copyright IEEE 2014

Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, including reprinting/republishing this material for advertising or promotional purposes, collecting new collected works for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Block-processing Soft-Input Soft-Output Demodulator for Coded PSK using DCT-based Phase Noise Estimation

Nele Noels, *Member, IEEE,*, Jabran Bhatti, Herwig Bruneel and Marc Moeneclaey, *Fellow, IEEE*

***Abstract*—This paper considers the detection of coded phase-shift keying signals subjected to additive white Gaussian noise and oscillator phase noise. We propose a detector that partitions the received frame into smaller blocks and models the unknown phasor variations over each block as a truncated discrete cosine transform (DCT) expansion. Detection and decoding are performed iteratively between a soft-input soft-output (SISO) demodulator, a SISO demapper and a SISO decoder based on the sum-product algorithm and the factor graph framework, while the expectation-maximization algorithm is used in the demodulator for DCT coefficients estimation. The resulting demodulator is shown to have an excellent performance/complexity trade-off and to be well-suited for parallel processing on multiple cores.**

***Index Terms*—Phase noise, Synchronization, Parallel algorithms, Parameter estimation, Error correction coding, Computation time, Communication systems, Digital communication, Digital signal processors, Discrete cosine transforms.**

I. Introduction

Over the last decade, several iterative receiver algorithms have been developed for bit interleaved coded symbols transmitted over an additive white Gaussian noise (AWGN) channel affected by oscillator phase noise (PN) [1]–[11]. In [1]–[4], iterative joint soft-input soft-output (SISO) demodulation, demapping and decoding algorithms are derived by applying the sum-product algorithm (SPA) to a suitable factor graph (FG). Among these algorithms, the one from [1] shows excellent performance at moderate complexity. A different approach is adopted in [5], [6], where the expectation-maximization algorithm (EMA) [12], [13] is used to estimate the PN, and a joint decoding and demapping procedure is employed which treats this PN estimate as the true value. In [5], the received data frame is divided into blocks and within each block the local time-average of the phase is estimated. In [6], the PN is tracked using a first-order decision-directed (DD) phase-locked loop (PLL). Similar receivers were also proposed in [2] and [8]; moreover, instead of a PLL, extended Kalman filtering techniques can also be used to track the phase, yielding a receiver with similar structure and performance [7], [14]. An important issue, with these PLL/Kalman receivers, is the phenomenon of cycle slip that may significantly compromise the error performance. In general, the robustness to cycle slip can be improved by inserting more pilot symbols into the data stream (at the cost of reduced power and bandwidth efficiency), and by optimizing the initial PN estimate used to start the iterations [8], [15]. The latter requires a priori knowledge of the PN spectrum which might not be available and significantly adds to the implementation and computation complexity of the receiver.

The present paper considers a new EMA-based receiver that models the unknown phasor as a truncated discrete cosine transform (DCT) expansion. Compared to a previously published conference paper of the present authors [9], here, a block-processing version of the phase estimator is considered, a performance and complexity analysis is performed, and the proposed receiver is compared to the receiver proposed in [1].

The paper is organized as follows. Section II describes the observation model and the main FG for the derivation of the proposed algorithm, which is detailed in Section III. Section IV provides an approximate theoretical performance analysis, Section V reports on the algorithm's computational complexity and Section VI addresses its parallel processing. Numerical results are shown in Section VII and conclusions are drawn in Section VIII.

Throughout this paper we use the following notations and conventions. Equality up to an irrelevant normalization factor is indicated as $\propto$. The mean of a random variable is denoted $\mathbb{E}[\cdot]$ and $|.|^2$ is the squared Euclidean norm of a vector of complex numbers. The trace of a matrix is referred to as $tr(\cdot)$; $[\cdot]^T$ and $[\cdot]^H$ are the transpose and Hermitian conjugate of a vector or matrix, and $diag\{\cdot\}$ returns a matrix diagonal as a column vector. The operators $\cdot^*$, $\Re\{\cdot\}$, $\Im\{\cdot\}$ and $\arg\{\cdot\}$ represent the conjugate, real part, imaginary part and angle of a complex number. $\jmath$ is a short-hand notation for $\sqrt{-1}$. The $D \times D$ identity matrix is represented as $\mathbf{I}_D$, and $\mathbf{0}_{D_1 \times D_2}$ denotes the $D_1 \times D_2$ matrix with all elements equal to zero. A vector $\mathbf{y} = (y_0, y_1, ..., y_{K-1})^T$ of dimension K will be decomposed into S subvectors of dimension L (with $LS = K$), i.e., $\mathbf{y} = \left(\mathbf{y}_{1;L}^T, \mathbf{y}_{2;L}^T, ..., \mathbf{y}_{S;L}^T\right)^T$ where $\mathbf{y}_{h;L} = (y_{h,0}, y_{h,1}, ..., y_{h,L-1})^T$, with $y_{h,l} = y_{(h-1)L+l}$ for $h = 1, 2, ..., S$ and $l = 0, 1, ..., L-1$, denotes the hth subvector[1].

Manuscript received November 14, 2013; revised March 21, 2014 and July 4, 2014.

Copyright (c) 2014 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

N. Noels, J. Bhatti, H. Bruneel and M. Moeneclaey are with the Department of Telecommunications and Information Processing, Ghent University, Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium, e-mail: nn,jb,hb,mm@telin.UGent.be.

[1]The restriction that all subvectors have equal length is easily relaxed.

0000–0000/00$00.00 © 2014 IEEE

II. System Description

A. Observation Model

We consider the transmission of a frame of K symbols over an AWGN channel that is affected by carrier PN. The resulting discrete-time received signal is described as: $\mathbf{r} = (r_0, r_1, ..., r_{K-1})^T$, with

$$r_k = a_k u_k + n_k, \tag{1}$$

where the index $k \in I = \{0, 1, ..., K-1\}$ refers to the kth symbol interval of length T, $\mathbf{a} = (a_0, a_1, ..., a_{K-1})^T$ is the transmitted symbol sequence, $\mathbf{u} = (u_0, u_1, ..., u_{K-1})^T$ collects the phasor samples and the vector $\mathbf{n} = (n_0, n_1, ..., n_{K-1})^T$ consists of K independent and identically distributed zero-mean circular symmetric complex-valued Gaussian random variables (ZMCSCGRVs) with $\mathbb{E}\left[|n_k|^2\right] = N_0$. The carrier phasor samples are modeled as $u_k = e^{j\theta_k}$, with $\{\theta_k\}$ a discrete-time PN process. The transmitted symbols belong to an M-PSK constellation $\Omega = \{\omega_0, \omega_1, ..., \omega_{M-1}\}$, with $|\omega_i|^2 = E_s$, for $i = 0, 1, ..., M-1$, where E_s denotes the symbol energy. To aid synchronization, K_p known pilot symbols have been periodically inserted in $\mathbf{a}$ at positions $k \in I_p = \{k_j : j = 0, 1, ..., K_p - 1\}$, with

$$k_j = \left\lfloor \frac{K(2j+1) - K_p}{2K_p} \right\rceil, \tag{2}$$

where $\lfloor x \rceil$ rounds x to the nearest integer[2]; when K is an odd multiple of K_p, the argument of the rounding function in (2) is integer, so that the rounding can be dropped. The remaining $K - K_p$ symbols a_k at positions $k \in I_d = I \backslash I_p$ represent unknown data symbols, resulting from encoding an information bit sequence $\mathbf{b}$ of size K_b and mapping the resulting coded bit sequence to a sequence of complex valued constellation points.

B. Main Factor Graph

In the following, we assume that the reader is familiar with the FG/SPA framework, see, e.g., [16]. Assuming independent equally-likely information bits, and dropping all factors not depending on $(\mathbf{b}, \mathbf{a})$, the joint a posteriori probability (APP) $p(\mathbf{b}, \mathbf{a}|\mathbf{r})$ of $\mathbf{b}$ and $\mathbf{a}$, given $\mathbf{r}$ from (1) factorizes as:

$$p(\mathbf{b}, \mathbf{a}|\mathbf{r}) \propto p(\mathbf{a}|\mathbf{b})\, p(\mathbf{r}|\mathbf{a}), \tag{3}$$

where $p(\mathbf{a}|\mathbf{b})$ equals 1 when the code and mapping constraints map $\mathbf{b}$ to $\mathbf{a}$ and equals 0 otherwise; for given $\mathbf{r}$, $p(\mathbf{r}|\mathbf{a})$ is the likelihood function (LF) of $\mathbf{a}$. The corresponding FG is shown in Fig. 1

and forms the starting point for the development of the proposed algorithm. The messages from the factor nodes (FN) $p(\mathbf{a}|\mathbf{b})$ and $p(\mathbf{r}|\mathbf{a})$ to a variable node (VN) a_k are denoted $L_{d,k}^{(\gamma)}(\omega_i)$ and $L_{u,k}^{(\gamma)}(\omega_i)$, respectively, where i = 0, 1, ..., $M-1$ and γ indicates the SPA iteration number[3]. It is assumed that

[2] x values with a fractional part of 0.5 round up to the nearest integer larger than x.

[3] Unless the FG is a tree, the SPA is iterative.

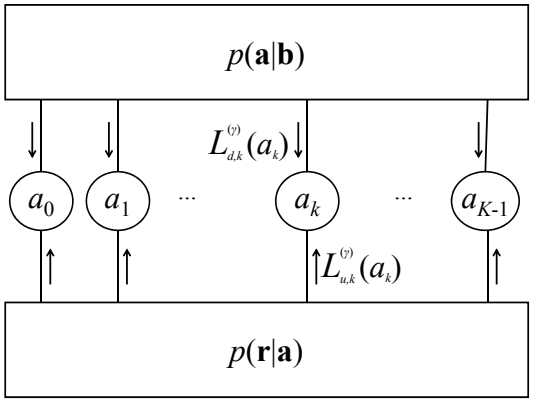

Figure 1. Graphical representation of iterative receiver for bit interleaved coded modulated signals.

these FG messages are in the "logarithmic domain", i.e.,

$$L_{e,k}^{(\gamma)}(\omega_i) = \ln\left(\frac{\mu_{e,k}^{(\gamma)}(\omega_i)}{\mu_{e,k}^{(\gamma)}(\omega_0)}\right)$$

for $e \in \{d, u\}$, where the "linear domain" FG messages $\mu_{e,k}^{(\gamma)}(\omega_i)$ are computed according to the SPA; note that $L_{e,k}^{(\gamma)}(\omega_0) = 0$. The focus in this paper is on applying the SPA to the FN $p(\mathbf{r}|\mathbf{a})$, which can be interpreted as a SISO demodulation process. This involves the computation of the upward messages $L_{u,k}^{(\gamma)}(\omega_i)$ for i = 0, 1, ..., $M-1$ and k = 0, 1, ..., $K-1$ from the downward messages $L_{d,k}^{(\gamma)}(\omega_i)$, i = 0, 1, ..., $M-1$ and k = 0, 1, ..., $K-1$. The downward messages, which result from the SPA applied to the FN $p(\mathbf{a}|\mathbf{b})$, are obtained as a by-product of the conventional joint SISO demapping and decoding operations[4].

