional phase shift keying (PSK) or quadrature amplitude modulation (QAM) signal constellations, but implicitly also by the specific combinations of the activated subcarrier indices. However, in contrast to the single-tap frequency-domain equalizer of classic OFDM, a joint PSK/QAM symbol and IM symbol detector is required for OFDM-IM systems. Bearing in mind that the values of frequency-domain symbols can be zero or non-zero, a reduced-complexity log-likelihood ratio (LLR) detector was proposed in [4] for reducing the excessive complexity of the joint maximum likelihood (ML) receiver. Inspired by the benefits of the greedy pursuit algorithms conceived for the estimation of sparse vectors, a low-complexity iterative residual check (IRC)-based detector is developed in [10], where the associated hard decisions are refined by making locally optimal choices at each step.

Nonetheless, as for all emerging techniques, there are also some open issues such as the design of powerful OFDM-IM detectors for realistic channel conditions, because in practice, OFDM-IM is extremely sensitive both to dispersive time-varying fading channels, as well as to carrier frequency offset (CFO) and to Phase-Noise (PHN) [11], just like classic OFDM. However, in most of the previous studies, quasi-static frequency-selective fading channels are considered, where the channel state information (CSI) is assumed to be perfectly known at the OFDM-IM receiver [10], [12]–[14]. As a further aggravation, in contrast to the channel impulse response (CIR) and CFO, which vary slowly with reference to the OFDM symbol rate, the PHN arising from the voltage-controlled oscillator's jitter fluctuates rapidly and thus cannot be eliminated with the aid of pilots. The existing papers mainly focus on the PHN mitigation techniques of classic OFDM [15]–[17], whilst PHN estimation has not been considered in OFDM-IM systems.

Indeed, it is quite a challenge to design a low-complexity OFDM-IM receiver in the presence of strong PHN. The unknown time-varying nonlinear effects imposed by PHN result in a loss of subcarrier orthogonality, which inevitably leads to inter-channel interference (ICI). Accordingly, the system implementation and performance are considerably affected, if the PHN is inadequately compensated. Specifically, if each index modulated symbol vector (IM-symbol) in each OFDM-IM subblock is detected independently, the performance erodes owing to the presence of correlated, i.e., colored, PHN. If

This work was supported by the “National Science Foundation of China (NSFC)” (Grant No.61571041,61971041,61471037). L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/N004558/1, EP/P034284/1, COALESCE, of the Royal Society’s Global Challenges Research Fund Grant as well as of the European Research Councils Advanced Fellow Grant QuantCom. (*Corresponding Author: Nan Wu.*)

Q. Shi is with the School of Information and Electronics, Beijing Institute of Technology, 100081, China, and the Faculty of Engineering and Information Technology, University of Technology Sydney, NSW 2007, Australia (e-mail: qlshi@bit.edu.cn).

N. Wu, and H. Wang are with the School of Information and Electronics, Beijing Institute of Technology, 100081, China (e-mail: {wunan, wanghua}@bit.edu.cn).

X. Ma is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA30332-0250, USA (e-mail: xiaoli@gatech.edu).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk).

however several IM-symbols are jointly detected for mitigating the PHN, the complexity of the receiver becomes prohibitive, even if reduced-search based detectors are used. Therefore, efficient PHN estimation techniques have to be designed to attain reliable BER performance and glean multiple advantages from OFDM-IM schemes. Moreover, wireless systems tend to rely on iterative channel decoding based on the powerful turbo principle to attain a low BER at SNRs close to channel capacity. However, the existing PHN estimation and OFDM symbol detection problems are predominantly formulated for uncoded systems. Against this background, our goal is to propose low-complexity methods for joint PHN estimation and OFDM-IM symbol detection. Very recently, message passing algorithms defined over factor graphs have gained popularity, including both belief propagation (BP) [18], as well as mean-field (MF) approximation based variational methods [19] and generalized approximate message passing (GAMP) techniques [20]. The BP and MF methods defined over factor graphs are derived from the stationary points of the corresponding variational free energy [21]. By forwarding messages from one node of the factor graph to neighboring nodes according to message updating rules (refer to [18] and [19] for BP and MF rules, respectively), the distance measure between a trial distribution and the original objective function to be estimated can be minimized. GAMP constitutes a state-of-the-art BP approximation, where the message passing between variable nodes and factor nodes is simplified by exploiting the central limit theorem and a second-order Taylor series (refer to [20] for GAMP rules). A structured approximate message passing (AMP) detector is developed for spatial modulation aided systems in [22], which is further generalized in [23], where perfectly known block fading Rayleigh channels are assumed, albeit no joint iterative turbo detection is invoked.

Explicitly, the contributions of this paper are summarized as follows:

- We propose a GAMP based low-complexity joint PHN estimation and decoding algorithm for OFDM-IM communicating in frequency-selective fading channels. The structured *a priori* information gleaned from the low-density parity-check (LDPC) soft-in soft-out (SISO) decoder in the turbo receiver [24] is specifically constructed to take into account the sparsity of the OFDM-IM symbols. Given that the consecutive frequency-domain IM symbols are not independent due to the existence of PHN, the GAMP algorithm is invoked for decoupling the vector estimation process into scalar operations. To circumvent the nonlinear and unknown transform matrix problem in the decoupling step of the GAMP [20], the MF method [19] is employed at the observation node (refer to [18] for details about the observation node defined over factor graphs) for approximating the output function f_{out} defined in GAMP.
- We consider two PHN models, i.e., the Wiener model [15], [16] and the truncated discrete cosine transform (DCT) expansion model [25], for representing the PHN. In contrast to linearizing the system model directly as in [15], the time-domain statistics of PHN are explicitly

exploited based on the Wiener model. In contrast to [25], where only hard estimates of the DCT coefficients are given by the expectation-maximization algorithm, the probability distributions of the DCT coefficients are provided. By exploiting the PHN estimates, the accuracy of data detection is gradually improved in an iterative fashion.

- We invoke the amalgamated BP-MF algorithm for joint PHN estimation and decoding as a benchmark. Parallel and sequential message passing aided schedules relying on the DCT model and Wiener model, respectively, are introduced for comparison. The complexity of the parallel scheduling is similar to that of the proposed GAMP based algorithms, since the efficiency of fast Fourier transform (FFT) can be exploited. However, some performance loss is observed due to its low convergence rate compared to both the sequential and to the proposed GAMP based methods. As for the Wiener model, the amalgamated BP-MF relying on sequential message passing outperforms the intrinsically parallel GAMP method, albeit at the expense of an increase of complexity.
- The performance of the proposed algorithms is evaluated and compared to that of classic OFDM for both perfect and realistic imperfect channel conditions. Both subcarrier interleaving and coordinate interleaving techniques are used for further improving the BER performance by increasing the diversity order. Our comparisons both to the conventional pilot based method, to the extended Kalman filter (EKF) of [15] and to the variational inference (VI) approach of [16] are also shown. Our simulation results: 1) confirm the robustness of the OFDM-IM scheme relying on the proposed algorithms both in the face of severe PHN, as well as realistic imperfect CSI and residual CFO compared to its classic OFDM counterpart; and 2) demonstrate the superiority of the proposed algorithm over existing methods.

The rest of the paper is organized as follows. In Section II, the system model of OFDM-IM symbols transmitted over frequency-selective fading channels in the presence of PHN is formulated. Then in Section III, both the Wiener model [15], [16] as well as the DCT model [25] are considered and then both the GAMP-Wiener and the GAMP-DCT algorithms are developed for joint PHN estimation and detection. In Section IV, the amalgamated BP-MF methods relying on sequential and parallel message passing are introduced and their complexity is compared to that of the GAMP based algorithms. Our simulation results and performance analysis are provided in Section V, where the proposed algorithms are also evaluated in the context of imperfect channel conditions. Finally, our conclusions are offered in Section VI.

Notation: x_n and $X_{m,n}$ represent the n th and the (m, n) th elements in vector \mathbf{x} and matrix \mathbf{X} , respectively. $\mathbf{X}_{:,n}$ and $\mathbf{X}_{n,:}$ indicate the n th column and the n th row of \mathbf{X} , respectively. \mathbf{I}_N and $\mathbf{1}$ denote the $N \times N$ identity matrix and the all-one column vector, respectively. $\mathcal{D}(\mathbf{x})$ is a diagonal matrix with the vector \mathbf{x} on its diagonal. $\mathcal{N}_c(x; \hat{m}_x, \hat{v}_x)$ represents the circular Gaussian probability density function (pdf) of the random variable x with mean \hat{m}_x and variance \hat{v}_x . The

relationship $f(x) = cg(x)$ for some positive constant c is written as $f(x) \propto g(x)$. $\lfloor \cdot \rfloor$ is the floor operation. $E_{b(x)}[\cdot]$ denotes the expectation with respect to the pdf $b(x)$. $\delta(\cdot)$ is the Dirac delta function. The operator \odot denotes element-wise product and $||\Omega||_0$ the size of set Ω . Φ is the DCT matrix with the (n, l) th entry $\Phi_{n,l}$ given by $\sqrt{\frac{2}{N}} \cos(\frac{\pi l}{N}(n + \frac{1}{2}))$.

II. SYSTEM MODEL

Let us consider an LDPC coded OFDM-IM system having N_q data subcarriers in each OFDM-IM block, whilst communicating over an L_h -tap frequency-selective fading channel. Fig. 1 depicts the transceiver schematic of the OFDM-IM system. The data bits $\mathbf{c}_q \triangleq [c_1^{(q)}, \dots, c_{M_q}^{(q)}]^T$ are generated by encoding M_i information bits $\mathbf{b} \triangleq [b_1, \dots, b_{M_i}]^T$ using a rate- R_c LDPC coder and then by partitioning the resultant $M_c = M_i/R_c$ bits $\mathbf{c} \triangleq [c_1, \dots, c_{M_c}]^T$ into an integer number of $Q \triangleq M_c/M_q$ OFDM-IM symbols. These M_q bits in the q th block are then splitted into G subgroups each containing M_g bits, i.e., we have $M_g = M_q/G$. Each subgroup of M_g -bits is mapped to an OFDM-IM subblock of length N_g , where $N_g = N_q/G$. Here, N_q is the size of the inverse fast Fourier transform (IFFT) of Fig. 1. In each OFDM-IM subblock, only K_g out of N_g available subcarriers determined by the first M_{g1} bits of the incoming M_g -bit sequence are activated for transmitting S -ary classic PSK/QAM symbols, while the remaining $(N_g - K_g)$ inactive subcarriers are set to zero. In addition to the M_{g1} bits transmitted by the active subcarriers, OFDM-IM also transmits $M_{g2} \triangleq M_g - M_{g1}$ bits by mapping them to K_g S -ary PSK/QAM symbols conveyed by the K_g active subcarriers. Therefore, we have $M_g = M_{g1} + M_{g2} = \left\lfloor \log_2 \binom{N_g}{K_g} \right\rfloor + K_g \log_2 S$, where $\binom{N_g}{K_g}$ denotes the binomial coefficient representing the number of subcarrier combinations. Then the M_{g1} bits determine $\Upsilon = 2^{\left\lfloor \log_2 \binom{N_g}{K_g} \right\rfloor}$ legitimate combinations of the active subcarriers. We define the indicator set $\Omega = \{\omega_1, \dots, \omega_\Upsilon\}$ associated with $\omega_\gamma = [\omega_{\gamma,1}, \dots, \omega_{\gamma,N_g}]^T \in \{0, 1\}^{N_g}$, where $\omega_{\gamma,i} = 1$ if the i th subcarrier is activated, and $\omega_{\gamma,i} = 0$ otherwise. The coded bits corresponding to ω_γ are represented by $[c_1^{\omega_\gamma}, \dots, c_{M_{g1}}^{\omega_\gamma}]^T$.

Specifically, we assume that the indices activated in the g th OFDM-IM subblock are given by $\mathcal{I}_g = \{i_{g,1}, \dots, i_{g,K_g}\}$, $g = 1, \dots, G$, where $i_{g,k_g} \in [1, \dots, N_g]$, and the non-zero symbols assigned to the K_g active subcarriers are collected as $\mathbf{x}_{g,d} = [x_{g,d}(1), \dots, x_{g,d}(K_g)]^T$, $g = 1, \dots, G$, where $x_{g,d}(k_g) \in \mathcal{S}$ and $\mathcal{S} = \{s_1, \dots, s_S\}$ denotes the normalized S -PSK/QAM symbol set. The coded bits corresponding to the m th constellation symbol s_m in the set are denoted by $[c_1^{s_m}, \dots, c_{\log_2 S}^{s_m}]^T$. Then the K_g data symbols are allocated to K_g active subcarriers of the index set in \mathcal{I}_g , yielding a subblock of symbols \mathbf{x}_g , which can be expressed as $\mathbf{x}_g = \mathbf{I}_g^0 \mathbf{x}_{g,d} \triangleq [x_g(1), \dots, x_g(N_g)]^T$, $g = 1, \dots, G$, where \mathbf{I}_g^0 consists of the K_g columns from \mathbf{I}_{N_g} having the indices given by \mathcal{I}_g . The index modulated symbol vector \mathbf{x}_g in the g OFDM-IM subblock is referred to as the g th IM-symbol. The activation pattern constraint here is that $\mathbf{x}_g^{\text{AP}} \in \Omega$, where $x_g^{\text{AP}}(n) = 0$ if $x_g(n) = 0$ and $x_g^{\text{AP}}(n) = 1$ otherwise. In other words, the set of Υ legitimate combinations is used at

both the transmitter and receiver sides, and the transmitter chooses an index combination from the set for each OFDM-IM subblock. After concatenating all symbol vectors from G subblocks by the OFDM-IM block assembler of Fig. 1, a main OFDM-IM block can be written as $\mathbf{x} = [\mathbf{x}_1^T, \dots, \mathbf{x}_G^T]^T \triangleq [x_1, \dots, x_{N_q}]^T \in \mathbb{C}^{N_q \times 1}$, where the n th element in \mathbf{x} obeys $x_n \in \{\mathcal{S}, 0\}$. Note that with a little abuse of notation, we may denote the bit vector transmitted in the g th subblock by $\mathbf{c}_g = [c_{g,1}^{\text{IM}}, \dots, c_{g,M_{g1}}^{\text{IM}}, c_{g,1}^{\text{S}}, \dots, c_{g,M_{g2}}^{\text{S}}]^T \in \{0, 1\}^{M_g}$, where $c_{g,i}^{\text{IM}}$, $i = 1, \dots, M_{g1}$, is the i th subcarrier selection bit, and $c_{g,j}^{\text{S}}$, $j = 1, \dots, M_{g2}$, the j th symbol mapping bit.

