Reduced-Complexity Joint Baseband Compensation of Phase Noise and I/Q Imbalance for MIMO-OFDM Systems

Payam Rabiei, *Student Member IEEE*, Won Namgoong *Senior Member IEEE*, and Naofal Al-Dhahir, *Fellow IEEE*,

***Abstract*— The maximum likelihood estimate of the impulse response of a frequency-selective channel in the presence of phase noise and I/Q imbalance is derived. The complexity of the joint estimator is reduced using approximate cost functions for both phase noise and I/Q imbalance. The proposed estimator is first applied to OFDM transmission in a single-input single-output system and then generalized to multi-input multi-output OFDM systems. The bit error rate performance of popular space-time codes for two and four transmit antennas is evaluated under zero-forcing, minimum mean squared error and maximum likelihood detection rules. An expression for the residual inter-carrier interference variance after phase noise and I/Q imbalance compensation is derived and compared to the uncompensated case. Significant improvement in signal-to-interference-noise is obtained with the proposed algorithm.**

***Index Terms*— OFDM, MIMO, phase noise, I/Q imbalance, channel estimation.**

I. Introduction

MULTI-INPUT MULTI-OUTPUT orthogonal frequency division multiplexing (MIMO-OFDM) is a bandwidth-efficient data transmission scheme which has been widely adopted in several broadband standards such as WLAN IEEE802.11a/g/n [1], WiMAX IEEE802.16, and LTE. Many OFDM receiver front-ends are based on the direct-conversion architecture which enables a low-power and highly-integrated implementation. However, analog circuit imperfections such as phase noise (PHN) and I/Q imbalance distort the received signal [2], motivating digital baseband techniques to compensate for these impairments at practical complexity.

PHN is a result of the random mismatch between the phases of the local oscillator (LO) and the input carrier signal [3] and has two effects on the received signal. The first effect is known as the common phase error (CPE) [4], which is the same random phase rotation in all sub-carriers regardless of their index. The second effect is the inter-carrier interference (ICI) caused by the loss of orthogonality among adjacent sub-carriers [5]. Since PHN changes continuously, it has to be estimated for every sample of an OFDM symbol in a data packet.

I/Q imbalance is the mismatch in amplitude and phase of the in-phase and quadrature-phase branches of the receive path. I/Q mismatch occurs [6] because of the error in the nominally 90° phase shifter and the mismatch between the amplitudes of the LO in-phase and quadrature-phase outputs.

A. Existing Work

The effect of PHN and I/Q imbalance on the performance of OFDM systems has been studied in several papers including [3], [7], [8]. Assuming perfect I/Q imbalance and channel knowledge, CPE estimation and compensation was considered in [4], [9]. Joint iterative estimation and compensation of the channel, CPE and PHN-induced ICI was considered in [5], [10]. In particular, when performing joint PHN and channel estimation in the full-pilot stage[1], the channel is estimated using an initial guess for PHN and then re-estimated based on the previous channel estimate. In the data transmission stage, CPE is estimated and removed using scattered pilots followed by data detection. Then, PHN is re-estimated based on the detected data from the previous step (i.e. in a decision-directed manner). These schemes iterate until convergence; hence, latency and implementation complexity are two major issues in these schemes. A time-domain data-aided joint PHN and channel estimation technique was presented in [11] where I/Q imbalance is ignored. Polynomial expansion of PHN is considered in [12] and the maximum likelihood (ML) estimate of the polynomial coefficients is provided. Again, the scheme in [12] is decision-directed and suffers from error propagation. The effects of PHN on OFDM systems are studied in a fast-fading environment in [13], [14], where perfect channel knowledge was assumed in [13].

Assuming perfect PHN, the effect of I/Q imbalance on OFDM systems was investigated in [6], [15], [16] for single-input single-output (SISO) channels and generalized to MIMO channels in [17], where the signal model was presented and the performance of different equalizers was investigated with perfect channel knowledge. Joint estimation of channel and I/Q imbalance parameters was studied in [18] but the interference from adjacent subcarriers due to PHN was ignored. Joint estimation and compensation of PHN and I/Q imbalance was considered in [19], but the channel was assumed known. In [20], [21], an iterative scheme was proposed to jointly estimate PHN, I/Q imbalance and the channel response. This scheme suffers from error propagation if un-coded data at the output of the slicer are used for iteration or from large latency and high implementation complexity if coded data at the output of the Viterbi decoder are used for iteration.

In [22], the channel was initially coarsely estimated in the full-pilot stage by ignoring the I/Q imbalance and PHN. Then, the I/Q imbalance parameters were computed using these channel estimates. The estimated I/Q imbalance parameters were subsequently used to refine the original channel estimate, in a non-iterative manner. In the data detection stage, after removing the I/Q imbalance, only CPE was estimated and compensated for prior to data detection.

This work was supported in part by the National Science Foundation (NSF) under contracts ECCS 10-02334 and CCF 07-33124 and by gifts from RIM Inc. This work was presented in part at IEEE Globecom 2009.

[1] In packet-based OFDM systems, a number of full-pilot OFDM symbols are transmitted prior to data transmission for synchronization and channel estimation purposes.

B. Our Contributions

1. We develop a low-complexity joint ML estimator of the channel, PHN and I/Q imbalance[2] in a SISO communication system, using the full-pilot OFDM symbol[3] (also known as the preamble) at the beginning of each packet. In contrast to the existing estimators, we derive *decoupled* expressions for PHN and I/Q imbalance by approximating the joint ML cost function using the assumption of small PHN and I/Q imbalance. Therefore, we are able to provide simple and non-iterative estimators which are practical and well-suited for implementation.