III. SISO Demodulator using Block-processing DCT-based Phasor Estimator

A. DCT-based Phasor Model

Because PN is typically a low-pass process [17], we propose using a compact representation of each of the $S = \frac{K}{L}$ size-L subvectors $\mathbf{u}_{h;L}$, h = 1, 2, ..., S, of $\mathbf{u} = \left(\mathbf{u}_{1;L}^T, \mathbf{u}_{2;L}^T, ..., \mathbf{u}_{S;L}^T\right)^T$ by means of only a few ($N \ll L$) parameters. The idea is to reduce the number of unknowns to be estimated from the observed signal.

It is always possible to write a vector $\mathbf{u}_{h;L}$ of size L as a linear combination of L orthonormal basis vectors $\mathbf{e}_{h,l}$ with coefficients $\lambda_{h,l}$, l = 0, 1, ..., $L-1$:

$$\mathbf{u}_{h;L} = \sum_{l=0}^{L-1} \lambda_{h,l} \mathbf{e}_{h,l}.$$

[4] A similar approach was taken in [1]–[6], [8], [9]. Receiver algorithms that, in order to cope with the PN, modify the demapping and decoding operation itself have also been developed [10], [11]; however, these algorithms are specifically designed for trellis-based detectors and can therefore not be used, for example, for LDPC coded signals.

A compact representation $\mathbf{u}_{h;L}^{(\mathrm{c})}$ of $\mathbf{u}_{h;L}$ then results from keeping in this expansion only N terms, with N significantly smaller than L:

$$\mathbf{u}_{h;L}^{(\mathrm{c})} = \sum_{l \in I^{(\mathrm{c})}} \lambda_{h,l} \mathbf{e}_{h,l}, \tag{4}$$

where $I^{(\mathrm{c})}$ denotes a subset of $\{0, 1, ..., L-1\}$ with cardinality N. The accuracy of the compact representation (4) depends on the statistics of $\mathbf{u}_{h;L}$, on the choice of the basis vectors $\mathbf{e}_{h,l}$, $l = 0, 1, ..., L-1$ and on the choice of $I^{(\mathrm{c})}$. For given N, the mean square error

$$E\left[\left|\mathbf{u}_{h;L} - \mathbf{u}_{h;L}^{(\mathrm{c})}\right|^2\right] \tag{5}$$

can be minimized by using the Karhunen-Loève (KL) expansion [18]; the corresponding basis vectors are the N eigenvectors of the correlation matrix $\mathbf{C}_{\mathbf{u}_{h;L}} = E[\mathbf{u}_{h;L}\mathbf{u}_{h;L}^H]$ that correspond to the N largest eigenvalues of $\mathbf{C}_{\mathbf{u}_{h;L}}$. However, the correlation matrix of $\mathbf{u}_{h;L}$ is usually unknown, in which case the KL expansion cannot be obtained. Instead, we propose to use the first N basis functions of the DCT for the compact representation of the phasor subvectors.

Using matrix notation, this yields, for $h = 1, 2, ..., S$:

$$\mathbf{u}_{h;L}^{(\mathrm{c})} = \mathbf{\Psi}_{L\times N} \mathbf{x}_{h;N}, \tag{6}$$

where $\mathbf{\Psi}_{L\times N}$ is an $L \times N$ DCT matrix, with

$$(\mathbf{\Psi}_{L\times N})_{l,n} = \begin{cases} \sqrt{\frac{1}{L}}, & n = 0 \\ \sqrt{\frac{2}{L}} \cos\left(\frac{\pi n}{L}\left(l + \frac{1}{2}\right)\right), & n = 1, 2, ..., N-1 \end{cases}, \tag{7}$$

for $l = 0, 1, ..., L-1$, and $\mathbf{x}_{h;N} = (x_{h,0}, x_{h,1}, ..., x_{h,N-1})^T$ contains the N coefficients of the truncated DCT expansion. Because the frequency content of the nth ($n = 0, 1, ..., L-1$) DCT basis function is centered at $\pm n/(2L)$, the first N DCT basis functions are well suited for the compact representation of a discrete-time lowpass process with a bandwidth $N/(2L)$, which explains their use in several audio and image compression algorithms. Moreover, fast algorithms exist for computing the DCT basis expansion coefficients for large L [19], [20]. In [21], it has been shown that for a given value of N, the DCT outperforms several other fast orthogonal transforms in terms of the MSE (5), and closely approaches KL performance. These observations motivate the use of the DCT for modeling $\mathbf{u}_{h;L}$.

B. EMA-based DCT Coefficients Estimation

Assuming that (6) is an accurate approximation of $\mathbf{u}_{h;L}$ for $h = 1, 2..., S$ ($S = K/L$), we propose to estimate the phasor vector $\mathbf{u}$ as $\hat{\mathbf{u}} = \left(\hat{\mathbf{u}}_{1;L}^T, \hat{\mathbf{u}}_{2;L}^T, ..., \hat{\mathbf{u}}_{S;L}^T\right)^T$ with

$$\hat{\mathbf{u}}_{h;L} = \mathbf{\Psi}_{L\times N} \hat{\mathbf{x}}_{h;N},$$

where $\hat{\mathbf{x}}_{h;N}$ is the hth subvector of an estimate $\hat{\mathbf{x}} = \left(\hat{\mathbf{x}}_{1;N}^T, \hat{\mathbf{x}}_{2;N}^T, ..., \hat{\mathbf{x}}_{S;N}^T\right)^T$ of the unknown deterministic vector parameter $\mathbf{x} = \left(\mathbf{x}_{1;N}^T, \mathbf{x}_{2;N}^T, ..., \mathbf{x}_{S;N}^T\right)^T$ that groups the DCT coefficient vectors of all subvectors $\mathbf{u}_{h;L}$ of $\mathbf{u}$. Note that $\mathbf{x}$ contains L/N times less elements than $\mathbf{u}$. We use the EMA [12] to produce a maximum-likelihood estimate of $\mathbf{x}$, in the presence of the symbol sequence $\mathbf{a}$, which acts as nuisance vector parameter. The EMA is an iterative procedure. Considering that also the SPA is iterative, we obtain a double-iterative receiver. In order to limit the number of SPA iterations, we intertwine both types of iterations [13], i.e., after each EMA iteration, we execute one SPA iteration (without resetting the FG messages); the resulting receiver operation can be described by means of a single iteration index γ. At the γth iteration, the estimate $\hat{\mathbf{x}}^{(\gamma)}$ for $\mathbf{x}$ is computed as

$$\hat{\mathbf{x}}^{(\gamma)} = \arg\max_{\mathbf{x}} Q\left(\mathbf{x}; \hat{\mathbf{x}}^{(\gamma-1)}\right), \tag{8}$$

with

$$Q\left(\mathbf{x}; \hat{\mathbf{x}}^{(\gamma-1)}\right) = \sum_{\mathbf{a}} p\left(\mathbf{a}\left|\mathbf{r}, \hat{\mathbf{x}}^{(\gamma-1)}\right.\right) \ln p\left(\mathbf{r}\left|\mathbf{a}, \mathbf{x}\right.\right). \tag{9}$$

the expectation of the log-LF $\ln p\left(\mathbf{r}\left|\mathbf{a}, \mathbf{x}\right.\right)$ of $(\mathbf{a}, \mathbf{x})$ with respect to the APP $p\left(\mathbf{a}\left|\mathbf{r}, \hat{\mathbf{x}}^{(\gamma-1)}\right.\right)$ of $\mathbf{a}$, given $\mathbf{r}$ and $\mathbf{x} = \hat{\mathbf{x}}^{(\gamma-1)}$. Using (6) and assuming that $\mathbf{u}_{h,L} = \mathbf{u}_{h,L}^{(\mathrm{c})}$ with equality, it follows from (1) that the LF $p\left(\mathbf{r}\left|\mathbf{a}, \mathbf{x}\right.\right)$ of $(\mathbf{a}, \mathbf{x})$ decomposes as:

$$p\left(\mathbf{r}\left|\mathbf{a}, \mathbf{x}\right.\right) = \prod_{h=1}^{S} p\left(\mathbf{r}_{h;L}\left|\mathbf{a}_{h;L}, \mathbf{x}_{h;N}\right.\right), \tag{10}$$

with

$$\begin{aligned} &p\left(\mathbf{r}_{h;L}\left|\mathbf{a}_{h;L}, \mathbf{x}_{h;N}\right.\right) \\ &= \exp\left\{-\frac{E_s}{N_0}\left|\frac{\mathrm{diag}\left\{\mathbf{r}_{h;L}\mathbf{a}_{h;L}^H\right\}}{E_s} - \mathbf{\Psi}_{L\times N}\mathbf{x}_{h;N}\right|^2\right\}, \end{aligned} \tag{11}$$

where we have exploited the fact that $|a_k|^2 = E_s$. Substituting (10)-(11) into (9) yields

$$\begin{aligned} &Q\left(\mathbf{x}; \hat{\mathbf{x}}^{(\gamma-1)}\right) \\ &= -\frac{1}{N_0}\sum_{h=1}^{S}\left(|\mathbf{r}_{h;L}|^2 - 2\Re\left\{\mathbf{v}_{h;L}^H \mathbf{\Psi}_{L\times N}\mathbf{x}_{h;N}\right\} + \mathbf{x}_{h;N}^H\mathbf{x}_{h;N}\right), \end{aligned} \tag{12}$$

with

$$\mathbf{v}_{h;L} = \sum_{\mathbf{a}} \Pr\left[\mathbf{a}\left|\mathbf{r}, \hat{\mathbf{x}}^{(\gamma-1)}\right.\right] \frac{\mathrm{diag}\left\{\mathbf{r}_{h;L}\mathbf{a}_{h;L}^H\right\}}{E_s}, \tag{13}$$

because $|a_k|^2 = E_s$ and $\mathbf{\Psi}_{L\times N}^T \mathbf{\Psi}_{L\times N} = \mathbf{I}_N$. Substituting (12) in (8) yields:

$$\hat{\mathbf{x}}^{(\gamma)} = \arg\max_{\mathbf{x}} \sum_{h=1}^{S}\left(2\Re\left\{\mathbf{v}_{h;L}^H \mathbf{\Psi}_{L\times N}\mathbf{x}_{h;N}\right\} - \mathbf{x}_{h;N}^H\mathbf{x}_{h;N}\right), \tag{14}$$

or, equivalently, with $\hat{\mathbf{x}}^{(\gamma)} = (\hat{\mathbf{x}}_{1;N}^{(\gamma)T}, \hat{\mathbf{x}}_{2;N}^{(\gamma)T}, ..., \hat{\mathbf{x}}_{S;N}^{(\gamma)T})^T$,

$$\hat{\mathbf{x}}_{h;N}^{(\gamma)} = \arg\max_{\mathbf{x}_{h;N}} \left(2\Re\left\{\mathbf{v}_{h;L}^H \mathbf{\Psi}_{L\times N}\mathbf{x}_{h;N}\right\} - \mathbf{x}_{h;N}^H\mathbf{x}_{h;N}\right), \tag{15}$$

$$= \left(\mathbf{\Psi}_{L\times N}^T \mathbf{\Psi}_{L\times N}\right)^{-1} \mathbf{\Psi}_{L\times N}^T \mathbf{v}_{h;L}, \tag{16}$$

$$= \mathbf{\Psi}_{L\times N}^T \mathbf{v}_{h;L}^{(\gamma)}, \tag{17}$$

for $h = 1, 2, ..., S$, where the last equality follows from $\mathbf{\Psi}_{L\times N}^T \mathbf{\Psi}_{L\times N} = \mathbf{I}_N$. The phasor estimate $\hat{\mathbf{u}}_{h;L}^{(\gamma)}$ related to the hth block is then obtained by substituting (17) into (6):

$$\hat{\mathbf{u}}_{h;L}^{(\gamma)} = \mathbf{\Psi}_{L\times N} \hat{\mathbf{x}}_{h;N}^{(\gamma)}. \tag{18}$$