By defining the total number of active indices in an OFDM-IM block as $K_q \triangleq GK_g$, we have $E[\mathbf{x}^H \mathbf{x}] = K_q$. At the transmitter, the IFFT having a normalization factor $\frac{N_q}{\sqrt{K_q}}$ is employed for generating the time-domain symbol vector $\mathbf{x}_T = \frac{1}{\sqrt{K_q}} \mathbf{F}^H \mathbf{x}$, where $\mathbf{F} \in \mathbb{C}^{N_q \times N_q}$ is the discrete Fourier transform (DFT) matrix associated with $\mathbf{F}^H \mathbf{F} = N_q \mathbf{I}_{N_q}$. Thus, the transmitted power is normalized, i.e., $E[\mathbf{x}_T^H \mathbf{x}_T] = N_q$. Then, an L_{cp} -length cyclic prefix (CP), which has to be longer than the channel's delay spread for the sake of eliminating the effects of the inter-symbol-interference (ISI) over frequency-selective fading channels, is added at the front of the time-domain symbols. Then the signal generated at the transmitter is sent over an L_h -tap frequency-selective fading channel. Initially, the CIRs are assumed to be perfectly known at the receiver, but this assumption will be relaxed in Section V, where the effects of imperfect CIRs will also be considered. In the presence of unknown PHN distortion, the N_q -point complex baseband received signal sequence $\mathbf{r} \in \mathbb{C}^{N_q \times 1}$ of an OFDM-IM block can be written as

$$\begin{aligned} \mathbf{r} &= \frac{1}{\sqrt{K_q}} \mathcal{D}(e^{j\theta}) \mathbf{R}_{\text{cp}} \mathbf{H}_t \mathbf{T}_{\text{cp}} \mathbf{F}^H \mathbf{x} + \mathbf{w} \\ &= \frac{1}{\sqrt{K_q}} \mathcal{D}(e^{j\theta}) \mathbf{F}^H \mathbf{H} \mathbf{x} + \mathbf{w}, \end{aligned} \quad (1)$$

where $\theta = [\theta_1, \dots, \theta_{N_q}]^T$ is the discrete PHN sample vector, $\mathbf{T}_{\text{cp}} = \begin{bmatrix} \mathbf{0}_{L_{\text{cp}} \times (N_q - L_{\text{cp}})} & \mathbf{I}_{L_{\text{cp}}} \\ & \mathbf{I}_{N_q} \end{bmatrix} \in \mathbb{R}^{(N_q + L_{\text{cp}}) \times N_q}$ inserts the L_{cp} -length CP, \mathbf{H}_t denotes the time-domain channel convolutional matrix consisting of the channel vector $\mathbf{h} = [h_1, \dots, h_{L_h}]^T \in \mathbb{C}^{L_h \times 1}$, $\mathbf{R}_{\text{cp}} = \begin{bmatrix} \mathbf{0}_{N_q \times L_{\text{cp}}} & \mathbf{I}_{N_q} \end{bmatrix} \in \mathbb{R}^{N_q \times (N_q + L_{\text{cp}})}$ removes the CP, and $\mathbf{w} \in \mathbb{C}^{N_q \times 1}$ is the zero-mean complex circularly symmetric Gaussian noise vector with a variance of $2\sigma^2$. In (1), $\mathbf{R}_{\text{cp}} \mathbf{H}_t \mathbf{T}_{\text{cp}} = \frac{1}{N_q} \mathbf{F}^H \mathbf{H} \mathbf{F}$ is a circular matrix and $\mathbf{H} = \mathcal{D}(\mathbf{F} \mathbf{I}_{L_h}^0 \mathbf{h})$ with $\mathbf{I}_{L_h}^0$ being formed by the first L_h columns from \mathbf{I}_{N_q} .

As illustrated in Fig. 1, at the receiver, the proposed joint PHN estimation and OFDM-IM decoding algorithm is performed by: 1) estimating the PHN using the received signal sequence and the OFDM-IM symbols' estimates from the previous iteration; 2) updating the OFDM-IM symbols' estimates using the received signal sequence and the PHN samples' most recent estimates; 3) calculating the *a posteriori* pdfs of information bits according to the IM and PSK/QAM modulation constraint; 4) extracting extrinsic information from the *a posteriori* information and then forwarding them to the LDPC SISO decoder to perform Bahl, Cocke, Jelinek,

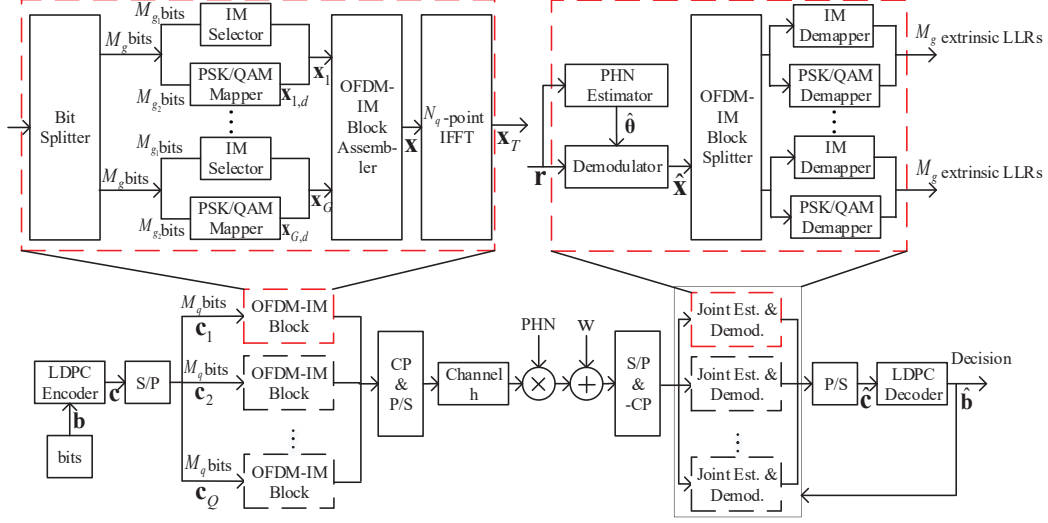


Fig. 1. The transceiver of the LDPC-coded OFDM-IM system.

and Raviv (BCJR) decoding [26]; and 5) calculating the *a priori* pdfs of information bits using the outputs of the decoder and then constructing the sparse *a priori* pdfs of the OFDM-IM symbols, which are further applied to update the PHN samples' and OFDM-IM symbols' estimates for the next iteration.

III. THE PROPOSED JOINT PHN ESTIMATION AND OFDM-IM SYMBOL DETECTION ALGORITHMS

A. Probabilistic Representation and Factor Graph

We employ a Wiener process [15]–[17], which is widely used for modelling a free-running oscillator, in order to approximate the variation of PHN, i.e.,

$$\theta_n = \theta_{n-1} + \Delta_n, \quad n = 2, \dots, N_q, \quad (2)$$

where Δ_n is a zero-mean white Gaussian noise process with a variance of σ_Δ^2 . Based on (2) and on the received signal model in (1), we obtain the joint *a posteriori* pdf of \mathbf{x} and $\boldsymbol{\theta}$

$$p(\mathbf{x}, \boldsymbol{\theta} | \mathbf{r}) \propto f_C(\mathbf{c}, \mathbf{b}) \prod_{g=1}^G f_{M,g}(\mathbf{x}_g, \mathbf{c}_g) P(\mathbf{x}) p(\theta_1) \cdot \prod_{n=2}^{N_q} f_{\theta_n}(\theta_n, \theta_{n-1}) \prod_{n=1}^{N_q} f_{r_n}(z_n, \theta_n) f_{\delta_n}(z_n, \mathbf{x}), \quad (3)$$

where $f_C(\mathbf{c}, \mathbf{b})$ represents the coding constraint; $f_{M,g}$ denotes the IM and PSK/QAM modulation constraint; $P(\mathbf{x})$ is the *a priori* pdf of the data symbol \mathbf{x} ; $p(\theta_1)$ is the *a priori* pdf of θ_1 and $f_{\theta_n}(\theta_n, \theta_{n-1}) \propto \mathcal{N}(\theta_n; \theta_{n-1}, \sigma_\Delta^2)$. In (3), the additional variable z_n is defined as $z_n \triangleq \Lambda_{n,:} \mathbf{x}$, where

$$\Lambda \triangleq \frac{1}{\sqrt{K_q}} \mathbf{F}^H \mathbf{H}. \quad (4)$$

Accordingly, the relationship between the variable z_n and the vector \mathbf{x} can be written as $f_{\delta_n}(z_n, \mathbf{x}) \triangleq p(z_n | \mathbf{x}) = \delta(z_n - \Lambda_{n,:} \mathbf{x})$, and the likelihood function can be calculated by $f_{r_n}(z_n, \theta_n) \triangleq p(r_n | z_n, \theta_n) = \mathcal{N}(r_n; z_n e^{j\theta_n}, 2\sigma^2)$. The factor graph corresponding to (3) is shown in Fig. 2, where the solid

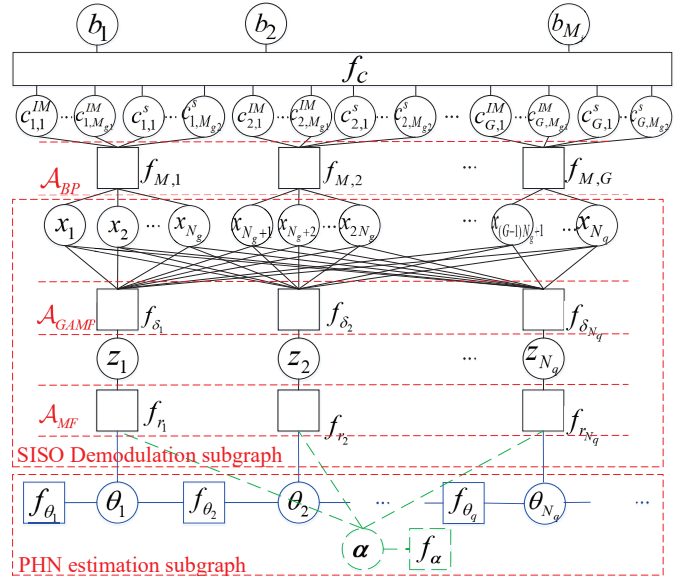


Fig. 2. Factor graph representation of the probabilistic model (3), where the solid lines and dashed lines in the PHN estimation subgraph are associated with the Wiener model and DCT model, respectively.

lines in the PHN estimation subgraph are associated with the Wiener model.

We classify the factor nodes into three disjoint subsets, i.e., $\mathcal{A}_{BP} \triangleq \{\{f_{M,g}\}_{g=1}^G, \{f_{\theta_n}\}_{n=1}^{N_q}\}$, $\mathcal{A}_{MF} \triangleq \{f_{r_n}\}_{n=1}^{N_q}$ and $\mathcal{A}_{GAMP} \triangleq \{f_{\delta_n}\}_{n=1}^{N_q}$. The messages can be readily calculated by splitting the factor graph in this way, because BP is compatible with hard constraints [21], whilst MF has simple message passing update rules for continuous variables belonging to an exponential family [27], and GAMP works well for estimating a vector transformed by a known linear mixing matrix [28], [29].

B. Soft Demodulation and Decoding for OFDM-IM Symbols

In frequency-selective fading channels which are free from PHN, the system model is formulated in the frequency domain, where frequency-domain received signals are subjected to

OFDM-IM detection. In this case, the complexity of the optimal receiver is on the order of $\mathcal{O}(G2^{M_g})$ [4], since the ML detector can be applied to each subblock independently. In the presence of correlated PHN, however, the received IM symbols become correlated and ICI occurs. In this case, the optimal detector has to consider all the legitimate IM-symbol combinations, hence the complexity order becomes $\mathcal{O}(2^{GM_g})$, which grows exponentially with G , K_g , N_g and S . Similarly to classic OFDM [15]–[17], the received time-domain signals are directly used for jointly estimating the PHN and for detecting the OFDM-IM symbols. As seen in (1), the frequency-domain OFDM-IM symbols are coupled by the transform matrix $\mathcal{D}(e^{j\theta})\mathbf{F}^H\mathbf{H}$. The relationship between the input and output is both nonlinear and unknown due to the existence of PHN, which imposes challenges on the detection. Moreover, the complexity of optimal OFDM-IM detection becomes much higher than that of the classic OFDM symbols in both cases, since the receiver has to detect both the subcarrier index-modulated and the PSK/QAM modulated symbols. In the following, we conceive low-complexity methods for joint PHN estimation and decoding of the OFDM-IM symbols affected by PHN.

A straightforward technique of reducing the complexity is to treat x_n , $n = 1, \dots, N_q$, as Gaussian variables [15], [16], [30], [31]. However, significant performance loss is imposed by this method, because x_n is actually a discrete variable and \mathbf{x}_g has a sparse structure. Hence we invoke the GAMP algorithm [20] for decoupling the signal vector estimation problem into N_q scalar problems, whilst taking into account its discrete and sparse nature. Unfortunately, the existing GAMP method of [20] cannot be directly invoked for OFDM-IM detection, because the nonlinear transform matrix imposing the PHN has to be known, which is not the case. Secondly, the sparsity of the signal cannot be readily exploited by the original GAMP [20]. Thirdly, it is more difficult to take full account of the activation pattern constraint if more than one subcarrier of the g th subgroup are activated. To circumvent these problems, we invoke the MF method to approximate the likelihood $p(r_n|z_n)$ for exploiting the available *a priori* information of the frequency-domain symbols provided by the LDPC SISO decoder. Hence appropriately adopt the GAMP algorithm for updating the *a posteriori* pdfs and the extrinsic information of frequency-domain symbols using our turbo receiver.