2. Assuming constant channel and I/Q imbalance over one OFDM packet, a phase-interpolation-based technique is proposed to estimate PHN using equally-spaced pilots in data OFDM symbols.

3. The above two algorithms for the SISO scenario are then generalized to the MIMO scenario where more channel taps, I/Q imbalance and PHN parameters are estimated [25].

The rest of this paper is organized as follows: in Section II, the notation is described and a key result is presented and used to derive the ML channel estimator. In Section III, the OFDM system model is presented and the joint estimator is derived during the full-pilot transmission stage. In Section IV, data detection and PHN interpolation during data OFDM symbol transmission is studied. In Section V, the extension to MIMO channel is presented. Numerical results are given in Section VI and the paper is concluded in Section VII.

II. Preliminaries

The notation of this paper is as follows: the lower/upper-case bold-face letters $\boldsymbol{a}/\boldsymbol{A}$, are frequency-domain vectors/matrices and the lower/upper-case letters with over-bar $\bar{a}/\bar{A}$, are time-domain vectors/matrices, respectively. Moreover, $\bar{a}^q_{im}(n)$ denotes the nth element of the time-domain vector $\bar{a}$, transmitted from the ith transmit antenna and received by the mth receive antenna over the qth time slot. The (m,n) element of a matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{A}(m,n)$.

In the following, we present and prove a key result which will be used to derive the channel estimator in the presence of PHN and I/Q imbalance.

Proposition:
Assuming a complex vector $\bar{r}$ has the following form

$$\bar{r} = \mu \bar{D}_1 \bar{x} + \nu \bar{D}_2 \bar{x}^* + \bar{z} \tag{1}$$

where $(\cdot)^*$ is the complex-conjugation operation, μ and ν are complex scalars, and $\bar{z} \sim \mathcal{CN}(0, \sigma_z^2 \boldsymbol{I})$. The ML estimate of the vector $\bar{x}$ is given by

$$\widehat{\bar{x}} = (|\mu|^2 \bar{D}_1^H \bar{D}_1 - |\nu|^2 \bar{D}_2^T \bar{D}_2^*)^{-1} (\mu^* \bar{D}_1^H \bar{r} - \nu \bar{D}_2^T \bar{r}^*) \tag{2}$$

where $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and Hermitian transpose operations, respectively.
Proof: see Appendix I. ■

III. SISO Full-Pilot Stage Estimations

A. Channel Estimation

Considering an OFDM system suffering from PHN and I/Q imbalance at the receiver, the received signal in the time-domain can be modeled as [22]

$$\bar{r} = \mu \bar{E} \bar{H} \bar{x} + \nu \bar{E}^H \bar{H}^* \bar{x}^* + \bar{z} \tag{3}$$

where $\bar{E}$ is the diagonal PHN matrix with (n,n) element given by $\exp(j\phi(n))$ where $\phi(n) = \phi(n-1) + \varepsilon$ and ε is a Gaussian random variable with zero mean and variance $2\pi\beta T_s$. The parameter β is the 3 dB PHN bandwidth and T_s is the OFDM sampling period [26]. Furthermore, $\bar{H}$ is the circulant channel matrix, $\bar{x}$ is the full-pilot OFDM symbol, $\bar{z}$ is the additive white Gaussian noise with variance σ_z^2 and μ and ν are the I/Q imbalance parameters defined as [2]

$$\begin{aligned} \mu &= \cos(\theta/2) + j\epsilon \sin(\theta/2) \\ \nu &= \epsilon \cos(\theta/2) - j \sin(\theta/2) \end{aligned} \tag{4}$$

where θ is the I/Q phase imbalance and ϵ is the amplitude imbalance. Taking the discrete Fourier transform (DFT) of $\bar{r}$, we can write the received signal in the frequency-domain as

$$\boldsymbol{r} = F\bar{r} = \mu \boldsymbol{P} \boldsymbol{X} W \bar{h} + \nu \boldsymbol{P}^H F F^T \boldsymbol{X}^H W^* \bar{h}^* + \boldsymbol{z} \tag{5}$$

where $\boldsymbol{r}$ is the received signal vector, F is the $N \times N$ DFT matrix, $\boldsymbol{P}$ is the PHN circulant matrix[4] which is constructed from cyclic shifts of the frequency-domain PHN row vector $\boldsymbol{p}$, $\boldsymbol{X}$ is the diagonal matrix of pilot symbols, $\bar{h}$ is the channel vector in the time-domain with memory L which is assumed to be less than the cyclic prefix length, and W is an $N \times L$ submatrix of the DFT matrix. The permutation matrix FF^T, in (5) selects the image sub-carrier which causes interference due to I/Q imbalance to fold to the desired sub-carrier. Using the Lemma in Section II, the ML estimate of the channel vector can be written as

$$\widehat{\bar{h}} = \frac{\mathcal{E}_{\boldsymbol{x}}^{-1}}{c} (\mu^* W^H \boldsymbol{X}^H \boldsymbol{P}^H \boldsymbol{r} - \nu W^H \boldsymbol{X}^* F F^T \boldsymbol{P}^* \boldsymbol{r}^*) \tag{6}$$

where $\mathcal{E}_{\boldsymbol{x}}$ is the energy of a pilot sub-carrier and $c = |\mu|^2 - |\nu|^2$. Note that to evaluate the ML channel estimate in (6), the receiver requires knowledge of μ, ν and $\boldsymbol{P}$.