Because the lth component $v_{h,l}^{(\gamma)}$ of $\mathbf{v}_h^{(\gamma)}$ from (13), only depends on the lth symbol $a_{h,l}$ in $\mathbf{a}_{h;L}$, computing the expectation (13), for $h = 1, 2, ..., S$, is equivalent to evaluating:

$$v_k^{(\gamma)} = \sum_{i=0}^{M-1} \tilde{r}_k(\omega_i)\, p\left(a_k = \omega_i \,|\mathbf{r};\ \hat{\mathbf{u}}^{(\gamma-1)}\right), \tag{19}$$

for $k = 0, 1, ..., K-1$, with $v_{h,l}^{(\gamma)} = v_{(h-1)L+l}^{(\gamma)}$, $l = 0, 1, ..., L-1$. Here, $\tilde{r}_k(\omega_i)$ is defined as

$$\tilde{r}_k(\omega_i) = r_k \omega_i^* \frac{1}{E_s},\ k = 0, 1, ...K-1 \tag{20}$$

and $p\left(a_k = \omega_i \,|\mathbf{r};\ \hat{\mathbf{u}}^{(\gamma)}\right) = p\left(a_k = \omega_i \,|\mathbf{r};\ \hat{\mathbf{x}}^{(\gamma)}\right)$ is the marginal APP of a_k being equal to ω_i, given $\mathbf{r}$ and $\mathbf{x} = \hat{\mathbf{x}}^{(\gamma)}$.

C. SPA-based Efficient Marginal Symbol APP Computation

The pilot symbols are known, yielding for $k \in I_p$ and for all γ: $p\left(a_k = \omega_i \,|\mathbf{r};\ \hat{\mathbf{u}}^{(\gamma-1)}\right)$ is 1 if $a_k = \omega_i$ and 0 otherwise. An approximation of the APPs $p\left(a_k = \omega_i \,|\mathbf{r};\ \hat{\mathbf{u}}^{(\gamma-1)}\right)$, for $k \in I_d$, can be efficiently computed by means of the SPA. Treating $\hat{\mathbf{u}}^{(\gamma-1)}$ as the true value of the carrier phasor, the LF $p(\mathbf{r}\,|\mathbf{a})$ of $\mathbf{a}$ in (3) further decomposes as:

$$p(\mathbf{r}\,|\mathbf{a}) \propto \prod_{k\in I_d} \exp\left\{F_k^{(\gamma)}(a_k)\right\}, \tag{21}$$

with

$$\begin{aligned} F_k^{(\gamma)}(a_k) &= \frac{1}{N_0}\left(2\Re\left\{r_k a_k^* \hat{u}_k^{(\gamma-1)*}\right\} - \left|\hat{u}_k^{(\gamma-1)}\right|^2\right), \\ &\approx \frac{2}{N_0}\Re\left\{r_k a_k^* \hat{u}_k^{(\gamma-1)*}\right\}, \end{aligned} \tag{22}$$

where the approximation holds for $|\hat{u}_k|^2 \approx 1$; note that $p(\mathbf{r}\,|\mathbf{a})$ from (21) depends on the iteration index through the estimate $\hat{\mathbf{u}}^{(\gamma-1)}$. The FG corresponding to the decomposition (21) of $p(\mathbf{r}\,|\mathbf{a})$ is shown in Fig. 2 . Replacing in Fig. 1 the FN $p(\mathbf{r}\,|\mathbf{a})$ with the FG from Fig. 2 and applying the SPA yields the APPs required in (19).

During the γth iteration, according to the rules of the SPA and the message conventions of this paper, given $\hat{\mathbf{u}}^{(\gamma-1)}$, the $M-1$ values $L_{u,k}^{(\gamma)}(\omega_i)$, $i = 1, 2, ..., M-1$, for $k \in I_d$ are updated as:

$$L_{u,k}^{(\gamma)}(\omega_i) = F_k^{(\gamma)}(\omega_i) - F_k^{(\gamma)}(\omega_0), \tag{23}$$

$$= \Re\left\{\Delta\dot{r}_k(\omega_i)\, \hat{u}_k^{(\gamma-1)*}\right\}, \tag{24}$$

with

$$\Delta\dot{r}_k(\omega_i) = \frac{2E_s}{N_0}\left(\tilde{r}_k(\omega_i) - \tilde{r}_k(\omega_0)\right). \tag{25}$$

Then, applying the standard SPA to the upper FN in Fig. 1 yields new values $L_{d,k}^{(\gamma)}(\omega_i)$, for $i = 1, 2, ..., M-1$ and $k \in I_d$, for the downward messages.

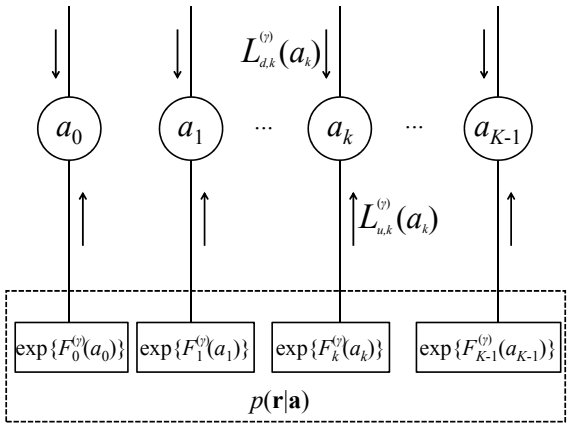

Figure 2. Graphical representation of SISO demodulator treating the estimates $\hat{\mathbf{u}}^{(\gamma-1)}$ as the true phasor values.

Having obtained $L_{u,k}^{(\gamma)}(\omega_i)$ and $L_{d,k}^{(\gamma)}(\omega_i)$, for $i = 1, 2, ..., M-1$ and $k \in I_d$, the marginal symbol APPs are approximated as:

$$p\left(a_k = \omega_i \,|\mathbf{r};\ \hat{\mathbf{u}}^{(\gamma)}\right) = \frac{p_k^{(\gamma)}(\omega_i)}{C_k^{(\gamma)}}, \tag{26}$$

for $i = 0, 1, ..., M-1$, with

$$p_k^{(\gamma)}(\omega_i) = \exp\left\{L_{d,k}^{(\gamma)}(\omega_i) + L_{u,k}^{(\gamma)}(\omega_i)\right\} \tag{27}$$

and

$$C_k^{(\gamma)} = \sum_{i=0}^{M-1} p_k^{(\gamma)}(\omega_i). \tag{28}$$

Substituting (26) and the known pilot symbols into (19) yields:

$$v_k^{(\gamma)} = \begin{cases} \frac{1}{C_k^{(\gamma)}} \sum_{i=0}^{M-1} \tilde{r}_k(\omega_i)\, p_k^{(\gamma)}(\omega_i), & k \in I_d \\ \tilde{r}_k(a_k), & k \in I_p \end{cases}. \tag{29}$$

The quantity $v_k^{(\gamma)}$ can be interpreted as a coarse estimate of u_k, based on the observation r_k and the APP of a_k that is available during the γth iteration; we have $v_k^{(\gamma)} = r_k \tilde{a}_k^{(\gamma)}/E_s$, where the soft decision $\tilde{a}_k^{(\gamma)}$ is the a posteriori expectation of a_k that results from the approximated APPs (26).

D. EMA Initialization

In order to start the iterations, an initial phasor estimate $\hat{\mathbf{u}}^{(0)}$ is required. This initial estimate is obtained from the observations r_k, $k \in I_p$ at the pilot symbol positions [22]. As far as the initial estimate is concerned, the received signal vector $\mathbf{r}$ is divided into S_p blocks of length L_p with $S_p L_p = K$; note that L_p might be different from the block length L used during the iterations. Each block contains B_p pilot symbols such that $S_p B_p = K_p$. We have:

$$\hat{\mathbf{u}}_{h;L_p}^{(0)} = \mathbf{\Psi}_{L_p\times N_p}\left(\mathbf{\Psi}_p^T \mathbf{\Psi}_p\right)^{-1} \mathbf{\Psi}_p^T \left(\mathbf{r}_{h;L_p} \circ \mathbf{a}_{h;L_p}^*\right), \tag{30}$$

$$\approx (K/K_p)\, \mathbf{\Psi}_{L_p\times N_p} \mathbf{\Psi}_p^T \left(\mathbf{r}_{h;L_p} \circ \mathbf{a}_{h;L_p}^*\right), \tag{31}$$

where $\circ$ denotes element-wise multiplication, $\mathbf{\Psi}_{L_p\times N_p}$ is obtained by substituting in (7) $(L,\ N)$ by $(L_p,\ N_p)$, and $\mathbf{\Psi}_p$

Algorithm 1 DCT receiver.

- Initialization:
  - Set all SPA messages to 0.
  - Compute $\{\tilde{r}_k(\omega_i)\}$ from (20).
  - Compute $\{\Delta\dot{r}_k(\omega_i)\}$ from (25).
  - Compute $\hat{\mathbf{u}}^{(0)}$ from (31).
- **For** $\gamma = 1, 2, ...\Gamma$:
  - Apply SPA to the FG of Fig. 1, with the FN $p(\mathbf{a}|\mathbf{b})$ suitably factorized and the FN $p(\mathbf{r}|\mathbf{a})$ decomposed as specified in Fig. 1.
    - Compute $\left\{L_{u,k}^{(\gamma)}(\omega_i)\right\}$ from (24).
    - Update all FG messages once; this includes computing $\left\{L_{d,k}^{(\gamma)}(\omega_i)\right\}$.
    - Compute information bit APPs and perform bit detection.

      If

      Information bit sequence correctly detected.