Specifically, instead of using the independent *a priori* information of the IM-symbols in the form $P(\mathbf{x}) = \prod_{n=1}^{N_q} P(x_n)$, we take into account their correlation for both the IM symbols and PSK/QAM symbols in the g th IM-symbol, using the extrinsic information gleaned from the LDPC SISO decoder (refer to [24] for details about the turbo detection principle), which can be written as $\{P(c_{g,1}^{\text{IM}}), \dots, P(c_{g,M_{g1}}^{\text{IM}}), P(c_{g,1}^{\text{S}}), \dots, P(c_{g,M_{g2}}^{\text{S}})\}$. Given this *a priori* information and the message $m_{x_i \rightarrow f_{M,g}}(x_i) \propto \mathcal{N}_c(x_i; \hat{\tau}_i, v_{\tau_i})$ forwarded from the variable node x_i to the factor node $f_{M,g}$, which will be later updated according to (12) and (13), we can calculate the *a posteriori* pdf $b(x_i)$ of x_i according to the associated IM and PSK/QAM modulation constraint. The detailed derivations of $b(x_i)$ are provided in

Appendix A. Given the *a posteriori* pdf $b(x_i)$, the *a posteriori* mean and variance of x_i are formulated as

$$\hat{x}_i = \sum_{s \in \mathcal{S}} sb(x_i = s), \quad v_{x_i} = \sum_{s \in \mathcal{S}} |s|^2 b(x_i = s) - |\hat{x}_i|^2. \quad (5)$$

By using the GAMP algorithm ([20, Eqs. (2a) and (2b)]), the message $m_{f_{\delta_n} \rightarrow z_n}(z_n) \propto \mathcal{N}_c(z_n; \hat{\xi}_n, v_{\xi_n})$ from the factor node f_{δ_n} to the variable node z_n is updated by

$$v_{\xi_n} = \sum_i |\Lambda_{n,i}|^2 v_{x_i}, \quad (6)$$

$$\hat{\xi}_n = \sum_i \Lambda_{n,i} \hat{x}_i - \hat{\xi}_n v_{\xi_n}, \quad (7)$$

where $\hat{\xi}_n$ related to the *a posteriori* mean of z_n will be later updated according to (11). In the original GAMP, the function F_{out} defined in [20] is given by $F_{\text{out}}(z_n, \hat{\xi}_n, v_{\xi_n}, r_n) = f_{\text{out}} - \frac{1}{v_{\xi_n}} |z_n - \hat{\xi}_n|^2$, where $f_{\text{out}} = \log p(r_n|z_n)$ [20, Eq. (7c)]. However, the relationship between z_n and r_n is nonlinear and unknown due to the presence of PHN. To this end, we use the MF method [19] for calculating the message $m_{f_{r_n} \rightarrow z_n}(z_n)$ forwarded from the factor node f_{r_n} to the variable node z_n , and use it for approximating $p(r_n|z_n)$. Given the *a posteriori* pdf of θ_n , i.e., $b(\theta_n) \propto \mathcal{N}(\theta_n; \hat{\theta}_n, v_{\theta_n})$, which will be later updated in (26), the message $m_{f_{r_n} \rightarrow z_n}(z_n)$ is calculated following the MF rule [19, Eq. (18)] and then approximated with the aid of its second-order Taylor expansion as:

$$m_{f_{r_n} \rightarrow z_n}(z_n) \propto \exp \{E_{b(\theta_n)} [\log f_{r_n}(z_n, e^{j\theta_n})]\} \approx \mathcal{N}_c(z_n; \hat{\eta}_n, v_{\eta_n}), \quad (8)$$

$$\text{where} \quad \hat{\eta}_n = r_n e^{-j\hat{\theta}_n} (1 - \frac{1}{2} v_{\theta_n}), \quad v_{\eta_n} = 2\sigma^2. \quad (9)$$

Bearing in mind that F_{out} is equivalent to the belief of z_n in its log-domain based on the probabilistic understanding of message passing, the *a posteriori* pdf of z_n can be expressed as $b(z_n) \propto \mathcal{N}_c(z_n; \hat{z}_n, v_{z_n})$, where

$$v_{z_n} = (v_{\xi_n}^{-1} + v_{\eta_n}^{-1})^{-1}, \quad \hat{z}_n = v_{z_n} (\hat{\xi}_n v_{\xi_n}^{-1} + \hat{\eta}_n v_{\eta_n}^{-1}). \quad (10)$$

The output scalar function $g_{\text{out}}(\hat{\xi}_n, v_{\xi_n}, \hat{\eta}_n, v_{\eta_n})$ is updated by ([20, Eqs. (3a) and (3b)])

$$\hat{\zeta}_n = g_{\text{out}}(\hat{\xi}_n, v_{\xi_n}, \hat{\eta}_n, v_{\eta_n}) = \frac{\hat{z}_n - \hat{\xi}_n}{v_{\xi_n}} = \frac{\hat{\eta}_n - \hat{\xi}_n}{v_{\eta_n} + v_{\xi_n}},$$

$$v_{\zeta_n} = -\frac{\partial}{\partial \hat{\xi}_n} g_{\text{out}}(\hat{\xi}_n, v_{\xi_n}, \hat{\eta}_n, v_{\eta_n}) = \frac{1}{v_{\xi_n}} \left(1 - \frac{v_{z_n}}{v_{\xi_n}}\right). \quad (11)$$

In the GAMP algorithm ([20, Eqs. (4a) and (4b)]), the variance and the mean of the message $m_{x_i \rightarrow f_{M,g}}(x_i) \propto \mathcal{N}_c(x_i; \hat{\tau}_i, v_{\tau_i})$ are calculated by

$$v_{\tau_i} = \left(\sum_n |\Lambda_{n,i}|^2 v_{\zeta_n} \right)^{-1}, \quad (12)$$

$$\hat{\tau}_i = v_{\tau_i} \sum_n \hat{\zeta}_n \Lambda_{n,i}^* + \hat{x}_i. \quad (13)$$

Note that the complexity of the structured GAMP algorithm is $\mathcal{O}(N_q^2)$, which is dominated by the matrix-vector multiplications in (6), (7), (12), and (13). This complexity can be reduced to $\mathcal{O}(N_q \log N_q)$ thanks to the FFT. Specifically, we

can rewrite (6) as

$$\mathbf{v}_\xi = 1/K_q |\mathbf{F}^H|^2 \mathbf{H}^H \mathbf{H} \mathbf{v}_x \approx 1/K_q \sum_i |h_i|^2 v_{x_i} \mathbf{1}_{N_q \times 1}, \quad (14)$$

where $\mathbf{v}_\xi = [v_{\xi_1}, \dots, v_{\xi_{N_q}}]^T$ and $\mathbf{v}_x = [v_{x_1}, \dots, v_{x_{N_q}}]^T$. Furthermore, we have

$$\hat{\xi} = 1/\sqrt{K_q} \mathbf{F}^H \mathbf{H} \hat{\mathbf{x}} - \hat{\zeta} \odot \mathbf{v}_\xi, \quad (15)$$

where $\hat{\xi} = [\hat{\xi}_1, \dots, \hat{\xi}_{N_q}]^T$, $\hat{\mathbf{x}} = [\hat{x}_1, \dots, \hat{x}_{N_q}]^T$, and $\hat{\zeta} = [\hat{\zeta}_1, \dots, \hat{\zeta}_{N_q}]^T$. Similarly, (12) can also be represented in a vectorial form, i.e.,

$$\mathbf{v}_\tau = \mathbf{1}_{N_q \times 1} \cdot / (\frac{1}{K_q} \mathbf{H}^H \mathbf{H} |\mathbf{F}|^2 \mathbf{v}_\zeta) \approx \frac{K_q}{\bar{v}_\zeta} (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{1}_{N_q \times 1}, \quad (16)$$

where $\mathbf{v}_\tau = [v_{\tau_1}, \dots, v_{\tau_{N_q}}]^T$, $\mathbf{v}_\zeta = [v_{\zeta_1}, \dots, v_{\zeta_{N_q}}]^T$, and $\bar{v}_\zeta = \sum_n v_{\zeta_n}$. Hence, (13) is reduced to the following equation

$$\hat{\tau} = 1/\sqrt{K_q} \mathbf{H}^H \mathbf{F} \hat{\zeta} \odot \mathbf{v}_\tau + \hat{\mathbf{x}}, \quad (17)$$

where $\hat{\tau} = [\hat{\tau}_1, \dots, \hat{\tau}_{N_q}]^T$ and $\hat{\zeta} = [\hat{\zeta}_1, \dots, \hat{\zeta}_{N_q}]^T$. Fortunately, some of the matrix-vector calculations within the GAMP iterations can be efficiently performed by the FFT. The complexity of the specific matrix-vector calculations relying on diagonal matrices is low.

Given the latest updated message $m_{x_i \rightarrow f_{M,g}}(x_i)$ in (12) and (13), the belief of x_i can be updated by (50) and (51), see Appendix A. Meanwhile, we can obtain the belief of \mathbf{x}_g as

$$b(\mathbf{x}_g) \propto P_{\mathbf{x}_g}(\mathbf{x}_g) \prod_{i \in I_g} \exp \left\{ -\frac{|x_i - \hat{\tau}_i|^2}{v_{\tau_i}} \right\}, \quad (18)$$

where $i \in I_g \triangleq \{(g-1)N_g + 1 : gN_g\}$. The extrinsic log-likelihood ratios (LLRs) of the i th subcarrier selection bit and of the j th symbol mapping bit in the g th IM-symbol, i.e. $L^{\text{extr}}(c_{g,i}^{\text{IM}})$ and $L^{\text{extr}}(c_{g,j}^{\text{S}})$, respectively, can be calculated according to the turbo principle [22]–[24].

C. PHN Estimation in the OFDM-IM Scheme

1) *PHN estimation based on the Wiener model:* The message emerging from the observation node to the PHN estimation subgraph can then be calculated by

$$m_{f_{r_n \rightarrow \theta_n}}(\theta_n) \propto \exp \{ E_{b(z_n)} [\log f_{r_n}(z_n, \theta_n)] \} \propto \left\{ \Re \left[\frac{r_n^* \hat{z}_n}{\sigma^2} e^{j\hat{\theta}_n} \right] \right\}. \quad (19)$$

It is readily seen that the message $m_{f_{r_n \rightarrow \theta_n}}(\theta_n)$ in (19) is not Gaussian due to the nonlinear effect imposed by PHN. However, this message is expected to be Gaussian, hence the messages updated in the PHN estimation subgraph are all Gaussian and thus low complexity can be achieved. We use the second-order Taylor expansion for approximating $m_{f_{r_n \rightarrow \theta_n}}(\theta_n)$, instead of directly linearizing the system model [32]–[34]. Given the *a posteriori* mean $\hat{\theta}_n$ and variance v_{θ_n} obtained from the previous iteration, $e^{j\hat{\theta}_n}$ can be approximated as

$$e^{j\hat{\theta}_n} = e^{j\hat{\theta}_n} \left(1 - j\hat{\theta}_n - \frac{1}{2}\hat{\theta}_n^2 \right) - \frac{1}{2}e^{j\hat{\theta}_n} \theta_n^2 + (j + \hat{\theta}_n)e^{j\hat{\theta}_n} \theta_n. \quad (20)$$

Thus, $m_{f_{r_n \rightarrow \theta_n}}(\theta_n)$ can be approximated by $\mathcal{N}(\theta_n; \hat{\theta}_n^\downarrow, v_{\theta_n^\downarrow})$, with

$$v_{\theta_n^\downarrow} = 1/\Re \left[\frac{r_n^* \hat{z}_n}{\sigma^2} e^{j\hat{\theta}_n} \right], \quad \hat{\theta}_n^\downarrow = v_{\theta_n^\downarrow} \Re \left[\frac{r_n^* \hat{z}_n}{\sigma^2} (j + \hat{\theta}_n) e^{j\hat{\theta}_n} \right]. \quad (21)$$

Given this Gaussian message forwarded from the MF subgraph to the PHN estimation subgraph, both $m_{f_{\theta_{n-1} \rightarrow \theta_n}}(\theta_n)$ and $m_{f_{\theta_{n+1} \rightarrow \theta_n}}(\theta_n)$ are Gaussian and can be recursively updated based on the BP algorithm ([18, Eqs. (5) and (6)]) as follows:

$$m_{f_{\theta_n \rightarrow \theta_n}}(\theta_n) \propto \int m_{f_{\theta_{n-1} \rightarrow \theta_{n-1}}}(\theta_{n-1}) f_{\theta_n}(\theta_n, \theta_{n-1}) \cdot m_{f_{r_{n-1} \rightarrow \theta_{n-1}}}(\theta_{n-1}) d\theta_{n-1} \propto \mathcal{N}(\theta_n; \hat{\theta}_n^\rightarrow, v_{\theta_n^\rightarrow}), \quad (22)$$

where

$$\hat{\theta}_n^\rightarrow = \frac{(v_{\theta_{n-1}^\rightarrow})^{-1} \hat{\theta}_{n-1}^\rightarrow + (v_{\theta_{n-1}^\downarrow})^{-1} \hat{\theta}_{n-1}^\downarrow}{(v_{\theta_{n-1}^\rightarrow})^{-1} + (v_{\theta_{n-1}^\downarrow})^{-1}}, \quad v_{\theta_n^\rightarrow} = \sigma_\Delta^2 + \left((v_{\theta_{n-1}^\rightarrow})^{-1} + (v_{\theta_{n-1}^\downarrow})^{-1} \right)^{-1}, \quad (23)$$

and

$$m_{f_{\theta_{n+1} \rightarrow \theta_n}}(\theta_n) \propto \int m_{f_{\theta_{n+2} \rightarrow \theta_{n+1}}}(\theta_{n+1}) f_{\theta_{n+1}}(\theta_{n+1}, \theta_n) \cdot m_{f_{r_{n+1} \rightarrow \theta_{n+1}}}(\theta_{n+1}) d\theta_{n+1} \propto \mathcal{N}(\theta_n; \hat{\theta}_n^\leftarrow, v_{\theta_n^\leftarrow}), \quad (24)$$

where

$$\hat{\theta}_n^\leftarrow = \frac{(v_{\theta_{n+1}^\leftarrow})^{-1} \hat{\theta}_{n+1}^\leftarrow + (v_{\theta_{n+1}^\downarrow})^{-1} \hat{\theta}_{n+1}^\downarrow}{(v_{\theta_{n+1}^\leftarrow})^{-1} + (v_{\theta_{n+1}^\downarrow})^{-1}}, \quad v_{\theta_n^\leftarrow} = \sigma_\Delta^2 + \left((v_{\theta_{n+1}^\leftarrow})^{-1} + (v_{\theta_{n+1}^\downarrow})^{-1} \right)^{-1}. \quad (25)$$

Given (19), (22) and (24), the belief of θ_n is expressed as ([18, Eq. (3)]):

$$b(\theta_n) \propto m_{f_{r_n \rightarrow \theta_n}}(\theta_n) m_{f_{\theta_n \rightarrow \theta_n}}(\theta_n) m_{f_{\theta_{n+1} \rightarrow \theta_n}}(\theta_n) \propto \mathcal{N}(\theta_n; \hat{\theta}_n, v_{\theta_n}), \quad (26)$$

where

$$v_{\theta_n} = (v_{\theta_n^\rightarrow}^{-1} + v_{\theta_n^\downarrow}^{-1} + v_{\theta_n^\leftarrow}^{-1})^{-1}, \quad \hat{\theta}_n = v_{\theta_n} (v_{\theta_n^\rightarrow}^{-1} \hat{\theta}_n^\rightarrow + v_{\theta_n^\downarrow}^{-1} \hat{\theta}_n^\downarrow + v_{\theta_n^\leftarrow}^{-1} \hat{\theta}_n^\leftarrow). \quad (27)$$

The message passing schedule for our proposed joint PHN estimation and decoding algorithm based on the GAMP method relying on the Wiener model of the OFDM-IM scheme, which is referred to as GAMP-Wiener, is summarized in Algorithm 1.