B. PHN Vector Estimation

To estimate the PHN vector, the log likelihood function of the received signal vector $\boldsymbol{r}$ given μ, ν, $\boldsymbol{P}$ and $\bar{h}$ has to be minimized over $\boldsymbol{p}$. This log likelihood function is the exponent of the conditional PDF (i.e. $f(\boldsymbol{r}|\bar{h}, \mu, \nu, \boldsymbol{p}) \propto \exp(-\|\boldsymbol{r} - \mu \boldsymbol{P}\boldsymbol{X}W\bar{h} - \nu \boldsymbol{P}^H F F^T \boldsymbol{X}^H W^* \bar{h}^*\|^2/\sigma_z^2)$). Therefore, we have

$$\widehat{\boldsymbol{p}} = \arg\min_{\boldsymbol{p}} \|\boldsymbol{r} - \mu \boldsymbol{P}\boldsymbol{X}W\bar{h} - \nu \boldsymbol{P}^H F F^T \boldsymbol{X}^H W^* \bar{h}^*\|^2 \tag{7}$$

Substituting back the channel estimate from (6), the dependence of (7) on the channel is removed. After straightforward algebra, the expression inside the squared norm in (7) can be written as

$$\begin{aligned} &\boldsymbol{r} - \mu \boldsymbol{P}\boldsymbol{X}W\widehat{\bar{h}} - \nu \boldsymbol{P}^H F F^T \boldsymbol{X}^H W^* \widehat{\bar{h}}^* \\ &\quad = \frac{\mathcal{E}_{\boldsymbol{x}}^{-1}}{c} (|\mu|^2 \boldsymbol{A}_1 \boldsymbol{r} - |\nu|^2 \boldsymbol{A}_2 \boldsymbol{r} + \mu\nu \boldsymbol{A}_3 \boldsymbol{r}^*) \end{aligned} \tag{8}$$

where

$$\begin{aligned} \boldsymbol{A}_1 &= \boldsymbol{P}\boldsymbol{X}VV^H\boldsymbol{X}^H\boldsymbol{P}^H \\ \boldsymbol{A}_2 &= \boldsymbol{P}^H F F^T \boldsymbol{X}^H V^* V^T \boldsymbol{X}^T F^* F^H \boldsymbol{P} \\ \boldsymbol{A}_3 &= \boldsymbol{P}^H F F^T \boldsymbol{X}^H V^* V^T \boldsymbol{X}^T \boldsymbol{P}^T \\ &\quad - \boldsymbol{P}\boldsymbol{X}VV^H\boldsymbol{X}^H F F^T \boldsymbol{P}^* \end{aligned} \tag{9}$$

[2]PHN and I/Q imbalance are only present at the receiver side while negligible at the transmitter. This is a valid assumption in most downlink transmission scenarios.

[3]Timing and carrier frequency offset (CFO) synchronization is achieved, regardless of the channel knowledge, using the correlation properties of the repetitive training sequences prior to the full-pilot OFDM symbol [1], [23], [24].

[4]The PHN matrix $\boldsymbol{P}$ is an orthonormal matrix i.e. $\boldsymbol{P}\boldsymbol{P}^H = \boldsymbol{I}$ where $\boldsymbol{I}$ is the identity matrix. Moreover, the PHN matrix can be well approximated by a banded matrix (around its main diagonal) since PHN is a low-pass process.

where $F = [W|V]$ is the DFT matrix. The expression in (8) can be approximated using two assumptions. First, for small I/Q imbalance, $|\nu|^2$ is close to zero[5] and, therefore, the second term in the right hand side (RHS) of (8) is negligible compared to the first term since $\|\boldsymbol{A}_1\boldsymbol{r}\|^2 = \|\boldsymbol{A}_2\boldsymbol{r}\|^2$. Second, $\boldsymbol{A}_3$ is a close-to-zero matrix and, therefore, $\nu\boldsymbol{A}_3$ in the third term of the RHS of (8) is close to zero. To show that $\boldsymbol{A}_3$ is a near zero matrix, assuming $L \ll N$ we have $V \approx F$. Then $\boldsymbol{A}_3 \approx \mathcal{E}_{\boldsymbol{x}}(\boldsymbol{P}^H F F^T \boldsymbol{P}^T - \boldsymbol{P} F F^T \boldsymbol{P}^*) = \mathcal{E}_{\boldsymbol{x}}(F\bar{E}^H\bar{E}^T F^T - F\bar{E}\bar{E}^* F^T) = \boldsymbol{0}$. Therefore, for the PHN estimation we use the following approximation

$$\begin{aligned} \boldsymbol{r} - \mu\boldsymbol{P}\boldsymbol{X}W\widehat{\bar{h}} - \nu\boldsymbol{P}^H F F^T \boldsymbol{X}^H W^* \widehat{\bar{h}}^* \\ \approx \frac{\mathcal{E}_{\boldsymbol{x}}^{-1}|\mu|^2}{c}\boldsymbol{A}_1\boldsymbol{r} \end{aligned} \tag{10}$$