      Then

      Quit For

      End If
  - Update phasor estimate.
    - Compute $p_k^{(\gamma)}(\omega_i)$ from (27).
    - Compute $\left\{C_{p,k}^{(\gamma)}\right\}$ from (28).
    - Compute $v_k^{(\gamma)}$from (29).
    - Compute $\hat{\mathbf{u}}^{(\gamma)}$ from (18) and (17).

End For

is an $L_p \times N_p$ matrix obtained by replacing in $\mathbf{\Psi}_{L_p \times N_p}$ the $(L_p - B_p)$ rows that correspond to data symbol positions with zero vectors (so that only (r_k, a_k) with $k \in I_p$ contributes to the initial estimate). The approximation (31) holds for the considered pilot symbol positions (2). When K is an odd multiple of K_p, we have $\mathbf{\Psi}_p^T \mathbf{\Psi}_p = (K_p/K)\,\mathbf{I}_{N_p}$, so that the approximation (31) becomes an equality. When K is not an odd multiple of K_p, we use (31) as an approximation of (30).

E. Receiver Scheduling

The receiver scheduling is summarized in Algorithm 1. In each iteration, first all upward and then all downward messages are updated. In addition, information bit APPs are computed and hard decisions about the information bits are made, after which a genie checks for bit errors; the receiver stops iterating after a maximum number of iterations Γ, or when all information bits have been detected correctly. In a practical system, the genie is replaced by a cyclic redundancy check.

IV. Estimator Performance Analysis

Proper operation of the DCT receiver requires an adequate choice for the values of the design parameters L, N, L_p and N_p. As an exact analytical performance analysis is not feasible, we resort to linearization, which holds for small estimation errors.

We assume that after convergence, the soft decisions $\tilde{a}_k^{(\infty)}$ are well approximated by the actual data symbols a_k; this holds for small additive noise and assuming small phase estimation error (PEE) after convergence of the iterative receiver. Taking into account that $\hat{\theta}_k^{(\infty)} = \arg\left(\hat{u}_k^{(\infty)}\right)$, we obtain:

$$\begin{aligned}\hat{\theta}_k^{(\infty)} - \theta_k &= \arg\left(1 + (\hat{u}_k^{(\infty)} - u_k)u_k^*\right), \\ &\approx \Im\{(\hat{u}_k^{(\infty)} - u_k)u_k^*\}. \end{aligned} \tag{32}$$

Using (1), (17) with $v_k^{(\infty)} = r_k a_k^* \frac{1}{E_s}$ and (18) a similar reasoning as in [9] yields:

$$\hat{\boldsymbol{\theta}}_{h;L}^{(\infty)} - \boldsymbol{\theta}_{h;L} \approx \mathbf{M}\Im\{\mathbf{w}_{h;L}\} - (\mathbf{I}_L - \mathbf{M})\,\boldsymbol{\theta}_{h:L} \tag{33}$$

where $\mathbf{w}_{h;L}$ consists of L independent ZMCSCGRVs with variance N_0/E_s and $\mathbf{M} = \mathbf{\Psi}_{L\times N}\mathbf{\Psi}_{L\times N}^T$. The first term in (33) is caused by the additive noise, whereas the second term in the expression of the PEE represents the DCT expansion modeling error. As $\mathbf{M}\mathbf{\Psi}_{L\times N} = \mathbf{\Psi}_{L\times N}$ and $(\mathbf{I}_L - \mathbf{M})\mathbf{\Psi}_{L\times N} = \mathbf{0}_{L\times N}$, the matrix operators $\mathbf{M}$ and $\mathbf{I}_L - \mathbf{M}$ can be interpreted as lowpass filter and a highpass filter, respectively, with (normalized) cut-off frequency $N/(2L)$. Hence, the PEE is determined mainly by the ratio N/L rather than by N and L separately. Making N/L larger (smaller) increases (decreases) the noise contribution but decreases (increases) the effect of the modeling error. The resulting mean square estimation error (MSEE) is given by

$$\begin{aligned}\text{MSEE} &= \frac{1}{L}\mathbb{E}\left[\left|\hat{\boldsymbol{\theta}}_{h;L}^{(\infty)} - \boldsymbol{\theta}_{h;L}\right|^2\right], \\ &= \frac{N}{L}\frac{N_0}{2E_s} + \frac{1}{L}tr\left((\mathbf{I}_L - \mathbf{M})\,\mathbf{C}_{\boldsymbol{\theta}_{h:L}}(\mathbf{I}_L - \mathbf{M})^T\right). \end{aligned} \tag{34}$$

The first term in (34), denoting the contribution from the AWGN, is proportional to the ratio N/L. The second term constitutes an MSEE floor for large E_s/N_0, caused by the modeling error. The PN statistics affect the MSEE floor through the PN correlation matrix $\mathbf{C}_{\boldsymbol{\theta}_{h:L}}$. When PN is absent, the MSEE is minimized by choosing $N = 1$ and $L = K$. In the presence of PN, an optimal value for N/L exists that minimizes the MSEE for the given set of channel statistics (PN and AWGN). In that case, the performance of the receiver is rather insensitive to the choice of L, for a fixed ratio N/L.

A similar reasoning can be applied to obtain the MSEE related to the initial pilot-based estimate $\hat{\boldsymbol{\theta}}_{h;L_p}^{(0)}$. One obtains

$$\begin{aligned}&\text{MSEE} \\ &= \frac{1}{L_p}\mathbb{E}\left[\left|\hat{\boldsymbol{\theta}}_{h;L_p}^{(0)} - \boldsymbol{\theta}_{h;L_p}\right|^2\right], \\ &= \frac{N_p}{B_p}\frac{N_0}{2E_s} + \frac{1}{L_p}tr\left((\mathbf{I}_{L_p} - \mathbf{M}_p)\,\mathbf{C}_{\boldsymbol{\theta}_{h;L_p}}(\mathbf{I}_{L_p} - \mathbf{M}_p)^T\right), \end{aligned} \tag{35}$$

where $\mathbf{M}_p = (K/K_p)\,\mathbf{\Psi}_{L_p\times N_p}\mathbf{\Psi}_p^T$. The contribution caused by the noise is proportional to N_p/B_p, whereas the contribution caused by the modeling error decreases with increasing

N_p/B_p. The MSEE (35) can be minimized by proper selection of the ratio N_p/B_p.

V. Complexity Analysis and Comparison

As a measurement of computational complexity we count the number of real additions (ADDs), real multiplications (MULs) and accesses to lookup tables (LUTs), *per symbol* and *per iteration*. Computations that are performed on pilot symbols only are ignored, as their complexity is negligible in the usual case where the ratio K/K_p is large (say, $K/K_p > 10$).

The computational complexity of the proposed DCT-based receiver consists of two main contributions. The first contribution results from updating the messages $L_{d,k}^{(\gamma)}(\omega_i)$, and depends on the type of code and mapping used. This step is common to all receivers that apply the SPA to the FN $p(\mathbf{a}|\mathbf{b})$ (e.g., [1]–[6], [8], [9]), and its efficiency originates from a further factorization of $p(\mathbf{a}|\mathbf{b})$ [13], [16]. The second contribution results from updating the messages $L_{u,k}^{(\gamma)}(\omega_i)$ in the opposite direction, and depends on how the receiver models $p(\mathbf{r}|\mathbf{a})$; our DCT-based receiver computes $p(\mathbf{r}|\mathbf{a})$ according to (21).

Let us first consider the complexity associated with computing $L_{d,k}^{(\gamma)}(\omega_i)$. Various reduced-complexity implementations of the SPA applied to the upper FN in Fig. 1 have been proposed for (bit-interleaved) coded modulation [23], [24]. By means of example, we consider the case of an LDPC code with rate R_c and parity check matrix with (average) column weight w_c and (average) row weight w_r. According to [23], the computational complexity of the associated SISO decoding is roughly equal to $(2w_c+1)\log_2 M + 3w_r(1-R_c)\log_2 M$. In addition, SISO demapping using the standard SPA updating rules involves $\left(\frac{M}{2}+2\right)\log_2 M$ ADDs and $2\left(\frac{M}{2}-1\right)\log_2 M$ evaluations (assumed to be implemented using a LUT) of the function $\max^*(x,y) = \log(e^x + e^y)$.

For the proposed DCT-based receiver, updating the messages $L_{u,k}^{(\gamma)}(\omega_i)$ involves the evaluation of (27)-(29), (18), (17) and (24).[5] Computing the APPs in (27), the normalization coefficients from (28) and the weighted sum (29) involves $4M-5$ ADDs, $2M-1$ MULs and $M-1$ LUTs, where we assume that the $\exp\{\cdot\}$ operation is implemented by a LUT. The straightforward evaluation of the DCT in (17)-(18) requires a number of operations that is proportional to N. For large N, fast DCT algorithms [19], [20] can be used that compute any number of DCT coefficients with a complexity order of $\log(L)$. In general, the evaluation of the DCTs in (17)-(18) requires $F_+(N,L)$ ADDs and $F_\times(N,L)$ MULs, with

$$F_+(N,L) \approx 4\min\left(N, \frac{3}{2}\log_2 L\right), \qquad (36)$$
$$F_\times(N,L) \approx 4\min\left(N, \frac{1}{2}\log_2 L\right),$$

where L is assumed a power of 2 and the approximations hold for small N/L and large L. In addition, $M-1$ ADDs and $2M-2$ MULs are required for computing $L_{u,k}^{(\gamma)}(\omega_i)$ from (24). This brings the total complexity associated with computing $L_{u,k}^{(\gamma)}(\omega_i)$ in case of DCT to $10M + 4\min\left(N, \frac{3}{2}\log_2 L\right) + 4\min\left(N, \frac{1}{2}\log_2 L\right) - 10$.