2) *PHN estimation based on the DCT model:* Inspired by the idea of reducing the number of unknown PHN samples and improving the degree of parallelism, we model the PHN variation as a truncated DCT expansion [25], i.e. the vector $\mathbf{u} \triangleq e^{j\theta}$ is approximately represented by

$$\mathbf{u} = \Phi \boldsymbol{\alpha}, \quad (28)$$

where Φ is the DCT matrix with size $N_q \times L_{\text{PN}}$, and $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_{L_{\text{PN}}}]^T$ collects the DCT coefficients to be

Algorithm 1 The proposed joint PHN estimation and decoding algorithm based on the GAMP method relying on the Wiener model.

- 1: Initialization: $t = 0$ (iteration index)
- 2: For $n = 1, \dots, N_q$, compute \hat{x}_n^0 and $v_{x_n}^0$; set $\hat{\zeta}_n^0 = 0$, $\hat{\eta}_n^0 = r_n$ and $v_{\eta_n}^0 = 2\sigma^2$.
- 3: Set $t = 1$ and repeat the following steps until $\sum_n |\hat{x}_n^{t+1} - \hat{x}_n^t|^2 < \epsilon \sum_n |\hat{x}_n^t|^2$ or $t > T_{\max}$, where ϵ is a pre-specified error tolerance and T_{\max} is the maximum number of iterations.
- 4: – PHN estimation step: For $n = 1, \dots, N_q$, compute $\{\mathbf{v}_\xi^t, \hat{\xi}^t\}$ and $\{\hat{z}_n^t, v_{z_n}^t\}$ from (14), (15) and (10), respectively.
- 5: For $n = 1, \dots, N_q$, compute $\{\hat{\theta}_n^\downarrow, v_{\theta_n^\downarrow}\}$, $\{\hat{\theta}_n^\rightarrow, v_{\theta_n^\rightarrow}\}$, $\{\hat{\theta}_n^\leftarrow, v_{\theta_n^\leftarrow}\}$ and $\{\hat{\theta}_n^t, v_{\theta_n^t}\}$ from (21), (23), (25) and (27).
- 6: – Decoupling step: For $n = 1, \dots, N_q$, compute $\{\hat{\eta}_n^t, v_{\eta_n}^t\}$, $\{\hat{z}_n^t, v_{z_n}^t\}$, $\{\hat{\zeta}_n^t, v_{\zeta_n}^t\}$, $\{\mathbf{v}_\tau^t, \hat{\tau}^t\}$ from (9), (10), (11), (16) and (17), respectively.
- 7: – Decoding step: For $g = 1, \dots, G$, $i = 1, \dots, M_{g1}$, $j = 1, \dots, M_{g2}$, $L^{\text{extr}}(c_{g,i}^{\text{IM}})$ and $L^{\text{extr}}(c_{g,j}^{\text{S}}) \rightarrow \text{Decoder}$.
- 8: Compute (46) by using the information from the decoder.
- 9: – Denoising step: For $n = 1, \dots, N_q$, compute $\{\hat{x}_n^t, v_{x_n}^t\}$ from (5).

estimated¹. In this way, the problem of estimating the N_q dimensional PHN vector has been transformed to the estimation of L_{PN} DCT coefficients. Based on (1) and (28), the received data vector can be alternatively expressed as:

$$\mathbf{r} = \frac{1}{\sqrt{K_q}} \mathcal{D}(\Phi \alpha) \mathbf{F}^H \mathbf{H} \mathbf{x} + \mathbf{w}. \quad (29)$$

The PHN estimation subgraph corresponding to the DCT model is shown on top of the Wiener model factor graph representation in Fig. 2 using dashed lines. The message arriving from the observation node to the DCT coefficient vector α is formulated as:

$$m_{f_{r_n} \rightarrow \alpha}(\alpha) \propto \exp \{E_{b(z_n)} [\log f_{r_n}(\alpha, z_n)]\} \\ \propto \exp \left\{ \frac{2\Re\{r_n^* \hat{z}_n(\Phi \alpha)_n\} - |(\Phi \alpha)_n|^2 (|\hat{z}_n|^2 + v_{z_n})}{2\sigma^2} \right\}.$$

Thus, the *a posteriori* pdf of α is given by

$$b(\alpha) \propto p(\alpha) \prod_{n=1}^{N_q} m_{f_{r_n} \rightarrow \alpha}(\alpha) \\ \propto \exp \left\{ \frac{2\Re\{(\mathbf{r} \odot \hat{\mathbf{z}}^*)^H \Phi \alpha\} - \alpha^H \Phi^H \mathcal{D}(\hat{\rho}_z) \Phi \alpha}{2\sigma^2} \right\},$$

where $\hat{\rho}_z = [|\hat{z}_1|^2 + v_{z_1}, \dots, |\hat{z}_{N_q}|^2 + v_{z_{N_q}}]^T$, while its covariance matrix and mean vector are

$$\mathbf{V}_\alpha = 2\sigma^2 (\Phi^H \mathcal{D}(\hat{\rho}_z) \Phi)^{-1}, \quad \hat{\alpha} = \frac{1}{2\sigma^2} \mathbf{V}_\alpha \Phi^H (\mathbf{r} \odot \hat{\mathbf{z}}^*). \quad (30)$$

Note that the computation of (30) involves a nontrivial matrix

inversion, which has a complexity order of $\mathcal{O}(L_{\text{PN}}^3)$.

Given $b(\alpha)$, the message $m_{f_{r_n} \rightarrow z_n}(z_n)$ is expressed as:

$$m_{f_{r_n} \rightarrow z_n}(z_n) \propto \exp \{E_{b(\alpha)} [\log f_{r_n}(z_n, \alpha)]\} \\ \approx \mathcal{N}_c(z_n; \hat{\eta}_n, v_{\eta_n}), \quad (31)$$

where

$$\hat{\eta}_n = \frac{r_n}{\hat{\rho}_{u_n}} (\Phi \hat{\alpha})_n^*, \quad v_{\eta_n} = 2\sigma^2 / \hat{\rho}_{u_n}, \quad (32)$$

with

$$\hat{\rho}_{u_n} = \sum_{i=1}^{L_{\text{PN}}} |\Phi_{n,i}|^2 (V_{i,i} + |\hat{\alpha}_i|^2) + \sum_{i=1}^{L_{\text{PN}}} \sum_{j=1, j>i}^{L_{\text{PN}}} \Phi_{n,i} \Phi_{n,j}^* \cdot 2\Re(V_{i,j} + \hat{\alpha}_i^* \hat{\alpha}_j). \quad (33)$$

Actually, the complexity of both (30) and (33) is low. As it will be demonstrated by our simulation results, an attractive BER performance can be attained when L_{PN} is selected to be sufficiently small compared to N_q . To further reduce this complexity, we approximate $\mathcal{D}(\hat{\rho}_z)$ by $\hat{\rho}_z \mathbf{I}_{N_q}$, where $\hat{\rho}_z$ is the average of the diagonal elements of $\mathcal{D}(\hat{\rho}_z)$. In this way, \mathbf{V}_α becomes a diagonal matrix, with $2\sigma^2 / \hat{\rho}_z$ being the diagonal element since $\Phi^H \Phi = \mathbf{I}_{L_{\text{PN}}}$.

The message passing schedule for the proposed joint PHN estimation and decoding algorithm based on the GAMP method using the DCT model for the OFDM-IM scheme, which is referred to as GAMP-DCT, is summarized in Algorithm 2.

Algorithm 2 The proposed joint PHN estimation and decoding algorithm based on the GAMP method using the DCT model.

- 1: The same as GAMP-Wiener except for the following modifications imposed on the PHN estimation step and the decoupling step:
- 2: – PHN estimation step: For $n = 1, \dots, N_q$, compute $\{\mathbf{v}_\xi^t, \hat{\xi}^t\}$ and $\{\hat{z}_n^t, v_{z_n}^t\}$ from (14), (15) and (10), respectively.
- 3: Compute $\{\mathbf{V}_\alpha^t, \hat{\alpha}^t\}$ from (30).
- 4: – Decoupling step: For $n = 1, \dots, N_q$, compute $\{\hat{\eta}_n^t, v_{\eta_n}^t\}$, $\{\hat{z}_n^t, v_{z_n}^t\}$, $\{\hat{\zeta}_n^t, v_{\zeta_n}^t\}$, $\{\mathbf{v}_\tau^t, \hat{\tau}^t\}$ from (32), (10), (11), (16) and (17), respectively.

IV. AMALGAMATED BP-MF ALGORITHM

For comparison, we also amalgamate the BP-MF algorithm of [21] and invoke it for the coded OFDM-IM system. The subgraph of the PHN estimation shown in Fig. 2 is also used in the BP-MF method. The difference with regard to the proposed GAMP based algorithms is that the BP-MF technique is invoked for soft demodulation, and the corresponding SISO demodulation subgraph shown in Fig. 2 is replaced by that in Fig. 3. In contrast to the traditional OFDM system, the sparsity of the OFDM-IM symbol should be explicitly exploited by the BP-MF method, which is similar to that of the GAMP based algorithms. However, instead of using central-limit-theorem based approximations (refer to [20] for details of original GAMP) for representing the noiseless output z_n ,

¹The design of the length of vector \mathbf{u} is flexible. The selection of N_q here is just for the sake of simplicity.

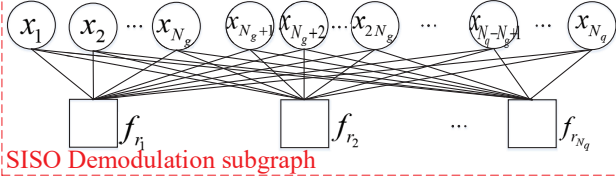


Fig. 3. The subgraph for soft demodulation using the amalgamated BP-MF algorithm.

$n = 1, \dots, N_q$, and assuming them to be independent of each other, we use the *a posteriori* pdf of z_n to estimate the PHN effect $e^{j\theta_n}$, the BP-MF method estimates \mathbf{x} directly and uses the *a posteriori* pdf of this vector for updating the PHN samples' estimates. As shown in Fig. 3, this part of the subgraph is densely connected. Therefore, there are multiple message-passing schedules to select. In the following, both a sequential and a parallel schedule are used. Bearing in mind that the DCT model based PHN estimation method operates in parallel, while the Wiener model based one is in a sequential recursion form, they are naturally embedded in parallel and sequential message-passing schedules, respectively. In both schedules, the extrinsic information extracted from the LDPC decoder is calculated simultaneously.

The message $m_{x_i \rightarrow f_{M,g}}^{\text{BP-MF}}(x_i)$ is formulated as:

$$m_{x_i \rightarrow f_{M,g}}^{\text{BP-MF}}(x_i) \propto \prod_{i=1}^{N_q} m_{f_{r_n} \rightarrow x_i}(x_i) \propto \mathcal{N}_c(x_i; \hat{\underline{\tau}}_i, v_{\underline{\tau}}_i), \quad (34)$$

where $\hat{\underline{\tau}}_i$ and $v_{\underline{\tau}}_i$ will be later updated according to (38) and (43) in the sequential and parallel schedules, respectively. Given the extrinsic information gleaned from the SISO decoder and (34), we can compute the *a posteriori* pdf of x_i , $\underline{b}(x_i)$ according to Appendix A, where $\{\hat{\tau}_i, v_{\tau_i}\}$ is replacing by $\{\hat{\underline{\tau}}_i, v_{\underline{\tau}}_i\}$. The *a posteriori* mean and variance of x_i are obtained as

$$\hat{x}_i = \sum_{s \in S} s \underline{b}(x_i = s), v_{x_i} = \sum_{s \in S} |s|^2 \underline{b}(x_i = s) - |\hat{x}_i|^2. \quad (35)$$

A. Wiener model based sequential BP-MF message passing schedule

According to the MF rule, the message $m_{f_{r_n} \rightarrow x_i}(x_i)$ can be expressed using $\underline{b}(x_i)$ and $b(\theta_n)$ by

$$m_{f_{r_n} \rightarrow x_i}(x_i) \propto \exp \left\{ E_{b(\theta_n)} \prod_{j \neq i} \underline{b}(x_j) [\log f_{r_n}(\mathbf{x}, \theta_n)] \right\} \\ \propto \mathcal{N}_c(x_i; \hat{\tau}_i', v_{\tau_i}'), \quad (36)$$

where

$$v_{\tau_i}' = 2\sigma^2 / |\Lambda_{n,i}|^2, \\ \hat{\tau}_i' = \frac{v_{\tau_i}' \Lambda_{n,i}^*}{2\sigma^2} \left(r_n e^{-j\hat{\theta}_n} (1 - \frac{1}{2} v_{\theta_n}) - \Lambda_{n,:} \hat{\mathbf{x}}_i \right). \quad (37)$$

In (37), $\hat{\mathbf{x}}_i$ is i.e., of the *a posteriori* mean of x_j , $j = 1 \dots N_q, j \neq i$, i.e., $\hat{\mathbf{x}}_i = [\hat{x}_1, \dots, \hat{x}_{i-1}, 0, \hat{x}_{i+1}, \dots, \hat{x}_{N_q}]^T$. Substituting (36) into (34), we have

$$v_{\underline{\tau}}_i = 2\sigma^2 (\Lambda_{:,i}^H \Lambda_{:,i})^{-1}, \hat{\underline{\tau}}_i = \frac{v_{\underline{\tau}}_i}{2\sigma^2} \Lambda_{:,i}^H (\mathcal{D}(\hat{\underline{\theta}}_\theta) \mathbf{r} - \Lambda \hat{\mathbf{x}}_i), \quad (38)$$

where $\hat{\underline{\theta}}_\theta = e^{-j\hat{\theta}} - \frac{1}{2} e^{-j\hat{\theta}} \odot \mathbf{v}_\theta$, $\hat{\theta} = [\hat{\theta}_1, \dots, \hat{\theta}_{N_q}]^T$ and $\mathbf{v}_\theta = [v_{\theta_1}, \dots, v_{\theta_{N_q}}]^T$.