The log-likelihood function in (7) can then be written as

$$\begin{aligned} \widehat{\boldsymbol{p}} &\approx \arg\max_{\boldsymbol{p}} \boldsymbol{r}^H \boldsymbol{A}_1^H \boldsymbol{A}_1 \boldsymbol{r} \\ &= \arg\max_{\boldsymbol{p}} \boldsymbol{p}\boldsymbol{R}^H \boldsymbol{X} V V^H \boldsymbol{X}^H \boldsymbol{R}\boldsymbol{p}^H = \arg\max_{\boldsymbol{p}} \boldsymbol{p}\boldsymbol{M}\boldsymbol{p}^H \end{aligned} \tag{11}$$

where $\boldsymbol{R}$ is a circulant matrix constructed from $\boldsymbol{r}$. Using the approximate cost function in (11), the PHN vector $\boldsymbol{p}$ can be estimated *regardless* of μ, ν and $\bar{h}$. The quadratic form in (11) is maximized subject to the small PHN constraint i.e. $\boldsymbol{\delta} \times \Re(\boldsymbol{p})^T = 1$ where $\boldsymbol{\delta}$ is the $1 \times N$ unit row vector with first entry equal to one and zeros otherwise and $\Re(\boldsymbol{p})$ is the real part of $\boldsymbol{p}$. The solution to this optimization problem is given by

$$\Re(\boldsymbol{p})^T = \lambda\boldsymbol{\Gamma}^{-1}(\boldsymbol{\Lambda}^T\boldsymbol{S} + \boldsymbol{I})\boldsymbol{\delta}^T \quad ; \quad \Im(\boldsymbol{p})^T = \lambda\boldsymbol{S}\boldsymbol{\delta}^T \tag{12}$$

where $\boldsymbol{\Gamma}$ and $\boldsymbol{\Lambda}$ are the real and imaginary parts of the Hermitian matrix $\boldsymbol{M}$, $\Im(\boldsymbol{p})$ is the imaginary part of $\boldsymbol{p}$ and $\boldsymbol{S} = [\boldsymbol{I} + \left(\boldsymbol{\Gamma}^{-1}\boldsymbol{\Lambda}\right)^2]^{-1}\boldsymbol{\Gamma}^{-1}\boldsymbol{\Lambda}\boldsymbol{\Gamma}^{-1}$. The Lagrange multiplier can be found by satisfying the constraint which gives

$$\lambda = [\boldsymbol{\delta}(\boldsymbol{\Gamma}^{-1}\boldsymbol{\Lambda}^T\boldsymbol{S} + \boldsymbol{\Gamma}^{-1})\boldsymbol{\delta}^T]^{-1} \tag{13}$$

The complexity of the above estimator can be further reduced by considering only the most significant elements of $\boldsymbol{p}$ and setting the rest to zero. In particular, $\boldsymbol{p}$ can be well approximated by estimating its $(d+1)$ elements only, i.e., $\boldsymbol{p}(N - d/2), \cdots \boldsymbol{p}(0), \cdots \boldsymbol{p}(d/2)$ since PHN can be modeled as a low-pass random process [26] where d is the number of dominant PHN spectral components. Therefore, all the matrices in the estimator of (12) can be reduced in size accordingly.

C. I/Q Imbalance Parameter Estimation

Now that the receiver has the PHN estimate, $\widehat{\boldsymbol{P}}$, the log-likelihood metric in (7) is updated using the PHN matrix estimate and is differentiated with respect to $\alpha = \nu/\mu^*$. The cost function of (7) can be written after substituting from (8) as follows

$$\begin{aligned} \widehat{\alpha} &= \arg\min_{\alpha} \|\boldsymbol{A}_1\boldsymbol{r} - |\alpha|^2\boldsymbol{A}_2\boldsymbol{r} + \alpha\boldsymbol{A}_3\boldsymbol{r}^*\|^2 \\ &= \arg\min_{\alpha} \{J(\alpha)\} \end{aligned} \tag{14}$$

where $\boldsymbol{A}_1$, $\boldsymbol{A}_2$ and $\boldsymbol{A}_3$ are evaluated using $\widehat{\boldsymbol{P}}$. Differentiating (14) with respect to α^* produces the desired estimate of α. However, since $|\nu|^2$ is assumed to be close to zero, the third and the forth powers of α are ignored in the optimization expression of (14) yielding the following approximate cost function

$$\begin{aligned} J(\alpha) \approx \boldsymbol{r}^H\boldsymbol{A}_1^H\boldsymbol{A}_1\boldsymbol{r} + |\alpha|^2(\boldsymbol{r}^T\boldsymbol{A}_3^H\boldsymbol{A}_3\boldsymbol{r}^* - \boldsymbol{r}^H\boldsymbol{A}_1^H\boldsymbol{A}_2\boldsymbol{r} \\ - \boldsymbol{r}^H\boldsymbol{A}_2^H\boldsymbol{A}_1\boldsymbol{r}) + \alpha\boldsymbol{r}^H\boldsymbol{A}_1^H\boldsymbol{A}_3\boldsymbol{r}^* + \alpha^*\boldsymbol{r}^T\boldsymbol{A}_3^H\boldsymbol{A}_1\boldsymbol{r} \end{aligned} \tag{15}$$

Differentiating with respect to α^* and solving for α yields

$$\widehat{\alpha} = \frac{-\boldsymbol{r}^T\boldsymbol{A}_3^H\boldsymbol{A}_1\boldsymbol{r}}{\boldsymbol{r}^T\boldsymbol{A}_3^H\boldsymbol{A}_3\boldsymbol{r}^* - \boldsymbol{r}^H\boldsymbol{A}_1^H\boldsymbol{A}_2\boldsymbol{r} - \boldsymbol{r}^H\boldsymbol{A}_2^H\boldsymbol{A}_1\boldsymbol{r}} \tag{16}$$