For the purpose of comparison, we also consider the complexity of updating $L_{u,k}^{(\gamma)}(\omega_i)$ in the case of the receiver proposed in [1]; this receiver will be referred to as Tikh receiver, because some of the FG messages are represented by Tikhonov distributions. For an MPSK constellation with $|a_k|^2 = E_s$, and following the notation from [1], Tikh computes the auxiliary quantities α_k, ξ_k, $a_{f,k}$ and $a_{b,k}$ for $k = 0, 1, ..., K-1$ and outputs the probabilities $P_u(a_k = \omega_i) \propto \exp\left\{L_{u,k}^{(\gamma)}(\omega_i)\right\}$. The variable α_k is the first order moment of a_k with respect to the probability $P_d(a_k = \omega_i) \propto \exp\left\{L_{d,k}^{(\gamma)}(\omega_i)\right\}$. The evaluation of α_k therefore requires $(M-1)$ evaluations of the function $\exp\{\cdot\}$ (using a LUT), $(M-1)$ ADDs to evaluate a normalization factor and $(2M-3)$ ADDs and $(2M-1)$ MULs to compute the weighted sum of the constellation symbols. Evaluating $\xi_k = r_k\alpha_k^*\left(N_0 + E_s - |\alpha_k|^2\right)^{-1}$ from α_k and r_k, further requires 4 ADDs and 8 MULs. The value of $a_{f,k}$ is computed recursively as $a_{f,k} = (a_{f,k-1} + \xi_{k-1})\left(1 + \sigma_{PN,Rx}^2|a_{f,k-1} + \xi_{k-1}|\right)^{-1}$, which involves 4 ADDs, 5 MULs and 1 LUT. Here, the quantity $\sigma_{PN,Rx}$ is a design parameter; Tikh is derived assuming a Wiener PN model (see Section VII) with $\sigma_{PN,Rx}$ denoting the standard deviation of the phase increment over a symbol interval. The value of $a_{b,k}$ is computed in a similar way. The probabilities $P_u(a_k = \omega_i)$ are computed as $I_0(|a_{f,k} + a_{b,k} + \dot{r}_k(\omega_i)|)$, where $\dot{r}_k(\omega_i) = \frac{2}{N_0}r_k\omega_i^*$ needs to be evaluated only once before the start of the iterations and $I_0(\cdot)$ is the modified Bessel function of the first kind of the 0th order. Using the approximation $\log(I_0(|x|)) \approx |x|$, the messages $L_{u,k}^{(\gamma)}(\omega_i)$ generated by the Tikh receiver can be computed as $|a_{f,k} + a_{b,k} + \dot{r}_k(\omega_i)| - |a_{f,k} + a_{b,k} + \dot{r}_k(\omega_0)|$, which requires $4M+1$ ADDs, $2M$ MULs and M LUTs.[6] Thus, the total complexity of updating the messages $L_{u,k}^{(\gamma)}(\omega_i)$ in the Tikh receiver is $13M + 27$.

The above results are summarized in Table I . The parameter W for Tikh will be introduced in the next section. For a straightforward adaptation of the algorithm from [1] to a block-wise processing, it suffices to set W equal to L. As we have pointed out in Section IV that the PN estimation performance mainly depends on the ratio N/L, N should be considered as roughly proportional to L. As a result, the computational complexity of the DCT receiver, per symbol and per receiver iteration, increases in proportion to L, for small L, and in proportion to $\log_2(L)$, for large L. Decreasing both N and L proportionally, decreases the number of computations without having a (large) effect on the performance. This property allows to control the receiver's complexity.

Table II illustrates that the complexity of updating $L_{u,k}^{(\gamma)}(\omega_i)$ for the Tikh receiver exceeds the complexity of updating $L_{d,k}^{(\gamma)}(\omega_i)$ (assuming the LDPC code used in Section VII). This

[5] The complexity of the initialization is neglected as compared to that of executing Γ iterations. This is motivated by the fact that the number of iterations is typically large, while the complexity of the initialization is usually of the same order as the complexity of a single iteration.

[6] When counting the number of operations for Tikh, we have assumed that computing the magnitude $|x| = \sqrt{\Re\{x\}^2 + \Im\{x\}^2}$ of a complex number x involves 2 MULs, 1 ADD and 1 LUT for the square-root operation.

update:	$L_{d,k}^{(\gamma)}(\omega_i)$	$L_{u,k}^{(\gamma)}(\omega_i)$	
		DCT	Tikh
ADDs	$\left((2w_c+1)+3w_r(1-R_c)+\frac{3}{2}M\right)\log_2 M$	$5M-6+4\min\left(N,\ \frac{3}{2}\log_2 L\right)$	$(4M+5)\frac{W}{L}+(3M+4)$
MULs		$4M-3+4\min\left(N,\ \frac{1}{2}\log_2 L\right)$	$(2M+5)\frac{W}{L}+(2M+12)$
LUTs		$M-1$	$(M+1)\frac{W}{L}+M$

Table I
COMPUTATIONAL COMPLEXITY PER SYMBOL AND PER ITERATION, ASSUMING MPSK MAPPING, AN LDPC CODE WITH RATE R_c AND AVERAGE WEIGHT w_c (w_r) OF THE COLUMNS (ROWS) OF THE PARITY CHECK MATRIX AND DCT OPERATING ON BLOCKS OF SIZE L AND ESTIMATING N DCT COEFFICIENTS PER BLOCK.

M	$L_{u,k}^{(\gamma)}(\omega_i)$	$L_{d,k}^{(\gamma)}(\omega_i)$
2	53	17
4	79	40
8	131	78
16	235	152

Table II
COMPUTATIONAL COMPLEXITY PER SYMBOL AND PER ITERATION: NUMERICAL RESULTS FOR TIKH RECEIVER WITH $W=L$, 8PSK MAPPING AND LDPC CODE WITH $(R_c, w_r, w_c)=(\frac{2}{3}, 10, \frac{4}{3})$.

justifies our attempt to decrease the complexity of the SISO demodulation step by introducing the DCT receiver.

DCT and Tikh have similar memory requirements. Both receivers require a lookup table for the $\exp\{\cdot\}$ operator. Both receivers also need to store $\dot{r}_k(\omega_i)$ for $k \in I$ and $i=1,2,...,M-1$, and $(L_{d,k}^{(\gamma)}(\omega_i),\ L_{u,k}^{(\gamma)}(\omega_i))$ for $k \in I_d$ and $i=1,2,...,M-1$. Tikh further requires a lookup table for the $\sqrt{\cdot}$ operator, the storage of r_k for $k \in I$ ($2K$ real-valued quantities), and, additional storage space for at least $4K$ real-valued numbers (because all ξ_k, $k \in I$ and all $a_{f,k}$ (or $a_{b,k}$) $k \in I$ need to be stored simultaneously at some point in the computation of $L_{u,k}^{(\gamma)}(\omega_i)$). In contrast, DCT requires storing N real-valued DCT basis vectors of size L (assuming a straightforward implementation of the DCT operations), the storage of $\Delta\dot{r}_k(\omega_i)$ $k \in I$ ($2KM$ real-valued quantities), and, for the computation of $L_{u,k}^{(\gamma)}(\omega_i)$, additional storage space for at least $2(L+N) \leq 4L \leq 4K$ real-valued numbers (for the simultaneous storage of vectors $\mathbf{v}_h$ and $\mathbf{x}_h$ for given h). For a fixed ratio N/L, decreasing the block size by a factor α, decreases the number of DCT basis vector elements by a factor α^2 and the number of elements in $\mathbf{v}_h$ and $\mathbf{x}_h$ by a factor α. This allows to limit the memory requirements of the proposed receiver.

VI. PARALLEL PROCESSING

In section V we have investigated the receiver complexity, expressed as the number of operations per symbol and per iteration. Multiplying this complexity figure by the symbol rate and by the number of iterations gives the required number of operations per unit of time. In recent years, parallel computing on a (large) number of cooperating processors, usually in the form of multiple cores on a single chip, has become an important tool to increase processing speed. When the receiver is equipped with multiple (say, S) cores operating in parallel, the following advantages result (at the expense of an increased hardware cost): (i) for a given symbol rate, encoding and mapping, the time to process a frame reduces by a factor of (at most) S, so that latency is decreased as compared to a single-core receiver; (ii) for a given encoding and mapping, the symbol rate that can be handled by the receiver increases by a factor of (at most) S as compared to a single-core receiver (in that case the channel bandwidth and the received power should also increase by the same factor in order to accomodate the increased symbol rate and to maintain the same E_b/N_0 for satisfactory error performance). The general trend in processor development has moved to chips with tens or even hundreds of cores [25], [26]. Today, manufacturers provide digital signal processing platforms for communication systems with up to 72 [27] and up to 273 [28] processing units.

The PN estimation and SISO demodulation of the proposed DCT-based receiver are well-suited to exploit the presence of multiple cores. Because operations are performed independently on different blocks, they can be split over different cores without need for these cores to act in synchrony or to exchange information. When S cores are available, for given frame size K the receiver selects the block size L as $L=K/S$ so that each core processes a single block per frame.

As far as the remaining receiver operations are concerned, the upward/downward scheduling proposed in Section III is generally well suited for a parallel implementation of the receiver. It is clear that the FNs in Fig. 2 can update their messages on a blockwise basis. In addition, demapping and decoding for LDPC coded modulation intrinsically has a high degree of parallelism. The operations involved in computing the messages $L_{d,k}^{(\gamma)}(\omega_i)$ naturally partition into FN operations and VN operations. All FN operations (all VN operations) are independent and therefore different subsets of FN operations (VN operations) can be executed simultaneously on different cores without affecting the error performance. Considering that the typical amount of both FN and VN operations is large, it is fair to assume that the number of operations per core required for computing $L_{d,k}^{(\gamma)}(\omega_i)$ scales inversely proportional with the number of cores. As FN operations depend on VN operations and vice versa, cores need to exchange information. The efficient organization of this information exchange is one of the practical challenges of a multicore implementation of LDPC decoders, see e.g., [29]–[31]. Unlike LDPC decoding, trellis-based decoding, as, for example, with convolutional codes and turbo codes, involves recursive calculations that prevent parallel computation.

A block-processing implementation of Tikh is also possible and requires the block-wise computation of the messages $L_{u,k}^{(\gamma)}(\omega_i)$. This can be realized by partitioning the received

data frame into blocks, and running *independent* recursions to compute the parameters $a_{f,k}$ and $a_{b,k}$ over each block [4]. Because the values of these parameters are unknown at the block boundaries, all these recursions are initialized with zeros which in each block gives rise to a transient during which the accuracy of $a_{f,k}$ and $a_{b,k}$ is poor. The effect of this transient on the overall performance of the receiver is negligible for blocks that are long as compared to the transient duration, but increases significantly for shorter blocks. A practical work-around, which reduces the performance degradation due to the transient effect at the expense of an increase in complexity, is described next. When S cores are available, the receiver for frame size K runs the recursions for computing $a_{f,k}$ and $a_{b,k}$ over blocks of size W larger than $L = \frac{K}{S}$. Each block overlaps with the previous one over $(W - L)$ symbol periods. The recursions for $a_{f,k}$ and $a_{b,k}$ are run over the entire block of size W, yet the $0.5\,(W - L)$ first and last values of $a_{f,k}$ and $a_{b,k}$ are discarded, and the messages $L_{u,k}^{(\gamma)}(\omega_i)$ are only computed for the L remaining values of k. As indicated in Table I, the computational complexity (per symbol and per iteration) of a block-processing Tikh receiver increases with the ratio W/L but does not depend on L and W separately. In the numerical results section we will show that the required value of W/L strongly depends on the ratio K/K_p.[7]

Parallel processing also affects the memory requirements of a receiver. Without going into details, it is clear that, depending on the specific implementation, the lookup tables employed by DCT and Tikh may need to be stored S times, i.e., once for every core. This is the case for the $\exp\{\cdot\}$-table (DCT and Tikh), for the $\sqrt{\cdot}$-table (Tikh only) and for the storage of the DCT basis vector elements (DCT only). The latter implies that, for a fixed ratio N/L, decreasing the block size by a factor of α, will decrease the total number of DCT basis vector elements to be stored by a factor of only α rather than α^2. Also note that, for Tikh, the total number of complex-valued quantities ξ_k that need to be stored simultaneously at some point in the computation of $L_{u,k}^{(\gamma)}(\omega_i)$ is W/L times as large for block-processing Tikh as for the conventional Tikh operating on the entire frame.