Using the second-order Taylor expansion of $e^{j\theta_n}$ in (20), the message passed on the MF subgraph to the PHN estimation subgraph can be approximated as

$$m_{f_{r_n} \rightarrow \theta_n}^{\text{BP-MF}}(\theta_n) \propto \exp \left\{ E_{\underline{b}(\mathbf{x})} [\log f_{r_n}(\mathbf{x}, \theta_n)] \right\} \\ \propto \left\{ \Re \left[\frac{r_n^* \Lambda_{n,:} \hat{\mathbf{x}}}{\sigma^2} e^{j\theta_n} \right] \right\} \\ \propto \mathcal{N}(\theta_n; \hat{\theta}_n^\perp, v_{\theta_n}^\perp), \quad (39)$$

where

$$v_{\theta_n}^\perp = 1 / \Re \left[\frac{r_n^* \Lambda_{n,:} \hat{\mathbf{x}}}{\sigma^2} e^{j\hat{\theta}_n} \right], \\ \hat{\theta}_n^\perp = v_{\theta_n}^\perp \Re \left[\frac{r_n^* \Lambda_{n,:} \hat{\mathbf{x}}}{\sigma^2} (j + \hat{\theta}_n) e^{j\hat{\theta}_n} \right]. \quad (40)$$

The amalgamated BP-MF using the sequential message passing schedule for the Wiener model based joint PHN estimation and decoding algorithm, which is referred to as BP-MF-s, is summarized in Algorithm 3.

Algorithm 3 Amalgamated BP-MF algorithm for joint PHN estimation and decoding based on the Wiener model using the sequential scheduling.

- 1: Initialization: $t = 0$ (iteration index)
- 2: For $n = 1, \dots, N_q$, compute \hat{x}_n^0 and $v_{x_n}^0$; set $\hat{\theta}_n = 0$.
- 3: Set $t = 1$ and repeat the following steps until $\sum_n |\hat{x}_n^{t+1} - \hat{x}_n^t|^2 < \epsilon \sum_n |\hat{x}_n^t|^2$ or $t > T_{\max}$.
- 4: – PHN estimation step: For $n = 1, \dots, N_q$, compute $\{\hat{\theta}_n^\perp, v_{\theta_n}^\perp\}$, $\{\hat{\theta}_n^\perp, v_{\theta_n}^\perp\}$, $\{\hat{\theta}_n^\perp, v_{\theta_n}^\perp\}$ and $\{\hat{\theta}_n^\perp, v_{\theta_n}^\perp\}$ from (40), (23), (25) and (27), respectively.
- 5: – Soft demodulation step: Compute $\hat{\mathbf{w}}^t = \mathcal{D}(\hat{\underline{\theta}}_\theta) \mathbf{r} - \Lambda \hat{\mathbf{x}}^{t-1}$.
- 6: For $n = 1, \dots, N_q$,
- 7: compute $v_{\underline{\tau}}_n^t$ from (38); compute $\hat{\underline{\tau}}_n^t = \frac{v_{\underline{\tau}}_n^t}{2\sigma^2} \Lambda_{:,n}^H \hat{\mathbf{w}}^t + \hat{x}_n^{t-1}$; compute $\{\hat{x}_n^t, v_{x_n}^t\}$; compute $\hat{\mathbf{w}}^t = \hat{\mathbf{w}}^t - \Lambda_{:,n} (\hat{x}_n^t - \hat{x}_n^{t-1})$.
- 8: End for n .
- 9: – Decoding step: For $n = 1, \dots, N_q$, $\{\hat{\underline{\tau}}_n^t, v_{\underline{\tau}}_n^t\} \rightarrow$ Decoder.
- 10: Compute (46) by using the extrinsic information from the decoder.
- 11: – Denoising step: For $n = 1, \dots, N_q$, compute $\{\hat{x}_n^t, v_{x_n}^t\}$ from (35); $\hat{\mathbf{x}}^t = [\hat{x}_1^t, \dots, \hat{x}_{N_q}^t]^T$.

B. DCT model based parallel BP-MF message passing schedule

Given $\underline{b}(x_i)$ and $b(\alpha)$ from the previous iteration, the message $m_{f_{r_n} \rightarrow x_i}(x_i)$ is formulated based on the MF method by

$$m_{f_{r_n} \rightarrow x_i}(x_i) \propto \exp \left\{ E_{b(\alpha)} \prod_{j \neq i} \underline{b}(x_j) [\log f_{r_n}(\mathbf{x}, \alpha)] \right\} \\ \propto \mathcal{N}_c(x_i; \hat{\tau}_i', v_{\tau_i}'), \quad (41)$$

where we have:

$$v_{\tau_i}' = 2\sigma^2 / (\hat{\theta}_{u_n} |\Lambda_{n,i}|^2), \\ \hat{\tau}_i' = \frac{v_{\tau_i}' \hat{\theta}_{u_n} \Lambda_{n,i}^*}{2\sigma^2} [(\Phi \hat{\alpha})_n^* r_n / \hat{\theta}_{u_n} - \Lambda_{n,:} \hat{\mathbf{x}}_i], \quad (42)$$

and the message $m_{x_i \rightarrow f_{M,g}}^{\text{BP-MF}}(x_i) \propto \mathcal{N}_c(x_i; \hat{\mathbf{r}}_i, v_{\mathbf{r}_i})$ is associated with

$$v_{\mathbf{r}_i} = 2\sigma^2 (\mathbf{\Lambda}_{:,i}^H \mathcal{D}(\hat{\mathbf{r}}_u) \mathbf{\Lambda}_{:,i})^{-1},$$

$$\hat{\mathbf{r}}_i = \frac{v_{\mathbf{r}_i}}{2\sigma^2} \mathbf{\Lambda}_{:,i}^H \mathcal{D}(\hat{\mathbf{r}}_u) (\mathcal{D}^H(\Phi \hat{\alpha}) \mathcal{D}^{-1}(\hat{\mathbf{r}}_u) \mathbf{r} - \mathbf{\Lambda} \hat{\mathbf{x}}_i), \quad (43)$$

where $\hat{\mathbf{r}}_u = [\hat{r}_{u1}, \dots, \hat{r}_{uN_q}]^T$.

The message emerging from the observation node to the PHN estimation subgraph can be expressed as:

$$m_{f_{r_n} \rightarrow \alpha}(\alpha) \propto \exp \left\{ E_b(\mathbf{x}) [\log f_{r_n}(\alpha, \mathbf{x})] \right\}$$

$$\propto \exp \left\{ \frac{2\Re\{r_n^* \mathbf{\Lambda}_{n,:} \hat{\mathbf{x}}(\Phi \alpha)_n\} - |(\Phi \alpha)_n|^2 \hat{\varrho}_{x_n}}{2\sigma^2} \right\},$$

where $\hat{\varrho}_{x_n} = \sum_{i=1}^{N_q} |\mathbf{\Lambda}_{n,i}|^2 (|\hat{x}_i|^2 + v_{x_i}) + 2 \sum_{i=1}^{N_q} \sum_{j>i}^{N_q} \Re\{\mathbf{\Lambda}_{n,i}^* \mathbf{\Lambda}_{n,j} \hat{x}_i^* \hat{x}_j\}$. Thus, the *a posteriori* pdf of the DCT coefficient vector is given by

$$b(\alpha) \propto p(\alpha) \prod_{n=1}^{N_q} m_{f_{r_n} \rightarrow \alpha}(\alpha)$$

$$\propto \exp \left\{ \frac{2\Re\{(\mathbf{r} \odot (\mathbf{\Lambda} \hat{\mathbf{x}})^*)^H \Phi \alpha\} - \alpha^H \Phi^H \mathcal{D}(\hat{\varrho}_x) \Phi \alpha}{2\sigma^2} \right\},$$

where $\hat{\varrho}_x = [\hat{\varrho}_{x1}, \dots, \hat{\varrho}_{xN_q}]^T$, and its covariance matrix and mean vector are approximated as

$$\mathbf{V}_\alpha = 2\sigma^2 / \bar{\varrho}_x \mathbf{I}_{L_{PN}}, \quad \hat{\alpha} = \frac{1}{\bar{\varrho}_x} \Phi^H (\mathbf{r} \odot (\mathbf{\Lambda} \hat{\mathbf{x}})^*), \quad (44)$$

where $\bar{\varrho}_x$ is the average of the elements in $\hat{\varrho}_x$.

The amalgamated BP-MF using the parallel message passing schedule for the DCT model based joint PHN estimation and decoding algorithm, which is referred to as BP-MF-p, is summarized in Algorithm 4.

Algorithm 4 Amalgamated BP-MF algorithm for joint PHN estimation and decoding based on the DCT model using the parallel scheduling.

- 1: Initialization: $t = 0$ (iteration index)
- 2: For $n = 1, \dots, N_q$, compute \hat{x}_n^0 and $v_{x_n}^0$.
- 3: Set $t = 1$ and repeat the following steps until $\sum_n |\hat{x}_n^{t+1} - \hat{x}_n^t|^2 < \epsilon \sum_n |\hat{x}_n^t|^2$ or $t > T_{\max}$.
- 4: – PHN estimation step: Compute $\{\mathbf{V}_\alpha^t, \hat{\alpha}^t\}$ from (44).
- 5: – Soft demodulation step: Compute $\hat{\mathbf{w}}^t = \mathcal{D}^H(\Phi \hat{\alpha}^t) \mathcal{D}^{-1}(\hat{\varrho}_u^t) \mathbf{r} - \mathbf{\Lambda} \hat{\mathbf{x}}^{t-1}$.
- 6: For $n = 1, \dots, N_q$, compute $v_{\mathbf{r}_n}^t$ from (43); compute $\hat{\mathbf{r}}_n^t = \frac{v_{\mathbf{r}_n}^t}{2\sigma^2} \mathbf{\Lambda}_{:,n}^H \mathcal{D}(\hat{\varrho}_u^t) \hat{\mathbf{w}}^t + \hat{x}_n^{t-1}$.
- 7: – Decoding step: For $n = 1, \dots, N_q$, $\{\hat{\mathbf{r}}_n^t, v_{\mathbf{r}_n}^t\} \rightarrow \text{Decoder}$.
- 8: Compute (46) by using the extrinsic information from the decoder.
- 9: – Denoising step: For $n = 1, \dots, N_q$, compute $\{\hat{\mathbf{x}}_n^t, v_{\mathbf{x}_n}^t\}$ from (35); $\hat{\mathbf{x}}^t = [\hat{\mathbf{x}}_1^t, \dots, \hat{\mathbf{x}}_{N_q}^t]^T$.

of the standard BCJR decoding, thus we only analyze the computational complexity of a single iteration between the soft demodulator and PHN estimator within an OFDM-IM block. For soft demodulation by the GAMP-Wiener scheme, the order of $\mathcal{O}(GN_qS)$ operations are needed for the calculation of (52) and (53) (see Appendix A), plus an N_q -point IFFT for evaluating (15) and an N_q -point FFT for evaluating (17). We point out that the complexity in this part grows linearly both with the number of IM symbols and the modulation order, while that in the optimal receiver has a complexity order of $\mathcal{O}(2^{GM_g})$. The PHN estimation, i.e. the computation of (19) and the messages in the PHN estimation subgraph based on the Wiener model and on the second-order Taylor expansion approximation is in the order of $\mathcal{O}(N_q)$. The GAMP-Wiener and GAMP-DCT scheme differ in the PHN estimation part. Explicitly, in GAMP-DCT, $\mathcal{O}(2N_qL_{PN})$ operations are required for the calculation of (30) and (32) following from the assumption of $\mathcal{D}(\hat{\varrho}_z) \approx \hat{\varrho}_z \mathbf{I}_{N_q}$. The BP-MF-s differs from the GAMP-Wiener scheme in the soft demodulation, where the complexity is increased from $\mathcal{O}(N_q \log N_q)$ to $\mathcal{O}(N_q^2)$, since the FFT cannot be directly invoked, when the messages are updated sequentially. By contrast, the BP-MF-p and GAMP-DCT have similar complexity in this part using the FFT, when passing messages in parallel. We also compare the complexity of the proposed algorithms to that of the conventional pilot based PHN estimation method of [35] combined with the conventional OFDM-IM detector of [4]. The calculations of a frequency-domain received sequence need an N_q -point FFT. The PHN estimation part in the pilot based method requires $\mathcal{O}(N_{\text{pilot}})$ operations, where N_{pilot} is the total number of pilots contained within an OFDM-IM packet. However, the complexity of these calculations is modest, which have already been performed before starting iterative decoding. After compensating the received samples using the PHN estimates, the conventional LLR detector of [4] is employed for soft demodulation, which has a complexity order of $\mathcal{O}(GM_gS)$. Furthermore, the existing PHN tracking algorithms of [15] and [16] are combined with the proposed demodulation method and the complexity of PHN estimation is studied for comparison with both its Wiener model and the DCT model based counterparts. For the EKF of [15], similar PHN estimation complexity is seen as that in the GAMP-Wiener scheme due to the local linearization of the system model using first-order Taylor expansion. The VI approach of [16] involves an inversion of an $(N_q \times N_q)$ -element matrix, which imposes a complexity order of $\mathcal{O}(N_q^3)$. In Table. 1, we give an example for $N_q = 64$, $L_{PN} = 2$, $G = 16$, $S = 4$, and $M_g = 4$. We observe that the proposed PHN estimation algorithms are computationally more efficient compared to the VI approach, for example, by a factor of 4096 at $N_q = 64$, when we employ the GAMP-Wiener method.

V. SIMULATION RESULTS

In this section, simulation results are provided for evaluating the performance of the proposed joint PHN estimation and decoding algorithms for the OFDM-IM scheme and for comparing it with conventional OFDM. For all simulations, a rate-

C. Complexity Analysis

In Table 1, we provide a brief complexity comparison of the proposed GAMP-Wiener, GAMP-DCT, BP-MF-s and BP-MF-p algorithms. All of these methods involve the calculations

TABLE I
COMPARISONS OF COMPUTATIONAL COMPLEXITY (FOR EXAMPLE: $N_q = 64$, $L_{PN} = 2$, $G = 16$, $S = 4$, AND $M_g = 4$.)