Based on $\widehat{\alpha}$, we can compute the approximate values of μ and ν as follows

$$\begin{aligned} \widehat{\alpha} = \frac{\nu}{\mu^*} &= \frac{\epsilon\cos(\theta/2) - j\sin(\theta/2)}{\cos(\theta/2) - j\epsilon\sin(\theta/2)} \\ &\approx \epsilon\left(1 + \tan^2(\theta/2)\right) + j(\epsilon^2 - 1)\tan(\theta/2) \end{aligned} \tag{17}$$

The approximation in (17) follows from the fact that $\epsilon^2\sin^2(\theta/2) \rightarrow 0$ for small I/Q mismatches. Therefore, from the real and imaginary parts of (17) we have

$$\widehat{\epsilon} \approx \frac{\Re(\widehat{\alpha})}{1 + \Im(\widehat{\alpha})^2} \tag{18}$$

and the approximate value of θ can be found to be

$$\widehat{\sin(\theta/2)} = \frac{-\Im(\widehat{\alpha})}{\sqrt{1 + \Im(\widehat{\alpha})^2}} \tag{19}$$

From (18) and (19), μ and ν are estimated by direct substitution into (4). The channel estimate is computed by substitution of $\widehat{\mu}$ and $\widehat{\nu}$ and $\widehat{\boldsymbol{P}}$ into (6). The estimation algorithm is summarized as follows

Step 1. After receiving $\boldsymbol{r}$ corresponding to the full-pilot OFDM symbol, $\boldsymbol{p}$ is estimated using (12) independent of μ, ν and $\bar{h}$.

Step 2. The I/Q imbalance parameters μ and ν are estimated using (16), (18) and (19) independent of $\bar{h}$ and by using the estimated PHN from Step 1.

Step 3. The channel vector is estimated by substituting $\widehat{\boldsymbol{P}}$, $\widehat{\mu}$ and $\widehat{\nu}$ into (6).

D. Complexity Analysis

As multiplication dominates the overall implementation complexity compared to additions, the number of real multiplications of the proposed scheme is compared with (6), (12), and (16). PHN estimation in (6) requires $8(\frac{d}{2}+1)^3 + 3(\frac{d}{2}+1)^2$ multiplications since it involves matrix inversion of size $d/2+1$. I/Q imbalance estimation in (12) requires $32(\frac{d}{2}+1)^2N^2 + 32(\frac{d}{2}+1)N + 4$ multiplications since it requires matrix vector multiplications of size N. Channel estimation in (6) requires $8N(\frac{d}{2}+L+2) + 8L$ multiplications depending on the time-domain channel length L. The sum of all these multiplications is compared with [11] and [21] in Fig. 12, which shows significant reduction in implementation complexity especially for large N.

IV. SISO Data Detection

Data detection is performed assuming that the channel vector and the I/Q imbalance parameters are constant for the entire OFDM packet while the PHN process changes from one OFDM symbol to the next; therefore, the estimated channel vector and I/Q imbalance parameters in the full-pilot stage can be used for data detection, but the PHN matrix is re-estimated for each data OFDM symbol. The received data OFDM symbol in the qth time slot, $\bar{r}^q$, and its complex conjugate are given by

$$\begin{aligned} \bar{r}^q &= \mu\bar{E}^q\bar{H}\bar{x}^q + \nu\bar{E}^{qH}\bar{H}^*\bar{x}^{q*} + \bar{z}^q \\ \bar{r}^{q*} &= \mu^*\bar{E}^{qH}\bar{H}^*\bar{x}^{q*} + \nu^*\bar{E}^q\bar{H}\bar{x}^q + \bar{z}^{q*} \end{aligned} \tag{20}$$

[5]Even for unrealistically large I/Q imbalance, this approximation is still valid. As a numerical example, let $20\log_{10}(1+\epsilon) = 1$ dB and $\theta = \pi/40$ which corresponds to large I/Q imbalance, then we have $|\nu|^2 = \epsilon^2\cos^2(\theta/2) + \sin^2(\theta/2) = 0.021 \ll 1$.

The image signal can be removed by combining $\bar{r}^q$ and $\bar{r}^{q*}$

$$\bar{y}^q = \frac{1}{|\widehat{\mu}|^2 - |\widehat{\nu}|^2} \left(\widehat{\mu}^* \bar{r}^q - \widehat{\nu} \bar{r}^{q*}\right) = \bar{E}^q \bar{H} \bar{x}^q + \bar{\tilde{z}}^q \tag{21}$$

where $\bar{\tilde{z}}^q = \widehat{\mu}^* \bar{z}^q - \widehat{\nu} \bar{z}^{q*} + \bar{\zeta}$ and $\bar{\zeta}$ is due to the I/Q imbalance estimation error. Taking the DFT of $\bar{y}^q$ we have

$$\boldsymbol{y}^q = F \bar{y}^q = \boldsymbol{P}^q \boldsymbol{H} \boldsymbol{x}^q + \tilde{\boldsymbol{z}}^q \tag{22}$$