As opposed to Tikh, DCT does not involve recursive calculations in each block. The latter receiver therefore offers more opportunities to split the workload associated with a particular block, in its turn, over multiple cores (simultaneously executing independent calculations). The investigation of the effect of using multiple cores per block falls beyond the scope of the present paper.

VII. Numerical Results and Discussion

We now present numerical results regarding the error performance and the computational load of the proposed receiver. The Wiener model [32] was used to characterize the PN, i.e., $\theta_k = \theta_{k-1} + \phi_k$, with the initial carrier phase θ_0 uniformly distributed in $[-\pi, \pi[$ and $\{\phi_k\}$ statistically independent Gaussian random variables with zero mean and standard deviation σ_{PN}, descriptive of the Wiener PN intensity. Conventional turbo and LDPC decoding procedures were employed. Blocks were processed sequentially, one after the other, without exchange of information between blocks. The resulting error performance is the same as for a receiver that would have processed the different blocks in parallel, simultaneously. We considered pseudo-random bit interleaved coded 8PSK modulation and pilot symbols were inserted according to (2). The error performance is expressed as the average fraction of erroneously detected information bits (bit error rate or BER) over as many codewords as required to encounter 100 erroneously detected codewords. The SNR is expressed as E_b/N_0, with $E_b = KE_s/K_b$ the energy per information bit. For the purpose of comparison, results are also presented for the Tikh receiver using $\sigma_{PN,Rx} = \sigma_{PN}$ [8].

To illustrate the potential benefits of DCT in a parallel implementation setting, we first consider a rate-2/3 irregular repeat accumulate LDPC code with $K_b = 43200$, specified in the DVB-S2 standard [33]. The check matrix $\mathbf{H}$ contains $\frac{1}{10}K_b$ columns with degree 13, $\frac{9}{10}K_b$ columns with degree 3 and $\frac{1}{2}K_b$ columns with degree 2. All $\frac{1}{2}K_b$ rows in $\mathbf{H}$ have degree 10. We take $K_p = 540$, yielding $K = 22140$ and $K/K_p = 41$, in agreement with [33]. A maximum of Γ=40 receiver iterations is performed. The standard deviation σ_{PN} of the phase noise increments is set to 3°. The perf. sync. receiver using perfect PN estimates (for a transmission with the same code and the same ratio K/K_p) achieves a BER of 10^{-6} at an E_b/N_0 value of about $(E_b/N_0)_{10^{-6},\mathrm{dB}}^{\mathrm{perf.\,sync.}}$ = 3.76 dB.

Fig. 3 illustrates the effect of the design parameters of DCT, i.e., $\{L_p,\ N_p\}$ and $\{L,\ N\}$ to be used during initialization (top figure) and iterations (bottom figure), respectively. The operating point is $E_b/N_0 = 4.55$ dB (i.e., 0.79 dB higher than $(E_b/N_0)_{10^{-6},\mathrm{dB}}^{\mathrm{perf.\,sync.}}$ to anticipate the power efficieny loss due to imperfect synchronization).[9] For a given number (N_p or N) of DCT coefficients to be estimated, an optimum block size (L_p or L) exists: a too small block size results in a high sensitivity to additive noise, while for a too large block size the PN cannot be accurately represented as a linear combination of the considered number of DCT basis functions. The minimum BER in Fig. 3 is achieved for a fixed ratio $N_p/L_p \approx 1/125$ ($\gamma = 0$) or $N/L \approx 1/30$ ($\gamma = 40$) and does not change much with N_p for $N_p \geq 4$ (top figure) or with N for $N \geq 4$ (bottom figure). For large L, several N give rise to BER values

[7] For completeness, we mention that, in contrast to Tikh, other FG-based receivers from [3] can be implemented in parallel without modification of the original algorithm; however, these receivers have been shown to be significantly more complex than the Tikh receiver and are therefore not further considered.

[8] It should be noted that this is usually not the optimal choice when suboptimal implementations are under consideration. Hence, especially for short blocks, a fine tuning of $\sigma_{PN,Rx}$ could result in a significant performance advantage.

[9] The bottom figure shows the BER after 40 iterations, assuming that, for given N, during initialization we have taken $N_p = N$ and L_p equal to the integer divider of K closest to the block length that minimizes the BER after the first iteration ($\gamma = 1$) for the considered value of N_p. This results in the following set of values for (L_p, N_p): $(123, 1)$, $(246, 2)$, $(369, 3)$, $(492, 4)$, $(738, 6)$ and $(1230, 10)$. This particular way of selecting (L_p, N_p) for given N essentially minimizes the BER after the first iteration ($\gamma = 1$), subject to the constraint that the initialisation complexity per block should be of the same order as the complexity for one iteration; it can be verified that for $N_p = N$ the DCT-related operations yield the same complexity per symbol when the block size is small.

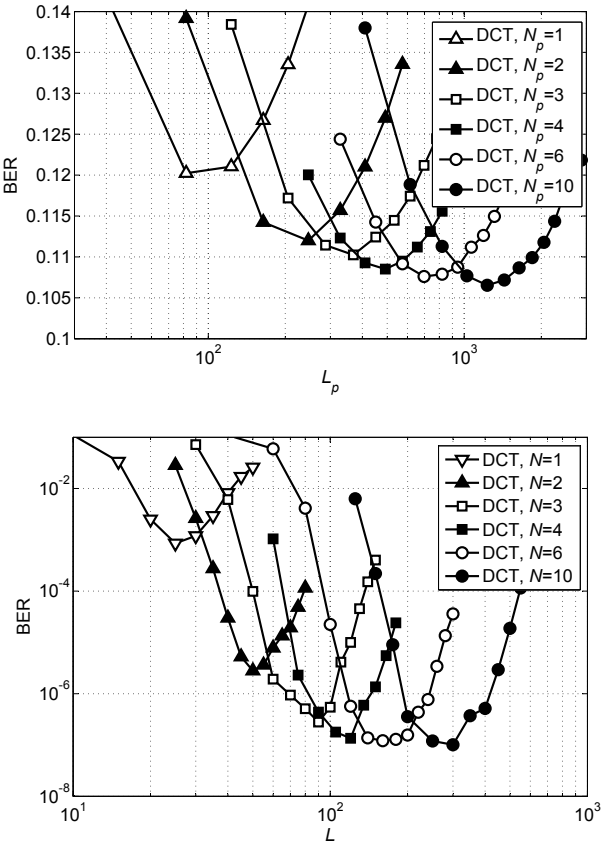

Figure 3. BER of DCT receiver after 1 iteration (top) and after 40 iterations (bottom), as a function of the block length L, at E_b/N_0 = 4.55 dB, for σ_{PN} = 3° and LDPC coded 8PSK signaling with K/K_p= 41, for several values of the number of estimated DCT coefficients (N_p or N).

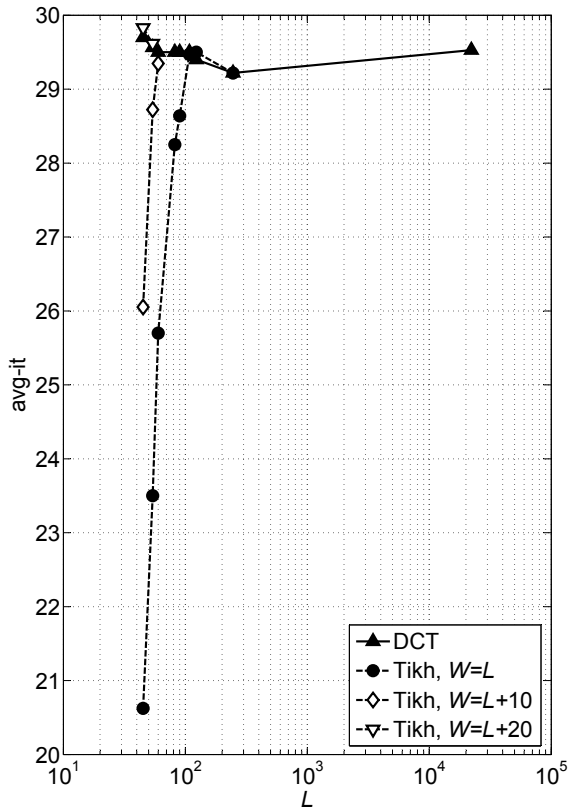

Figure 4. Average number of iterations (top) and avg-comp/bl (bottom) of DCT and Tikh with $\Gamma = 40$, at a target BER of 10^{-6} after a maximum of 40 iterations, as a function of the block size L, for σ_{PN} = 3° and LDPC coded 8PSK signaling with K/K_p = 41.

that are close to the minimum BER for the considered L; e.g., for $L = 200$, close-to-optimum BER values are achieved for $N \in [6, 10]$ (or, equivalently, for $N/L \in [0.03, 0.05]$). This indicates that, for given σ_{PN}, the performance is rather insensitive to moderate variations of the estimator's cut-off frequency around its optimum value.

When the receiver is equipped with multiple cores and each core processes a single block per frame, the parallel execution time for processing a frame is proportional to the average computational load per block (avg-comp/bl): this load is the product of the complexity per block and per iteration (L times the sum of the relevant entries in Table I for the computation of $L_{d,k}^{(\gamma)}(\omega_i)$ and $L_{u,k}^{(\gamma)}(\omega_i)$) with the average number of iterations (avg-it) performed by the respective receivers. Fig. 4 shows avg-it as a function of L for DCT and Tikh with $\Gamma = 40$, operating at BER = 10^{-6}. The avg-it for DCT lies between 29 and 30. Tikh with $W = L$ also performs 29 to 30 iterations on average when operating on blocks larger than $L = 100$, but its avg-it is significantly less for smaller L. For small values of L, the avg-it for Tikh increases significantly with increasing W.

We now consider the error performance degradation of DCT and Tikh with respect to perf. sync., expressed as the power efficiency loss. This loss is defined as

$$(\Delta E_b/N_0)_{10^{-6},\mathrm{dB}} = (E_b/N_0)_{10^{-6},\mathrm{dB}} - (E_b/N_0)_{10^{-6},\mathrm{dB}}^{\mathrm{perf.\,sync.}}, \tag{37}$$

with $(E_b/N_0)_{10^{-6},\mathrm{dB}}$ denoting the value of E_b/N_0 in dB required for DCT or Tikh to obtain a BER of 10^{-6}. Fig. 5 shows the avg-comp/bl as a function of the power efficiency loss for various L that are dividers of $K = 22140$; in addition, for Tikh, we also take $(W - L) \in \{0, 10, 20\}$. By varying the value of L (and also W in case of Tikh), one can trade-off execution time (indicated by avg-comp/bl) against error performance (indicated by $(E_b/N_0)_{10^{-6},\mathrm{dB}}$) and parallel system implementation cost (indicated by the number of cores, $S = 22140/L$).