Algorithm	Complexity of the demodulator	Value	Complexity of the PHN estimator	Value
GAMP-Wiener	$\mathcal{O}(GN_gS + 2N_q \log N_q)$	$\mathcal{O}(1024)$	$\mathcal{O}(N_q)$	$\mathcal{O}(64)$
GAMP-DCT	$\mathcal{O}(GN_gS + 2N_q \log N_q)$	$\mathcal{O}(1024)$	$\mathcal{O}(2N_q L_{PN})$	$\mathcal{O}(256)$
BP-MF-s	$\mathcal{O}(GN_gS + N_q \log N_q + N_q^2)$	$\mathcal{O}(4736)$	$\mathcal{O}(N_q)$	$\mathcal{O}(64)$
BP-MF-p	$\mathcal{O}(GN_gS + 2N_q \log N_q)$	$\mathcal{O}(1024)$	$\mathcal{O}(2N_q L_{PN})$	$\mathcal{O}(256)$
Optimal Detector	$\mathcal{O}(2^{GM_g})$	$\mathcal{O}(1.8 \times 10^{19})$	—	—
EKF [15]	—	—	$\mathcal{O}(N_q)$	$\mathcal{O}(64)$
VI [16]	—	—	$\mathcal{O}(N_q^3)$	$\mathcal{O}(2.6 \times 10^5)$

0.66 (i.e., $R_c = 0.66$) LDPC code² using a block size 8064 is used. The maximum number of iterations in the LDPC decoder is 50, while it is $T_{\max} = 20$ between the SISO demodulation and SISO decoding, unless otherwise stated. The tolerance value is set to be $\epsilon = 10^{-12}$. The number of subcarriers in an OFDM-IM symbol is kept constant at 64, i.e., $N_q = 64$. A ten-tap (i.e., $L_h = 10$) Rayleigh fading channel is considered. The CIRs are circularly symmetric complex Gaussian random variables following the distribution $\mathcal{N}(h_l; 0, 1/L_h)$. The CP length L_{CP} is selected to be 13. The SNR is expressed as E_b/N_0 , with $E_b = \frac{N_q + L_{CP}}{R_c G(M_{g1} + M_{g2})}$ being the energy per bit. The effective throughput is $R_c \frac{G(M_{g1} + M_{g2})}{N_q + L_{CP}}$ [bits/symbol]. For characterizing the PHN, a Wiener model with $\sigma_\Delta = 0.5^\circ$ and a severer one with $\sigma_\Delta = 3^\circ$ are used for generating PHN samples. The initial phase θ_1 is uniformly distributed in $[0, 2\pi)$. The number of DCT coefficients L_{PN} is designed to be 2, unless otherwise specified. A pilot is inserted every 64 subcarriers to assist the iterative decoder to bootstrap and to aid synchronization³. Bearing in mind that OFDM-IM scheme tends to exhibit benefits over its OFDM counterpart when the throughput is below 2 [bits/symbol] [12], we only consider low-order modulation schemes⁴ such as binary phase shift keying (BPSK), i.e., $S = 2$, and quadrature phase shift keying (QPSK), i.e., $S = 4$.

Fig. 4 shows the BERs of the proposed GAMP-Wiener and GAMP-DCT algorithms for different OFDM-IM schemes characterized by $\{S = 4, N_g = 4, K_g = 1\}$ (denoted by “(4, 4, 1)”) and $\{S = 2, N_g = 4, K_g = 2\}$ (denoted by “(2, 4, 2)”), respectively. For comparison, the BER curves of classic BPSK modulated OFDM (denoted by “OFDM BPSK”), yielding the same transmission rate as that of OFDM-IM, are also plotted. The BERs of the OFDM-IM and OFDM schemes with known PHN, denoted by “known PHN”, are also shown as benchmarks. A constellation rotation technique with a rotation angle of $\pi/12$ and the coordinate interleaved orthogonal design (CIOD) approach of [14] are used for improving the diversity gain. Additionally, for OFDM-IM with known PHN, the subcarrier-level interleaving scheme

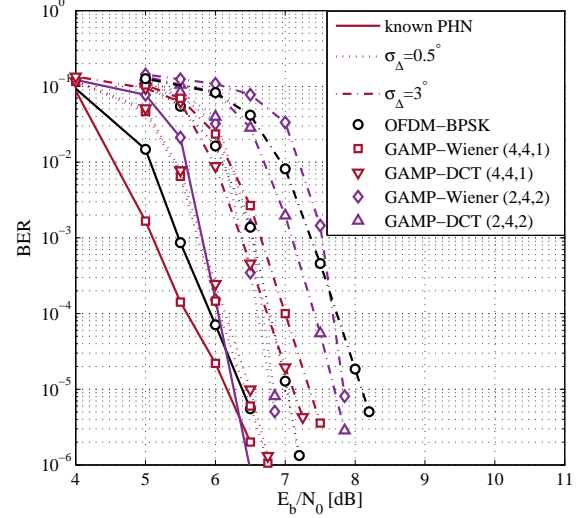


Fig. 4. BER performance of classic BPSK-modulated OFDM and OFDM-IM (characterized by (4,4,1) and (2,4,2)) systems communicating over a 10-tap Rayleigh fading channel affected by PHN. Wiener PHN model with standard deviations of $\sigma_\Delta = 0.5^\circ$ and 3° are considered. Both proposed GAMP-Wiener and GAMP-DCT are evaluated. The effective throughput is 0.55 bits/symbol.

of [13] is also invoked for attaining multipath diversity by reducing the correlation amongst the subcarriers 1 to N_q . As shown in Fig. 4, in case of known PHN, OFDM-IM outperforms its OFDM counterpart in the high-SNR region, in agreement with the findings of [4], [12], [13]. This is because the performance advantage of the bits transmitted in the IM domain is more significant at high SNRs, especially for $\sigma_\Delta = 0.5^\circ$. Specifically, the performance loss of both the GAMP-Wiener and GAMP-DCT algorithms recorded in case of known PHN is modest, which confirms the efficiency of the Wiener model and DCT model based PHN estimation methods, when $\sigma_\Delta = 0.5^\circ$. By contrast, about 0.7dB gap is seen for classic OFDM, if the target BER is 10^{-5} . This means that using the proposed GAMP based algorithms, the OFDM-IM schemes transmitting part of the information bits by the indices of activated subcarriers indeed succeed in improving the performance over classic OFDM with PHN distortion. For $\sigma_\Delta = 3^\circ$, the GAMP-DCT is slightly superior to GAMP-Wiener for OFDM-IM (4, 4, 1) at a slightly higher complexity. This indicates that the accuracy of the Taylor expansion in the GAMP-Wiener receiver degrades upon increasing the PHN intensity. Meanwhile, we point out that the correlation between the PHN samples is not exploited across different OFDM-IM symbols, which is also a reason for the performance degradation of GAMP-Wiener. Explicitly, the mean values of the messages $m_{f_{\theta_n \rightarrow \theta_n}}(\theta_n)$ and $m_{f_{\theta_{n+1} \rightarrow \theta_n}}(\theta_n)$ at the boundaries

²The variable and check node degree distributions (refer to [36] for details) are $X_1 = 0.00003 + 0.1909X + 0.4070X^2 + 0.4020X^7$ and $X_2 = 0.5029X^9 + 0.4971X^{10}$, respectively. The parameter definitions follow [36].

³The power overhead due to the pilots is included in the simulation results. The gap between cases with known PHN and unknown PHN only comes from different algorithms and is not due to pilots, since the increase of required SNR is added artificially in known PHN cases.

⁴As evidenced by our additional extensive simulations, the BER performance advantage of the OFDM-IM scheme over its classic OFDM counterpart tends to erode for higher order modulations.

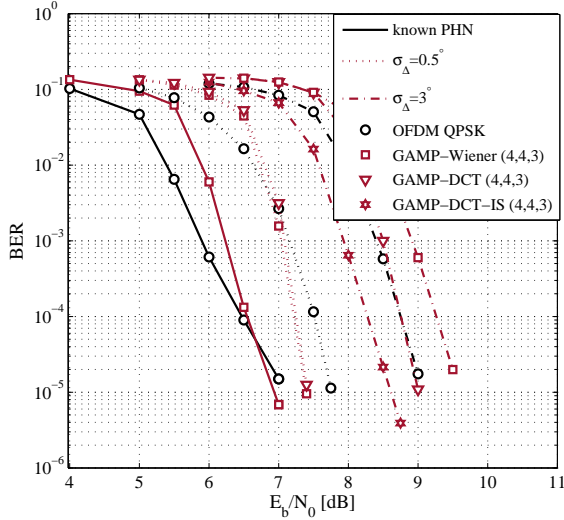


Fig. 5. BER performance of classic QPSK-modulated OFDM and OFDM-IM (characterized by (4,4,3)) systems communicating over a 10-tap Rayleigh fading channel affected by PHN. Wiener PHN model with standard deviations of $\sigma_{\Delta} = 0.5^{\circ}$ and 3° are considered. Both proposed GAMP-Wiener and GAMP-DCT are evaluated. The effective throughput is 1.1 bits/symbol.

of an OFDM-IM symbol are initialized with zeros, which will introduce a transient for the PHN estimation, especially when N_q is not long enough.

In Fig. 5, we compare the BER performance of the OFDM-IM (4,4,3) and classic QPSK modulated OFDM both at an effective throughput of 1.1 bits/symbol. As seen from Fig. 5, the proposed GAMP-Wiener achieves better BER performance than classic OFDM beyond a certain SNR when $\sigma_{\Delta} = 0.5^{\circ}$. This trend is similar to that in Fig. 4. However, this performance advantage diminishes for a relatively severe PHN of $\sigma_{\Delta} = 3^{\circ}$. This can be explained by the fact that the accuracy of data detection and PHN estimation is reduced upon increasing the PHN intensity, since more ICI is introduced. To improve the BER performance of OFDM-IM (4,4,3) for $\sigma_{\Delta} = 3^{\circ}$, we add a depth- N_q subcarrier interleaver to rearrange the contaminated subcarriers in frequency domain. The BER curve of this interleaving aided GAMP-DCT method is labelled by “GAMP-DCT-IS”. Specifically, the interleaved frequency domain signals can be expressed as $\tilde{\mathbf{x}} = [x_1(1), x_2(1), \dots, x_G(1), \dots, x_1(N_q), \dots, x_G(N_q)]^T$. Accordingly, after de-interleaving at the receiver, we have

$$\mathbf{r} = \frac{1}{\sqrt{K_q}} \mathcal{D}(e^{j\theta}) \tilde{\mathbf{F}}^H \tilde{\mathbf{H}} \mathbf{x} + \mathbf{w}, \quad (45)$$

where $\tilde{\mathbf{F}}^H$ and $\tilde{\mathbf{H}}$ are de-interleaved versions of \mathbf{F}^H and \mathbf{H} , respectively. In this way, the performance degradation of IM-modulated bits due to the frequency-domain correlation is reduced. It is shown that as expected, about 0.4dB gain is achieved at a BER of 10^{-5} for the GAMP-DCT-IS over GAMP-DCT. However, the OFDM-IM having realistic unknown PHN and using the interleaving loses the low complexity advantage of the FFT, which is in contrast to the known PHN case due to the matrix $\tilde{\mathbf{F}}^H$ in (45). Using Equations (6), (7), (12) and (13) instead of (14), (15), (16) and (17), the complexity is increased from $\mathcal{O}(N_q \log N_q)$ to $\mathcal{O}(N_q^2)$. Hence, there is a flexible performance vs. complexity tradeoff in the

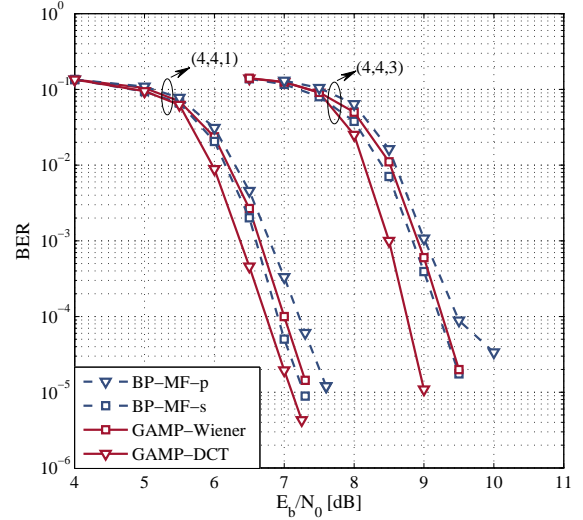


Fig. 6. BER performance of the proposed GAMP-DCT, GAMP-Wiener and BP-MF algorithms. Both parallel and sequential schedules for the BP-MF methods are studied for comparison. Wiener PHN model with $\sigma_{\Delta} = 3^{\circ}$ is considered. The effective throughput of OFDM-IM (4,4,1) and (4,4,3) is 0.55 bits/symbol and 1.1 bits/symbol, respectively.

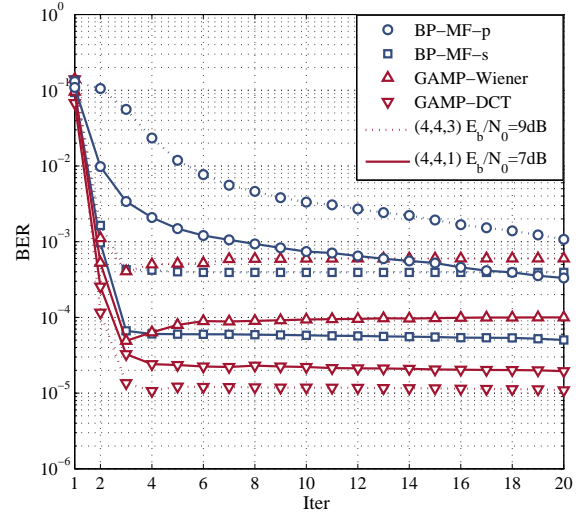


Fig. 7. BER performance of the proposed GAMP-DCT, GAMP-Wiener and BP-MF algorithms versus iteration index. Both parallel and sequential schedules for the BP-MF methods are studied for comparison. Wiener PHN model with $\sigma_{\Delta} = 3^{\circ}$ is considered. The effective throughputs of OFDM-IM (4,4,1) and (4,4,3) are 0.55 bits/symbol and 1.1 bits/symbol, respectively.

GAMP based OFDM-IM receiver.

In Fig. 6, the BER performance of the BP-MF-s and BP-MF-p based receivers is compared to that of the proposed GAMP-Wiener and GAMP-DCT algorithms at a PHN of $\sigma_{\Delta} = 3^{\circ}$. The OFDM-IM schemes characterized by (4,4,1) and (4,4,3) and having throughputs of 0.55 bits/symbol and 1.1 bits/symbol are considered. The number of turbo iterations between SISO demodulation and SISO decoding is fixed to $T_{\max} = 20$. We observe that GAMP-DCT exhibits the best BER amongst the aforementioned algorithms. Compared with GAMP-Wiener, BP-MF-s achieves slightly lower BER values, which is attributed to exploiting the *a posteriori* estimates of \mathbf{x}_g in the estimation of the remaining symbols during iteration t . By contrast, a significant performance loss is seen for the BP-MF-p method, especially for OFDM-IM (4,4,3).