where $\boldsymbol{H}$ is the diagonal channel matrix in the frequency-domain. Since the channel estimate is available from the full-pilot stage, the equi-distant pilots in $\boldsymbol{x}^q$ are used to estimate the diagonal element of $\boldsymbol{P}^q$ at the pilot positions. The average of all these estimates is used as the CPE estimate of the qth data OFDM symbol i.e. $\overline{\phi^q}$. The generalized form of the CPE estimate will be presented when we extend our scheme to the MIMO case. Instead of detecting the data using $\overline{\phi^q}$ only as in the conventional receiver, the qth OFDM symbol is stored and the CPE of the next data OFDM symbol, $\overline{\phi^{q+1}}$, is computed and a linear interpolation between $\overline{\phi^{q-1}}$, $\overline{\phi^q}$ and $\overline{\phi^{q+1}}$ is performed. This simple operation helps reduce the ICI level and improves the BER performance compared to the conventional approach of using $\overline{\phi^q}$ only. After estimating the PHN matrix using this interpolation technique, the data symbols are detected as follows

$$\boldsymbol{x}^q(k) = \frac{1}{\boldsymbol{H}(k,k)} \sum_{i=-d/2}^{d/2} \widehat{\boldsymbol{p}}^{q*}(i) \boldsymbol{y}^q(k-i) \tag{23}$$

To keep the complexity low, no more than one OFDM symbol is stored for interpolation. Moreover, the convolution in (23) is performed over few PHN spectral components due to the "approximately" banded structure of $\boldsymbol{P}$.[6]

V. Extension to MIMO Scenarios

The extension to MIMO scenarios is carried out using the following three assumptions:

1. In the channel estimation stage, full-pilot OFDM symbols are transmitted from different transmit antennas. To minimize the mean squared error of the MIMO channel estimate [25], the full-pilot OFDM symbols are designed such that the transmitted pilots from different transmit antennas are orthogonal to each other. Moreover, the number of time slots over which the full-pilot OFDM symbols are transmitted should be greater than or equal to the number of transmit antennas to avoid having an under-determined system of equations.

2. In the data transmission stage, OFDM symbols are transmitted according to a predefined space-time encoding scheme.

3. The channel impulse response vectors are spatially independent, i.e. the elements of the MIMO channel matrix are independent samples of complex Gaussian random variables and there is no correlation neither between the transmit antennas nor the receive antennas [27].

A. Channel Estimation

Assume that the number of transmit and receive antennas are n_T and n_R, respectively, where $n_R \geq n_T$ and that the maximum number of channel taps across any of the paths between receive and transmit antennas is $L_{\max}$. Assume also that Q OFDM training symbols are transmitted in the full-pilot stage from the ith transmit antenna for $i = \{1, \cdots, n_T\}$ and are collected and processed at the mth receive antenna for $m = \{1, \cdots, n_R\}$. We define $\bar{h}_{im}$ as the time-domain channel vector from the ith transmit to the mth receive antenna which is considered to be constant during one packet of OFDM symbols. The corresponding signal model for the MIMO case can be written by generalizing (5) as follows

$$\begin{aligned} \boldsymbol{r}_m^q = F \bar{r}_m^q = \mu_m \boldsymbol{P}_m^q \sum_{i=1}^{n_T} \boldsymbol{X}_i^q W \bar{h}_{im} \\ + \nu_m \boldsymbol{P}_m^{qH} F F^T \sum_{i=1}^{n_T} \boldsymbol{X}_i^{qH} W^* \bar{h}_{im}^* + \boldsymbol{z}_m^q \end{aligned} \tag{24}$$

where $\boldsymbol{r}_m^q$ and $\boldsymbol{P}_m^q$ are the received vector and the PHN circulant matrix in the mth receive branch at the qth time slot in the frequency-domain, $\boldsymbol{X}_i^q$ is a diagonal matrix of pilot symbols transmitted from the ith transmit antenna at the qth time slot and μ_m and ν_m are the I/Q imbalance parameters for the mth receive branch, respectively. Given that Q is greater than or equal to n_T, the system of equations in (24) can be written in a compact matrix form which can be used to estimate $\bar{h}_{im}$ for $i = 1, \cdots, n_T$. In particular, we assume that Q is large enough so that the channels from the mth receive antenna to all transmit antennas can be estimated separately from the other receive branches. To write the system of equations for the mth receive antenna, we define $\boldsymbol{r}_m = [\boldsymbol{r}_m^{1T} \cdots \boldsymbol{r}_m^{QT}]^T$ which is a size $QN \times 1$ vector constructed by concatenating the received vectors over Q time slots and $\bar{h}_m = [\bar{h}_{1m}^T \cdots \bar{h}_{n_T m}^T]^T$ as a size $n_T L_{\max} \times 1$ channel vector in the time domain which is to be estimated. The input-output model can be expressed as follows

$$\boldsymbol{r}_m = \mu_m \boldsymbol{\mathcal{P}}_m \boldsymbol{\mathcal{X}} \boldsymbol{W} \bar{h}_m + \nu_m \boldsymbol{\mathcal{P}}_m^H \boldsymbol{\mathcal{J}} \boldsymbol{\mathcal{X}}^* \boldsymbol{W}^* \bar{h}_m^* + \boldsymbol{z}_m \tag{25}$$

where $\boldsymbol{\mathcal{P}}_m = \text{diag}\{[\boldsymbol{P}_m^1 \cdots \boldsymbol{P}_m^Q]\}$, $\boldsymbol{W} = \{\boldsymbol{I}_{n_T} \odot W\}$, $\boldsymbol{\mathcal{J}} = \{\boldsymbol{I}_Q \odot F F^T\}$ and

$$\boldsymbol{\mathcal{X}} = \begin{bmatrix} \boldsymbol{X}_1^1 & \cdots & \boldsymbol{X}_{n_T}^1 \\ \vdots & \ddots & \vdots \\ \boldsymbol{X}_1^Q & \cdots & \boldsymbol{X}_{n_T}^Q \end{bmatrix}_{QN \times n_T N} \tag{26}$$

where $\odot$ is the Kronecker product between two matrices and $\text{diag}\{\cdot\}$ is the operation which constructs a diagonal matrix from its vector input or a block diagonal matrix from its matrix input.