- The vertical line at $(\Delta E_s/N_0)_{10^{-6},\mathrm{dB}}$= 0.58 dB indicates the power efficiency loss of the conventional Tikh operating on the entire frame without block-processing, and serves as a practical lower bound on the loss of the considered receivers. Relative to that, with $0.73\,\mathrm{dB} \leq (\Delta E_s/N_0)_{10^{-6},\mathrm{dB}} \leq 0.82\,\mathrm{dB}$, DCT yields only a very modest additional degradation.
- Reducing the execution time by processing blocks inde-

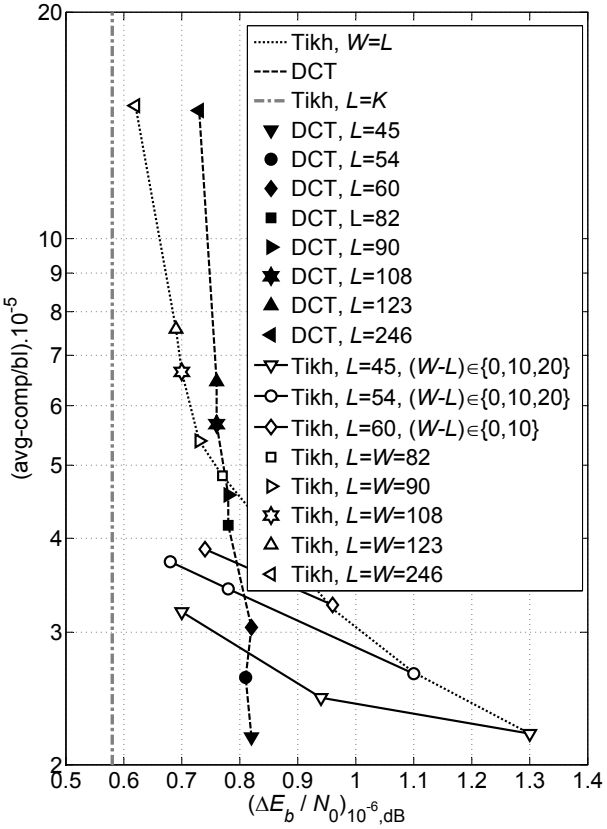

Figure 5. Avg-comp/bl as a function of power efficiency loss of DCT and Tikh receivers after maximum 40 iterations with respect to perf. sync. receiver, in terms of E_b/N_0 (in dB), at a target BER of 10^{-6}, for several block lengths L, for σ_{PN} = 3° and LDPC coded 8PSK signaling with K/K_p = 41.

pendently in parallel gives rise to some BER performance degradation; this is because the correlation between the PN values in different blocks cannot be exploited. The effect is rather small for DCT, quite severe for Tikh with $W = L$ (due to the zero initialization of the forward/backward recursions over each block), and significantly reduces for Tikh with given L when $(W - L)$ increases. However, the improved error performance of Tikh for $W > L$ comes at the expense of a larger avg-comp/bl. For example, for $L = 54$, increasing W from $W = L = 54$ to $W = L + 20 = 74$ yields a power efficiency gain of 0.45 dB but also an increase in avg-comp/bl of almost 50%.

- For given L, Tikh yields a larger avg-comp/bl than DCT when both receivers use a high degree of parallelization ($45 \leq L \leq 246$). Hence, using DCT instead of Tikh allows to realize a target avg-comp/bl using a significantly smaller amount of parallel cores $S = K/L$, at the limited cost of a small additional power efficiency loss.
- For given L (or, equivalently, given S), Tikh requires an at least 20% larger avg-comp/bl for achieving a power loss smaller than or equal to that of DCT.

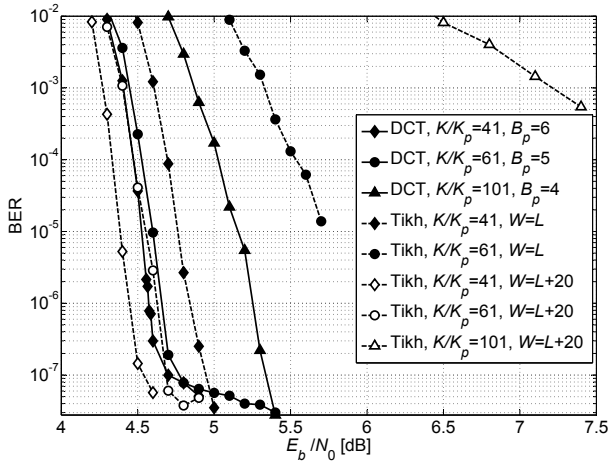

Figure 6. BER performance of Tikh with $W \geq L$ and of DCT with $N = 2$, as a function of E_b/N_0, for several values of K/K_p, and for $L = 54$, $\Gamma = 40$, σ_{PN} = 3° and LDPC coded 8PSK signaling.

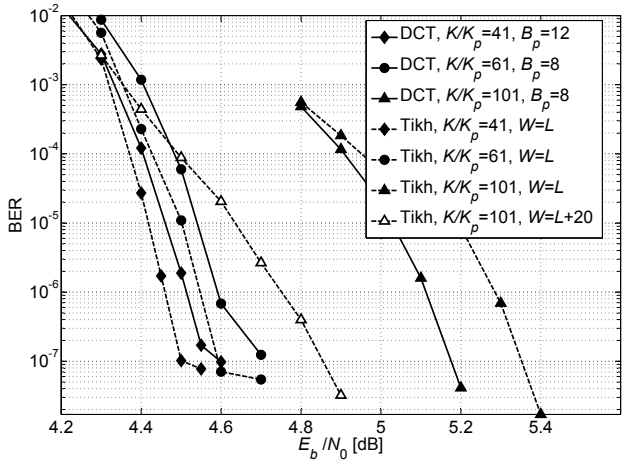

Figure 7. BER performance of Tikh with $W \geq L$ and of DCT with $N = 2$, as a function of E_b/N_0, for several values of K/K_p, and for $L = 108$, $\Gamma = 40$, σ_{PN} = 3° and LDPC coded 8PSK signaling.

Fig. 6 (for $L = 54$) and Fig. 7 (for $L = 108$) show the BER of DCT and Tikh with $W \in \{L, L + 20\}$ as a function of E_b/N_0. Results are presented for K/K_p = 41, 61, 101. The considered LDPC code is known to have an error floor below 10^{-7} [34]. However, increasing the ratio K/K_p degrades the BER performance above 10^{-7} of both Tikh and DCT. For both $L = 54$ and $L = 108$, this degradation is more significant for Tikh with $W = L$ than for DCT. The largest degradations occur for Tikh with W smaller than K/K_p, in which case not all blocks of size W contain pilot symbols. The DCT receiver can cope with large K/K_p by taking during initialization a large enough blocksize L_p so that the corresponding blocks contain a sufficient number B_p of pilot symbols. Furthermore, for given L, the minimum value of W that is required for Tikh to yield a performance comparable to that of DCT increases with K/K_p. This implies that the advantage of DCT over Tikh in terms of the computational load per block in the case of

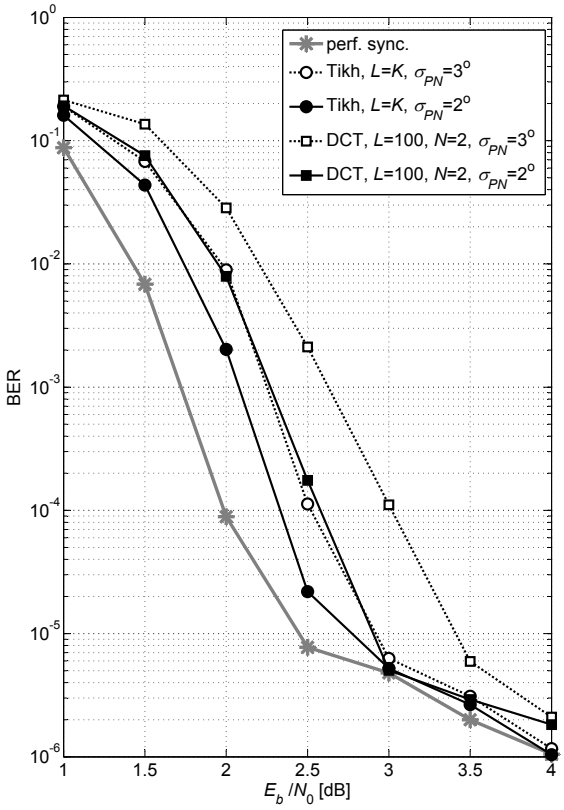

Figure 8. BER performance of Tikh operating on complete frame ($L = K$) and of DCT estimating $N = 2$ DCT coefficients per block of size $L = 100$, after 10 iterations, as a function of E_b/N_0, for Wiener PN with standard deviation σ_{PN} and turbo coded 8PSK signaling.

high parallelisation increases with increasing K/K_p.

Above we have compared DCT and Tikh in a parallel implementation setting, where the receiver is equipped with multiple cores and each core processes a single block per frame. However, in conventional single-core receivers, all blocks are processed sequentially rather than in parallel. In such a case, there is no reason for using Tikh with $L < K$, which causes only performance degradation and no complexity reduction as compared to Tikh with $L = K$. On the other hand, DCT with small L may provide an interesting low-complexity alternative to Tikh with $L = K$. To illustrate this, let us consider a rate 1/3 turbo code with $K_b = 946$ information bits (followed by 4 termination bits) that consists of the parallel concatenation of two identical rate 1/2 recursive systematic convolutional codes with generator sequences (11111) and (10001). $K_p = 50$ pilot symbols are inserted, such that $K = 1000$ and $K/K_p = 20$. Always a total of Γ=10 receiver iterations is carried out. In this case, a multi-core receiver implementation is cost-inefficient because the trellis-based turbo decoding algorithm involves recursive calculations that prevent parallel computation. Fig. 8 compares, for $\sigma_{PN} = 2^\circ, 3^\circ$, the BER of Tikh with $L = K$ to that of DCT with $N_p = N = 2$ and (L_p,L) = $(L_{p,\mathrm{opt}},L_{\mathrm{opt}})$ selected to minimize the BER at E_b/N_0=3 dB. For both values of σ_{PN}, $(L_{p,\mathrm{opt}},L_{\mathrm{opt}}) \approx (360,\ 100)$[10]. This indicates that the optimum design parameters N_p, L_p, N and L are robust against moderate changes in the PN intensity. According to Table I the computational complexity of the SISO demodulation process (per symbol and per iteration) is significantly higher for Tikh with $L = K$ than for DCT with $L = 100$ and $N = 2$ (131 versus 86 operations per symbol and per iteration, i.e., about 50% difference). Fig. 8 shows that, for $\sigma_{PN} = 3^\circ$, Tikh and DCT have a power efficiency loss at a BER of 10^{-5}, as compared to the perf. sync. receiver, that amounts to about 0.5 dB and 1.0 dB, respectively, indicating that DCT is only 0.5 dB worse than Tikh; for $\sigma_{PN} = 2^\circ$, DCT is only about 0.3 dB worse than Tikh. This indicates that the error performance degradation caused by the complexity reduction of the DCT receiver is rather small.