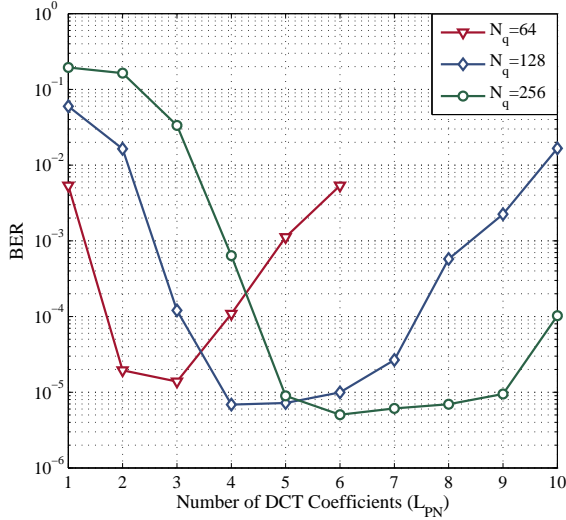


Fig. 8. The impact of the number of DCT coefficients on BER performance of the GAMP-DCT algorithm at $E_b/N_0 = 7$ dB. The block length of DCT model is fixed at 64, 128, and 256, respectively. Wiener PHN model with $\sigma_\Delta = 3^\circ$ is considered.

For example, at a BER of 10^{-5} , about 0.6dB loss is observed for BP-MF-p over GAMP-DCT for OFDM-IM (4, 4, 1), but an error floor around the BER value of 10^{-4} for OFDM-IM (4, 4, 3). Furthermore, Fig. 7 plots the BERs of the receivers as a function of the iteration index under the conditions of Fig. 6. The SNR is fixed to 7dB (dotted line) and 9dB (cont. line) for OFDM-IM (4, 4, 1) and OFDM-IM (4, 4, 3), respectively. Observe that the GAMP-Wiener, GAMP-DCT and BP-MF-s algorithms all converge after about 5 iterations. Therefore, 5 iterations on average are sufficient to meet the pre-specified stopping criterion. By contrast, the convergence of BP-MF-p is slow compared to both its sequential counterpart and to the GAMP based receivers, which explains its poor BER performance at high SNRs in Fig. 6. We also observe that although the complexity of GAMP-Wiener is lower than that of BP-MF-s and its performance is inferior to BP-MF-s after convergence, the BER of GAMP-Wiener is slightly lower than that of BP-MF-s at the first few iterations.

In Fig. 8, the BER performance vs. the number of DCT coefficients is studied. Observing that similar to the phenomenon shown in [25] for a single-carrier system, given the OFDM-IM symbol length, there is an appropriate number of DCT coefficients. After optimizing the relationship between the number of DCT coefficients and OFDM-IM symbol length, we find that a slightly better BER performance can be attained for higher N_q .

The proposed algorithm relying on residual CFO $\Delta\phi$ and imperfect channel information is evaluated in Fig. 9. In practical systems, the channel estimation error at the receiver is assumed to be a zero-mean Gaussian variable with a variance of v_{ϵ_h} [38]. Fig. 9 compares the BER performance of the proposed GAMP-DCT both for OFDM-IM and for classic OFDM, when $\Delta\phi = 0.1$ [39], $v_{\epsilon_h} = 0.01$ and $\sigma_\Delta = 3^\circ$. The simulation settings are kept the same as in Fig. 4. In order to take the channel uncertainty and the residual CFO into account, the proposed algorithm is modified by recalculating the equivalent mean and power of the residual

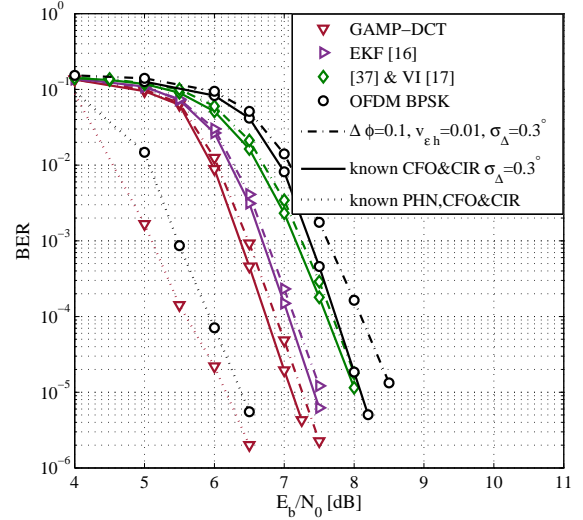


Fig. 9. The impact of residual CFO of $\Delta\phi = 0.1$ and imperfect CSI with parameter $v_{\epsilon_h} = 0.01$ on BER performance of the GAMP-DCT algorithm, the existing EKF [15], and the [37] combined with VI [16]. Wiener PHN model with $\sigma_\Delta = 3^\circ$ is considered.

interference imposed by residual CFO and imperfect channel estimation using the PHN estimates and symbols estimates at each iteration. Additionally, the BER curves of the pilot based joint CFO, CIR and PHN estimation methods of [15] (denoted by “EKF [15]”) and of [37] are now combined with the PHN tracking algorithm of [16] (denoted by “[37] & VI [16]”), which are shown for comparison. We observe that the proposed GAMP-DCT method outperforms the existing pilot based EKF and [37] & VI, even in the presence of residual CFO and channel estimation errors. Moreover, compared to classic OFDM, OFDM-IM relying on the proposed GAMP-DCT method is more robust both to the residual CFO and to the channel uncertainty. For example, for $\Delta\phi = 0.1$, $v_{\epsilon_h} = 0.01$, and $\sigma_\Delta = 3^\circ$, only a modest performance gap of 0.15dB is seen for OFDM-IM compared to the perfect CFO and CSI based BER curve at $\text{BER} = 10^{-5}$, while 0.4dB is observed for classic OFDM. However, this gap widens at higher SNR values.

In Fig. 10, the BER performance of the proposed GAMP-Wiener and GAMP-DCT algorithms is compared to existing methods. The conventional pilot based PHN estimation method of [35] is now combined with the conventional OFDM-IM detector of [4] (denoted by “pilot based & [4]”). We consider different ratios of pilots, e.g., 6%, 12.5% and 25%. The extended Kalman filter method (denoted by “EKF [16]”) and the variational inference approach (denoted by “VI [17]”) were proposed for joint PHN estimation and uncoded OFDM data detection in [15] and [16], respectively. These are now extended to the coded OFDM-IM scheme by combining it with the proposed GAMP based data detection algorithm. The system parameters are the same as in Fig. 5, where two different PHN variances, i.e., $\sigma_\Delta = 0.5^\circ$ and 3° are considered. Additionally, the BER performance recorded without PHN estimation (denoted by “No PHN cancel.”) is also shown for comparison. It is seen that the OFDM-IM performance degrades significantly, if the effect of PHN is ignored. The conventional pilot based method improves the BER performance by increasing the ratio

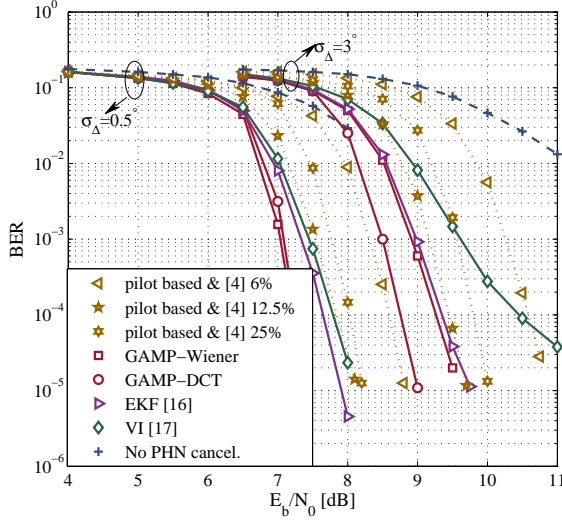


Fig. 10. The BER performance of proposed GAMP-Wiener and GAMP-DCT compared with existing EKF [15] and VI [16] algorithms. The OFDM-IM is characterized by $(4, 4, 3)$. Wiener PHN model with $\sigma_{\Delta} = 0.5^{\circ}$ and 3° are considered. The effective throughput is 1.1 bits/symbol.

of pilots from 6% to 12.5%. However, a performance erosion is imposed by further increasing this ratio, since due to the power assigned to the pilots, the E_b/N_0 reduction becomes significant. We also observe that the pilot based algorithm is inferior to other data-aided methods, due to the fact that the PHN varies rapidly compared the OFDM-IM symbol rate and hence the PHN cannot be sufficiently mitigated using pilots. By using the proposed GAMP-Wiener and GAMP-DCT, the BER performance improves considerably even for severe PHN. The GAMP-DCT scheme outperforms the others due to the fact that the first L_{PN} DCT coefficients are well suited for the compact representation of the lowpass process of discrete-time PHN samples [25]⁵. However, the VI method exhibits poor BER performance compared to the proposed ones and the EKF, especially at high SNRs, where the performance is mainly limited by the PHN estimation accuracy instead of the additive white noise. Explicitly, it suffers from an error floor when $\sigma_{\Delta} = 3^{\circ}$. This outcome can be attributed to the first-order Taylor expansion approximation of the PHN vector over the whole OFDM-IM symbol. The linearization approximation directly imposed on the system model substantially deviates from the true system parameters in the presence of strong PHN. By contrast, both the proposed GAMP-Wiener method and the EKF updates the current PHN sample estimation by using previous PHN estimates. The essential difference between the EKF and the GAMP-Wiener is that the EKF linearizes the system model locally using the first-order Taylor series, while the GAMP-Wiener approximates the messages exchanged in the factor graph by second-order Taylor expansion. This explains the superiority of the GAMP-Wiener over the EKF for two different PHN variance values.

VI. CONCLUSIONS

This paper proposes several low-complexity joint PHN estimation and decoding algorithms for OFDM-IM operating

in frequency-selective fading channels. Two models are considered for approximating the PHN process. The Wiener model exploits the time dependent statistics, while the DCT model dramatically reduces the number of unknown variables. Based on these pair of PHN models, a factor graph is constructed and the GAMP method is invoked, where the unknown nonlinear transform matrix problem imposed by PHN is circumvented by using MF approximation. For comparison, the amalgamated BP-MF methods are presented for soft demodulation and for both sequential as well as for parallel message passing schedules. The performance vs. complexity of the proposed algorithms is compared to that of conventional pilot based method, the existing EKF and VI approach. Our numerical results show that the proposed GAMP-DCT method has the best BER performance at a slightly increased computational complexity, as shown in Table 1. We also observe that OFDM-IM outperforms its OFDM counterpart even for strong PHN, imperfect CSI and residual CFO.

APPENDIX A

DERIVATION OF THE *a posteriori* PDF $b(x_i)$

According to the associated IM and PSK/QAM modulation constraint, we have

$$P(\mathbf{x}_g^{\text{AP}} = \boldsymbol{\omega}_{\gamma}) = \prod_{i=1}^{M_{g1}} P(c_{g,i}^{\text{IM}} = c_i^{\omega_{\gamma}}), \gamma = 1, \dots, \Upsilon,$$

$$P(x_{g,d}(k) = s_m) = \prod_{i=1}^{\log_2 S} P(c_{g,(k-1)\log_2 S + i}^S = c_i^{s_m}),$$

$$m = 1, \dots, S, k = 1, \dots, K_g. \quad (46)$$

Thus, the *a priori* pdf of \mathbf{x} provided by the LDPC SISO decoder can be formulated as

$$P(\mathbf{x}) = \prod_{g=1}^G P_{\mathbf{x}_g}(\mathbf{x}_g),$$

$$P_{\mathbf{x}_g}(\mathbf{x}_g) = \sum_{\gamma=1}^{\Upsilon} P(\mathbf{x}_g^{\text{AP}} = \boldsymbol{\omega}_{\gamma}) \left[\prod_{i \in \mathcal{I}^{\gamma}} \left(\sum_{s \in \mathcal{S}} \delta(x_i - s) P(x_i) \right) \cdot \prod_{j \in \bar{\mathcal{I}}^{\gamma}} \delta(x_j) \right], \quad (47)$$

where $i, j \in I_g \triangleq \{(g-1)N_g + 1 : gN_g\}$, \mathcal{I}^{γ} is the support set of vector $\boldsymbol{\omega}_{\gamma}$, i.e., $\mathcal{I}^{\gamma} = \{i | \omega_{\gamma,i} = 1, i \in I_g\}$, $\bar{\mathcal{I}}^{\gamma}$ is the complementary set of \mathcal{I}^{γ} , and $x_i \in \{x_{g,d}(1), \dots, x_{g,d}(K_g)\}$.