The frequency-domain signal model in (24) assumes that there is an independent oscillator for each receive branch. Although the PHN processes at the LO output are in practice correlated, considering their joint probability density function would increase the complexity of the estimation algorithm significantly and, therefore, they are estimated separately. There are n_R different sets of I/Q imbalance and PHN parameters to be estimated. Over the full-pilot OFDM symbol transmission stage, the total number of unknowns to be estimated is $n_T n_R L_{\max} + n_R QN + 2n_R$ complex variables. The total number of available equations, however, is just $n_R QN$ complex equations. Since this system of equations is under-determined, only few spectral components of the PHN process are estimated. Assuming d off-diagonal elements of $\boldsymbol{P}_m^q$ are estimated, then the following condition must be satisfied

$$n_T L_{\max} + Q(d+1) + 2 \leq NQ \tag{27}$$

This condition is satisfied for typical values of n_T and channel memory lengths since the DFT size N is large. Following the same procedure for channel estimation in the SISO case, the estimate of $\bar{h}_m$ is given by

$$\widehat{\bar{h}}_m = \frac{\mathcal{E}_{\boldsymbol{x}}^{-1}}{c_m} (\mu_m^* \boldsymbol{W}^H \boldsymbol{\mathcal{X}}^H \boldsymbol{\mathcal{P}}_m^H \boldsymbol{r}_m - \nu_m \boldsymbol{W}^H \boldsymbol{\mathcal{X}}^* \boldsymbol{\mathcal{J}} \boldsymbol{\mathcal{P}}_m^* \boldsymbol{r}_m^*) \tag{28}$$

[6] Since the frequency-domain PHN matrix is circulant, orthonormal and approximately banded, PHN compensation consists of a simple FIR matched-filtering operation by the same short FIR filter for all sub-carriers.

where $c_m = |\mu_m|^2 - |\nu_m|^2$. Substituting (25) into (28) we find that $\widehat{\tilde{h}}_m = \bar{h}_m + \tilde{\tilde{z}}_m$ if and only if the pilot matrix $\boldsymbol{\mathcal{X}}$ satisfies

$$\boldsymbol{\mathcal{X}}^H \boldsymbol{\mathcal{X}} = \mathcal{E}_{\boldsymbol{x}} \boldsymbol{I} \quad \text{or} \quad \sum_{q=1}^{Q} \boldsymbol{X}_i^{qH} \boldsymbol{X}_j^q = \begin{cases} \mathcal{E}_{\boldsymbol{x}} \boldsymbol{I}_N & i = j \\ \boldsymbol{0} & i \neq j \end{cases} \tag{29}$$

where $\tilde{\tilde{z}}_m$ is the equivalent noise vector given by

$$\tilde{\tilde{z}}_m = \frac{\mathcal{E}_{\boldsymbol{x}}^{-1}}{c_m} (\mu_m^* \boldsymbol{W}^H \boldsymbol{\mathcal{X}}^H \boldsymbol{\mathcal{P}}_m^H \boldsymbol{z}_m - \nu_m \boldsymbol{W}^H \boldsymbol{\mathcal{X}}^* \boldsymbol{\mathcal{J}} \boldsymbol{\mathcal{P}}_m^* \boldsymbol{z}_m^*) \tag{30}$$

Note that the condition in (29) produces the channel estimate with the minimum mean square error [25].

B. PHN and I/Q Imbalance Parameters Estimation

Using the same procedure given in Subsections III-B and III-C, we show that the ML estimator has the same quadratic form as in (11) as follows

$$\begin{aligned} \widehat{\boldsymbol{p}}_m &\approx \arg\max_{\boldsymbol{p}_m} \|\boldsymbol{\mathcal{P}}_m \boldsymbol{\mathcal{X}} \boldsymbol{V} \boldsymbol{V}^H \boldsymbol{\mathcal{X}}^H \boldsymbol{\mathcal{P}}_m^H \boldsymbol{r}_m\|^2 \\ &= \arg\max_{\boldsymbol{p}_m} \boldsymbol{p}_m \boldsymbol{\mathcal{M}}_m \boldsymbol{p}_m^H \quad : \quad m = 1, \cdots, n_R \end{aligned} \tag{31}$$

where $\boldsymbol{p}_m = [\boldsymbol{p}_m^1 \cdots \boldsymbol{p}_m^Q]$ and $\boldsymbol{p}_m^q$ is a row vector and $\boldsymbol{V} = \{\boldsymbol{I}_{n_T} \odot V\}_{n_T N \times n_T (N - L_{\max})}$. Estimating only $(d+1)$ elements of $\boldsymbol{p}_m^q$, the complexity of this estimation algorithm is reduced from computing the inverse of a $QN \times QN$ matrix to that of a $Q(d+1) \times Q(d+1)$ matrix. The constraint to solve (31) is the same as the constraint in (11). Now that the receiver has the estimates for every PHN vector over Q time slots for each receive antenna, the I/Q imbalance parameters for the mth receive branch can be computed. Following the derivations in (14)-(19) for the SISO case and defining $\alpha_m = \nu_m / \mu_m^*$, the approximate ML estimate of α_m is obtained as follows