VIII. Conclusions and Remarks

The paper proposes a novel iterative SISO demodulator for coded PSK signals in AWGN with PN, which makes use of the DCT for compactly representing the time-varying phasor. Splitting the received frame into multiple blocks and processing these blocks independently allows to control the receiver's complexity and memory requirements by selecting suitable values for the design parameters. As opposed to the receivers from [1], [2], [4], [6], [8], the DCT-based demodulator does not involve recursive calculations, which makes it better suited for a parallel implementation on multiple cores. The block-processing gives rise to a performance degradation that depends of the type of coded modulation and the number of pilot symbols used. Results obtained for LDPC coded 8PSK with only 2.5% pilot symbols show that, for a large degree of parallelisation (many short blocks per frame and each core processes a single block), the state-of-the-art Tikh receiver from [1] requires an at least 20% larger average computational load per block for achieving a power loss smaller than or equal to that of DCT. This advantage of DCT over Tikh even slightly increases when the relative amount of pilot symbols decreases. Results obtained for turbo coded 8PSK with 5% pilot symbols further indicate that DCT can also provide a computationally efficient alternative to Tikh in a conventional sequential implementation, at the expense of only a small power efficiency loss.

Acknowledgment

This research has been funded by the Interuniversity Attraction Poles Program initiated by the Belgian Science Policy Office. The first author would like to acknowledge the financial support of the Fund for Scientific Research - Flanders (FWO Vlaanderen).

References

[1] G. Colavolpe *et al.*, "Algorithms for iterative decoding in the presence of strong phase noise," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 9, pp. 1748–1757, Sep. 2005.

[10]Because 360 is not an integer divisor of 1000, for DCT initialization we split the frame into two blocks of size 360 and one block of size 280.

[2] J. Dauwels and H.-A. Loeliger, "Phase estimation by message passing," in *2004 IEEE Int. Conf. Commun.*, Paris, France, 2004, pp. 523–527.
[3] G. Colavolpe, "On LDPC codes over channels with memory," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1757–1766, Jul. 2006.
[4] A. Barbieri *et al.*, "Joint iterative detection and decoding in the presence of phase noise and frequency offset," *IEEE Trans. Commun.*, vol. 55, pp. 171–179, Jan. 2007.
[5] L. Benvenuti *et al.*, "Code-aware carrier phase noise compensation on turbo-coded spectrally-efficient high-order modulations," in *8-th Int. Workshop on Signal Process. Space Commun.*, Catania, Italy, 2003, pp. 177–184.
[6] N. Noels *et al.*, "Performance analysis of ML-based feedback carrier phase synchronizers for coded signals," *IEEE Trans. Signal Process.*, vol. 55, pp. 1129–1136, Mar. 2007.
[7] R. Krishnan *et al.*, "Soft metrics and their performance analysis for optimal data detection in the presence of strong oscillator phase noise," *IEEE Trans. Commun.*, vol. 61, pp. 2385–2395, Jun. 2013.
[8] N. Kamiya and E. Sasaki, "Pilot-symbols assisted and code-aided phase error estimation for high-order QAM transmission," *IEEE Trans. Commun.*, vol. 61, pp. 4369–7380, Oct. 2013.
[9] J. Bhatti *et al.*, "Algorithms for iterative PN estimation based on a truncated DCT expansion," in *2011 IEEE 12th Int. Workshop Signal Process. Advances in Wireless Commun.*, San Francisco, CA, 2011, pp. 51–55.
[10] G. Ferrari *et al.*, "On linear predictive detection for communications with phase noise and frequency offset," *IEEE Trans. Veh. Technol.*, vol. 56, pp. 2073–2085, Jul. 2007.
[11] A. Anastasopoulos and K. M. Chugg, "Adaptive iterative detection for phase tracking in turbo coded systems," *IEEE Trans. Commun.*, vol. 49, pp. 2135–2144, Dec. 2001.
[12] T. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Mag.*, vol. 13, no. 6, pp. 47–60, Nov. 1996.
[13] N. Noels *et al.*, "A theoretical framework for soft information based synchronization in iterative (turbo) receivers," *EURASIP J. Wireless Commun. and Network.*, vol. 2005, no. 2, pp. 117–129, Apr. 2005.
[14] A. Patapoutian, "On phase-locked loops and Kalman filters," *IEEE Trans. Commun.*, vol. 47, pp. 670–672, May 1999.
[15] A. Spalvieri and L. Barletta, "Pilot-aided carrier recovery in the presence of phase noise," *IEEE Trans. Commun.*, vol. 59, pp. 1966–1974, Jul. 2011.
[16] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Processing Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.
[17] A. Chorti and M. Brookes, "A spectral model for RF oscillators with power-law phase noise," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 53, pp. 1989–1999, Sep. 2006.
[18] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. Colombus, OH: McGraw-Hill, 1984.
[19] G. Bi and L. W. Yu, "DCT algorithms for composite sequence lengths," *IEEE Trans. Signal Process.*, vol. 46, pp. 554–562, Mar. 1998.
[20] G. Bi, "Fast algorithms for type-III DCT of composite sequence lengths," *IEEE Trans. Signal Process.*, vol. 47, pp. 2053–2059, Jul. 1999.
[21] T. Ahmed *et al.*, "Discrete cosine transform," *IEEE Trans. Comput.*, vol. 23, pp. 90–93, Jan. 1974.
[22] J. Bhatti and M. Moeneclaey, "Feedforward data-aided phase noise estimation from a DCT basis expansion," *EURASIP J. Wireless Commun. and Network., Special Issue on Synchronization in Wireless Commun.*, vol. 2009, Jan. 2009.
[23] J. Chen *et al.*, "Reduced-complexity decoding of LDPC codes," *IEEE Trans. Commun.*, vol. 53, pp. 1288–1299, Aug. 2005.
[24] D. Fertonani *et al.*, "Reduced-complexity BCJR algorithm for turbo equalization," *IEEE Trans. Commun.*, vol. 55, pp. 2279–2287, Dec. 2007.
[25] K. Asanovic *et al.*, "The landscape of parallel computing research: A view from Berkeley," University of California, Berkeley, CA, Tech. Rep. UCB/EECS-2006-183, Dec. 2006.
[26] L. Karam *et al.*, "Trends in multi-core DSP platforms," *IEEE Signal Processing Mag.*, vol. 26, no. 6, pp. 38–49, Nov. 2009.
[27] (2013) Pc205. [Online]. Available: http://www.picochip.com/
[28] (2013) tile-gx8072. [Online]. Available: http://www.tilera.com/
[29] A. Tarable *et al.*, "Mapping interleaving laws to parallel turbo and LDPC decoder architectures," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2002–2009, Sep. 2004.
[30] J. Dielissen *et al.*, "Low cost LDPC decoder for DVB-S2," in *Proc. Design, Automation and Test in Europe*, Munich, Germany, 2006.
[31] M. Awais *et al.*, "A novel architecture for scalable, high throughput, multi-standard LDPC decoder," in *Euromicro Conf. Digital Syst. Design*, Oulu, Finland, Aug. 2011, pp. 340–347.
[32] H. Meyr *et al.*, *Digital Communication Receivers*, 1st ed. Hoboken, NY: John Wiley & Sons Inc., 1998.
[33] *Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications (DVB-S2)*, ETSI Std. EN 302 307 v1.3.1, 2013.
[34] B. Butler and P. Siegel, "Error floor approximation for LDPC codes in the AWGN channel," in *49th Annu. Allerton Conf. Commun., Control, and Computing*, Monticello, IL, Sep. 2011, pp. 204–211.

Nele Noels (S'04, M'11) received the diploma of electrical engineering and the Ph.D. degree in electrical engineering from Ghent University, Gent, Belgium in 2001 and 2009, respectively. She is currently part-time Professor at the Department of Telecommunications and Information Processing (TELIN), Gent University and a postdoctoral researcher supported by the Research Foundation-Flanders (FWO Vlaanderen). Her main research interests are in statistical communication theory, carrier and symbol synchronization, bandwidth-efficient modulation and coding, massive MIMO, optical OFDM, satellite and mobile communication. She is (co-)author of over 40 academic papers in international journals and conference proceedings and recipient of several scientific awards. In 2010, she received the Scientific Award Alcatel Lucent Bell for the best Belgian thesis concerning an original study of information and communication technology, concepts and/or applications.

Jabran Bhatti received the Masters degree in electrical engineering and the Ph.D. degree in electrical engineering from Ghent University, Gent, Belgium, in 2006 and 2012, respectively. From 2006 till 2012 he was a member of the Department of Telecommunications and Information Processing at Ghent University.

Herwig Bruneel was born in Zottegem, Belgium, in 1954. He received the Masters degree in electrical engineering, the Masters degree in computer science, and the Ph.D. degree in computer science, all from Ghent University, Belgium, in 1978, 1979, and 1984, respectively. He is full Professor in the Faculty of Engineering and head of the Department of Telecommunications and Information Processing at the same university. He also leads the SMACS Research Group within this department. His main personal research interests include stochastic modelling and analysis of communication systems, discrete-time queueing theory, and the study of ARQ protocols. He has published more than 400 papers on these subjects and is coauthor, with B. G. Kim, of the book Discrete-Time Models for Communication Systems Including ATM (Boston, MA: Kluwer, 1993). From October 2001 to September 2003, he served as the Academic Director for Research Affairs at Ghent University. Dr. Bruneel has held a career-long Methusalem grant from the Flemish Government at Ghent University, specifically on Stochastic Modelling and Analysis of Communication Systems since 2009.

Marc Moeneclaey (M'93, SM'99, F'02) received the diploma of electrical engineering and the Ph.D. degree in electrical engineering from Ghent University, Gent, Belgium, in 1978 and 1983, respectively. He is Professor at the Department of Telecommunications and Information Processing (TELIN), Gent University. His main research interests are in statistical communication theory, (iterative) estimation an detection, carrier and symbol synchronization, bandwidth-efficient modulation and coding, spread-spectrum, satellite and mobile communication. He is the author of more than 400 scientific papers in international journals and conference proceedings. Together with Prof. H. Meyr (RWTH Aachen) and Dr. S. Fechtel (Siemens AG), he co-authors the book "Digital Communication Receivers - Synchronization, channel estimation, and signal processing." During the period 1992-1994, was Editor for Synchronization, for the IEEE Transactions on Communications. He served as co-guest editor for special issues of the Wireless Personal Communications Journal (on Equalization and Synchronization in Wireless Communications) and the IEEE Journal on Selected Areas in Communications (on Signal Synchronization in Digital Transmission Systems) in 1998 and 2001, respectively.