Given (47) and the message $m_{x_i \rightarrow f_{M,g}}(x_i) \propto \mathcal{N}_c(x_i; \hat{\tau}_i, v_{\tau_i})$, which will be later updated according to (12) and (13), we can calculate the *a posteriori* pdf $b(x_i)$ of x_i by (48) for $x_i = s_m, s_m \in \mathcal{S}$, and by (49) for $x_i = 0$. After further algebra in manipulations, we arrive at

$$b(x_i = s_m) = \frac{D(x_i = s_m)}{\sum_{s \in \mathcal{S}} D(x_i = s) + E}, \quad (50)$$

for $x_i = s_m, s_m \in \mathcal{S}$, and

$$b(x_i = 0) = \frac{E}{\sum_{s \in \mathcal{S}} D(x_i = s) + E}, \quad (51)$$

⁵Similar applications of DCT expansion technique can be found in various audio and image compression algorithms.

$$b(x_i = s_m) = \sum_{\gamma: i \in \mathcal{I}^\gamma} P(\mathbf{x}_g^{\text{IM}} = \boldsymbol{\omega}_\gamma) \exp \left\{ -\frac{|s_m - \hat{\vartheta}_i|^2}{v_{\hat{\vartheta}_i}} \right\} P(x_i = s_m) \prod_{j \in \mathcal{I}^\gamma, j \neq i} \left(\sum_{s \in \mathcal{S}} \exp \left\{ -\frac{|s - \hat{\vartheta}_j|^2}{v_{\hat{\vartheta}_j}} \right\} P(x_j = s) \right) \prod_{k \in \mathcal{I}^\gamma} \exp \left\{ -\frac{|\hat{\vartheta}_k|^2}{v_{\hat{\vartheta}_k}} \right\}. \quad (48)$$

$$b(x_i = 0) = \exp \left\{ -\frac{|\hat{\vartheta}_i|^2}{v_{\hat{\vartheta}_i}} \right\} \cdot \sum_{\gamma: i \in \mathcal{I}^\gamma} P(\mathbf{x}_g^{\text{IM}} = \boldsymbol{\omega}_\gamma) \prod_{j \in \mathcal{I}^\gamma, j \neq i} \exp \left\{ -\frac{|\hat{\vartheta}_j|^2}{v_{\hat{\vartheta}_j}} \right\} \prod_{k \in \mathcal{I}^\gamma} \left(\sum_{s \in \mathcal{S}} \exp \left\{ -\frac{|s - \hat{\vartheta}_k|^2}{v_{\hat{\vartheta}_k}} \right\} P(x_k = s) \right). \quad (49)$$

$$D(x_i = s_m) = \sum_{\gamma: i \in \mathcal{I}^\gamma} P(\mathbf{x}_g^{\text{IM}} = \boldsymbol{\omega}_\gamma) \exp \left\{ -\frac{|s_m|^2 - 2\Re\{s_m \hat{\vartheta}_i^*\}}{v_{\hat{\vartheta}_i}} \right\} P(x_i = s_m) \prod_{j \in \mathcal{I}^\gamma, j \neq i} \left(\sum_{s \in \mathcal{S}} \exp \left\{ -\frac{|s|^2 - 2\Re\{s \hat{\vartheta}_j^*\}}{v_{\hat{\vartheta}_j}} \right\} P(x_j = s) \right). \quad (52)$$

$$E = \sum_{\gamma: i \in \mathcal{I}^\gamma} P(\mathbf{x}_g^{\text{IM}} = \boldsymbol{\omega}_\gamma) \prod_{j \in \mathcal{I}^\gamma} \left(\sum_{s \in \mathcal{S}} \exp \left\{ -\frac{|s|^2 - 2\Re\{s \hat{\vartheta}_j^*\}}{v_{\hat{\vartheta}_j}} \right\} P(x_j = s) \right). \quad (53)$$

for $x_i = 0$, where $D(x_i)$ and E are given by (52) and (53), respectively. The *a posteriori* mean and variance of x_i are formulated as

$$\hat{x}_i = \sum_{s \in \mathcal{S}} s b(x_i = s), v_{x_i} = \sum_{s \in \mathcal{S}} |s|^2 b(x_i = s) - |\hat{x}_i|^2. \quad (54)$$

Note that there is no *a priori* information arriving from the decoder during the first iteration, and the *a priori* pdf of \mathbf{x}_g can be directly written as

$$P_{\mathbf{x}_g}^0(\mathbf{x}_g) = \frac{1}{\Upsilon} \sum_{\gamma=1}^{\Upsilon} \left[\prod_{i \in \mathcal{I}^\gamma} \left(\frac{1}{M} \sum_{s \in \mathcal{S}} \delta(x_i - s) \right) \prod_{j \in \mathcal{I}^\gamma} \delta(x_j) \right].$$

Accordingly, the *a posteriori* mean and variance of x_i for the first iteration is given by

$$\hat{x}_i^0 = \sum_{s \in \mathcal{S}} s \frac{||\boldsymbol{\Omega}_{i,1}||_0}{||\boldsymbol{\Omega}||_0 M}, \quad v_{x_i}^0 = \frac{1}{M} \sum_{s \in \mathcal{S}} |s|^2 \frac{||\boldsymbol{\Omega}_{i,1}||_0}{||\boldsymbol{\Omega}||_0} - |\hat{x}_i^0|^2, \quad (55)$$

where $\boldsymbol{\Omega}_{i,1} = \{\boldsymbol{\omega}_\gamma | \omega_{\gamma,i} = 1, \gamma = 1, \dots, \Upsilon\}$.

REFERENCES

- [1] S. Sasaki, J. K. Zhu, and G. Marubayashi, "Performance of parallel combinatorial spread spectrum multiple access communication systems," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 1991, pp. 204–208.
- [2] N. Ishikawa, S. Sugiura, and L. Hanzo, "50 years of permutation, spatial and index modulation: From classic RF to visible light communications and data storage," *IEEE Commun. Surveys Tut.*, vol. 20, no. 3, pp. 1905–1938, thirdquarter 2018.
- [3] D. Tsonev, S. Sinanovic, and H. Haas, "Enhanced subcarrier index modulation (SIM) OFDM," *Proc. IEEE GLOBECOM Workshops*, pp. 728–732, Dec 2011.
- [4] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5536–5549, Nov 2013.
- [5] R. Fan, Y. J. Yu, and Y. L. Guan, "Generalization of orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5350–5359, Oct 2015.
- [6] M. Wen, B. Ye, E. Basar, Q. Li, and F. Ji, "Enhanced orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4786–4801, July 2017.
- [7] M. Wen, E. Basar, Q. Li, B. Zheng, and M. Zhang, "Multiple-mode orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Commun.*, vol. 65, no. 9, pp. 3892–3906, Sep. 2017.
- [8] Q. Li, M. Wen, E. Basar, H. V. Poor, B. Zheng, and F. Chen, "Diversity enhancing multiple-mode OFDM with index modulation," *IEEE Trans. Commun.*, vol. 66, no. 8, pp. 3653–3666, Aug 2018.
- [9] M. Wen, X. Chen, Q. Li, E. Basar, Y. Wu, and W. Zhang, "Index modulation aided subcarrier mapping for dual-hop OFDM relaying," *IEEE Trans. Commun.*, vol. 67, no. 9, pp. 6012–6024, Sep. 2019.
- [10] H. Zhang, L. Yang, and L. Hanzo, "Compressed sensing improves the performance of subcarrier index-modulation-assisted OFDM," *IEEE Access*, vol. 4, pp. 7859–7873, 2016.
- [11] L. Piazzo and P. Mandarini, "Analysis of phase noise effects in OFDM modems," *IEEE Trans. Commun.*, vol. 50, no. 10, pp. 1696–1705, Oct. 2002.
- [12] N. Ishikawa, S. Sugiura, and L. Hanzo, "Subcarrier-index modulation aided OFDM - will it work?" *IEEE Access*, vol. 4, pp. 2580–2593, 2016.
- [13] Y. Xiao, S. Wang, L. Dan, X. Lei, P. Yang, and W. Xiang, "OFDM with interleaved subcarrier-index modulation," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1447–1450, Aug 2014.
- [14] E. Basar, "OFDM with index modulation using coordinate interleaving," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 381–384, Aug 2015.
- [15] O. H. Salim, A. A. Nasir, H. Mehrpouyan, W. Xiang, S. Durrani, and R. A. Kennedy, "Channel, phase noise, and frequency offset in OFDM systems: Joint estimation, data detection, and hybrid cram-cao lower bound," *IEEE Trans. Commun.*, vol. 62, no. 9, pp. 3311–3325, Sep. 2014.
- [16] D. D. Lin and T. J. Lim, "The variational inference approach to joint data detection and phase noise estimation in OFDM," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1862–1874, May. 2007.
- [17] R. Wang, H. Mehrpouyan, M. Tao, and Y. Hua, "Channel estimation, carrier recovery, and data detection in the presence of phase noise in OFDM relay systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1186–1205, Feb 2016.
- [18] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
- [19] J. Winn and C. M. Bishop, "Variational message passing," *Journal of Machine Learning Research*, vol. 6, no. Apr, pp. 661–694, 2005.
- [20] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *Proc. IEEE Inf. Symp. Inf. Theory*, pp. 2168–2172, 2011.
- [21] E. Riegler, G. E. Kirkelund, C. N. Manchon, M. Badiu, and B. H. Fleury, "Merging belief propagation and the mean field approximation: A free energy approach," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 588–602, Jan. 2013.
- [22] X. Meng, S. Wu, L. Kuang, D. Huang, and J. Lu, "Multi-user detection for spatial modulation via structured approximate message passing," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1527–1530, Aug 2016.
- [23] L. Wei, J. Zheng, and Q. Liu, "Approximate message passing detector for index modulation with multiple active resources," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 972–976, Jan 2019.
- [24] 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.
- [25] N. Noels, J. Bhatti, H. Bruneel, and M. Moeneclaey, "Block-processing soft-input soft-output demodulator for coded PSK using DCT-based phase noise estimation," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2939–2950, Aug. 2014.
- [26] V. Franz and J. B. Anderson, "Concatenated decoding with a reduced-

search BCJR algorithm,” *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 186–195, Feb 1998.

- [27] W. Yuan, N. Wu, A. Zhang, X. Huang, Y. Li, and L. Hanzo, “Iterative receiver design for FTN signaling aided sparse code multiple access,” *IEEE Trans. Wireless Commun.*, pp. 1–1, 2019.
- [28] Q. Guo, D. Huang, S. Nordholm, J. Xi, and Y. Yu, “Iterative frequency domain equalization with generalized approximate message passing,” *IEEE Signal Process. Lett.*, vol. 20, no. 6, pp. 559–562, Jun. 2013.
- [29] Q. Shi, N. Wu, X. Ma, and H. Wang, “Frequency-domain joint channel estimation and decoding for faster-than-nyquist signaling,” *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 781–795, Feb. 2018.
- [30] M. D. Renzo, H. Haas, and P. M. Grant, “Spatial modulation for multiple-antenna wireless systems: a survey,” *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 182–191, December 2011.
- [31] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, “Spatial modulation,” *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, July 2008.
- [32] X. Wang and B. Hu, “A low-complexity ML estimator for carrier and sampling frequency offsets in OFDM systems,” *IEEE Commun. Lett.*, vol. 18, no. 3, pp. 503–506, March 2014.
- [33] W. Wang, Z. Wang, C. Zhang, Q. Guo, P. Sun, and X. Wang, “A BP-MF-EP based iterative receiver for joint phase noise estimation, equalization, and decoding,” *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1349–1353, Oct. 2016.
- [34] M. Morelli and M. Moretti, “Fine carrier and sampling frequency synchronization in OFDM systems,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1514–1524, April 2010.
- [35] S. Wu and Y. Bar-Ness, “OFDM systems in the presence of phase noise: consequences and solutions,” *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1988–1996, Nov 2004.
- [36] J. Xu, L. Chen, I. Djurdjevic, S. Lin, and K. Abdel-Ghaffar, “Construction of regular and irregular LDPC codes: Geometry decomposition and masking,” *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 121–134, Jan 2007.
- [37] D. D. Lin, R. A. Pacheco, Teng Joon Lim, and D. Hatzinakos, “Joint estimation of channel response, frequency offset, and phase noise in ofdm,” *IEEE Trans. Signal Processing*, vol. 54, no. 9, pp. 3542–3554, Sep. 2006.
- [38] P. Aquilina and T. Ratnarajah, “Performance analysis of IA techniques in the MIMO IBC with imperfect CSI,” *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1259–1270, April 2015.
- [39] A. Almradi and K. A. Hamdi, “Spectral efficiency of OFDM systems with random residual CFO,” *IEEE Trans. Commun.*, vol. 63, no. 7, pp. 2580–2590, July 2015.



Qiaolin Shi received the B. S. degree in 2014 from Beijing Institute of Technology, Beijing, China. She is currently working toward the Ph.D. degree with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China, and the Faculty of Engineering and Information Technology, University of Technology Sydney, NSW, Australia. Her research interests include statistical learning on graphical models and its application to wireless communications. Mrs. Shi has served as a Reviewer for several international journals and conferences.



Nan Wu (M’11) received his B.S., M.S. and Ph.D. degrees from Beijing Institute of Technology (BIT), Beijing, China in 2003, 2005 and 2011, respectively. From 2008 to 2009, he was a Visiting Ph.D. Student with the Department of Electrical Engineering, Pennsylvania State University, USA. He is currently a Professor with the School of Information and Electronics, BIT. His research interests include signal processing in wireless communication networks. He was a recipient of the National Excellent Doctoral Dissertation Award by MOE of China in 2013. He

serves as an Editorial Board Member of the *IEEE Access*, the *International Journal of Electronics and Communications*, the *KSII Transactions on Internet and Information Systems*, and the *IEICE Transactions on Communications*.



Hua Wang received his Ph.D. degree from Beijing Institute of Technology (BIT), Beijing, China, in 1999. He is now a Professor with School of Information and Electronics, BIT. From Feb. 2009 to Jan. 2010, he was a visiting professor with Department of Electrical Engineering, Arizona State University, USA. His research interests are in the fields of communication theory and signal processing, wireless networking, modem design and implementation for satellite communication.



Xiaoli Ma (F’16) received the B.S. degree in automatic control from Tsinghua University, Beijing, China, in 1998, the M.S. degree in electrical engineering from the University of Virginia, Charlottesville, VA, USA, in 2000, and the Ph.D. degree in electrical engineering from the University of Minnesota, Minneapolis, MN, USA, in 2003. She joined the Department of Electrical and Computer Engineering, Auburn University, where she was an Assistant Professor until December 2005. Since January 2006, she has been with the School of Electrical and Computer Engineering, Georgia Tech, Atlanta, GA, USA, where she is currently a Professor. Her research focuses on wireless networking and communications, including network performance analysis, transceiver designs for wireless time- and frequency-selective channels, channel estimation and equalization algorithms, carrier frequency synchronization for OFDM systems, Internet of Things, and performance analysis for blockchains. She was the recipient of the Lockheed Martin Aeronautics Company Deans Award for Teaching Excellence by the College of Engineering in 2009, and Outstanding Junior Faculty Award by the School of Electrical and Computer Engineering in 2010, at Georgia Tech. She was a Senior Area Editor for *IEEE Signal Processing Letters* (2014–2017) and *Elsevier Digital Signal Processing* (2012–2016), and has been an Associate Editor for *IEEE Signal Processing Letters* (2007–2009) and the *IEEE Transactions on Wireless Communications* (2008–2013).



Lajos Hanzo (M91-SM92-F04) (<http://www-mobile.ecs.soton.ac.uk>) FREng, FIEEE, FIET, Fellow of EURASIP, DSc received his degree in electronics in 1976 and his doctorate in 1983. In 2009 he was awarded an honorary doctorate by the Technical University of Budapest and in 2015 by the University of Edinburgh. In 2016 he was admitted to the Hungarian Academy of Science. During his 40-year career in telecommunications he has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has

been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has successfully supervised 120+ PhD students, co-authored 18 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published 1910 research contributions at IEEE Xplore, acted both as TPC and General Chair of IEEE conferences, presented keynote lectures and has been awarded a number of distinctions. Currently he is directing an academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) UK, the European Research Council’s Advanced Fellow Grant and the Royal Society’s Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE ComSoc and VTS. During 2008 - 2012 he was the Editor-in-Chief of the *IEEE Press* and a Chaired Professor also at Tsinghua University, Beijing. For further information on research in progress and associated publications please refer to <http://www-mobile.ecs.soton.ac.uk>