$$\widehat{\alpha}_m = \frac{-\boldsymbol{r}_m^T \boldsymbol{A}_{m3}^H \boldsymbol{A}_{m1} \boldsymbol{r}_m}{\boldsymbol{r}_m^T \boldsymbol{A}_{m3}^H \boldsymbol{A}_{m3} \boldsymbol{r}_m^* - \boldsymbol{r}_m^H \boldsymbol{A}_{m1}^H \boldsymbol{A}_{m2} \boldsymbol{r}_m - \boldsymbol{r}_m^H \boldsymbol{A}_{m2}^H \boldsymbol{A}_{m1} \boldsymbol{r}_m} \tag{32}$$

where $\boldsymbol{A}_{mi}$ for $m = 1, \cdots, n_R$ are the same as the expressions in (9) evaluated for the mth receive branch PHN process. The I/Q imbalance parameters are computed using the same procedure in (17)-(19) and then used in (28) to produce the channel estimate.

C. Data Detection Stage

During the data transmission stage, equally-spaced pilots are available within each data OFDM symbol. The data OFDM symbols, $\boldsymbol{s}_i^q$, are generated in the frequency-domain and are mapped into a corresponding space-time code according to a predefined rule. Depending on the transmission rate, a generator matrix is derived to construct a linear combination of $\boldsymbol{s}_i^q$ and encodes it into $\boldsymbol{x}_i^q$ such that $\boldsymbol{x} = \boldsymbol{G}\boldsymbol{s}$ where $\boldsymbol{s} = [\boldsymbol{s}_1^{1T} \cdots \boldsymbol{s}_1^{QT} \cdots \boldsymbol{s}_{n_T}^{1T} \cdots \boldsymbol{s}_{n_T}^{QT}]^T$, $\boldsymbol{x} = [\boldsymbol{x}_1^{1T} \cdots \boldsymbol{x}_1^{QT} \cdots \boldsymbol{x}_{n_T}^{1T} \cdots \boldsymbol{x}_{n_T}^{QT}]^T$. The generator matrices used in this paper are for the Alamouti code [28], spatial multiplexing (SM) [29] and the Golden code [30].

Applying the processing operations in (20)-(21) to remove the effect of I/Q imbalance for every received vector $\bar{r}_m^q$ yields

$$\bar{y}_m^q = \frac{\widehat{\mu}_m^* \bar{r}_m^q - \widehat{\nu}_m \bar{r}_m^{q*}}{|\widehat{\mu}_m|^2 - |\widehat{\nu}_m|^2} = \bar{E}_m^q \sum_{i=1}^{n_T} \bar{H}_{im} \bar{x}_i^q + \tilde{\tilde{z}}_m^q \tag{33}$$

where $\bar{H}_{im}$ is the time-domain circulant channel matrix from the ith transmit to the mth receive antenna. Transforming to the frequency-domain, (33) can be written as

$$\boldsymbol{y}_m^q = F \bar{y}_m^q = \boldsymbol{P}_m^q \sum_{i=1}^{n_T} \boldsymbol{H}_{im} \boldsymbol{x}_i^q + \tilde{\boldsymbol{z}}_m^q \tag{34}$$

Collecting the received vectors for $m = 1, \cdots, n_R$ and $q = 1, \cdots, Q$ and stacking them in a vector, (34) can be written as

$$\boldsymbol{y} = \boldsymbol{\mathcal{P}} \boldsymbol{\mathcal{H}} \boldsymbol{x} + \tilde{\boldsymbol{z}} = \boldsymbol{\mathcal{P}} \boldsymbol{\mathcal{H}} \boldsymbol{G} \boldsymbol{s} + \tilde{\boldsymbol{z}} \tag{35}$$

where $\boldsymbol{\mathcal{H}}$ is the frequency-domain MIMO channel and

$$\begin{aligned} \boldsymbol{y} &= [\boldsymbol{y}_1^{1T}, \cdots, \boldsymbol{y}_1^{QT} \cdots \boldsymbol{y}_{n_R}^{1T}, \cdots, \boldsymbol{y}_{n_R}^{QT}]^T \\ \boldsymbol{\mathcal{P}} &= \text{diag}\{\boldsymbol{P}_1^1 \cdots \boldsymbol{P}_1^Q \cdots, \boldsymbol{P}_{n_R}^1 \cdots \boldsymbol{P}_{n_R}^Q\} \end{aligned} \tag{36}$$

The equally-spaced pilots in every $\boldsymbol{x}_i^q$ are used to estimate the CPE term of $\boldsymbol{P}_m^q$. Then, linear interpolation as in [31] is employed (separately for each receive branch) to connect these CPEs over all transmission time slots (i.e. $q = 1, \cdots, Q$).

Using (34), the kth frequency-domain sample of the qth OFDM symbol at the mth receive antenna is given by

$$\boldsymbol{y}_m^q(k) = \boldsymbol{p}_m^q(0) \sum_{i=1}^{n_T} \boldsymbol{x}_i^q(k) \boldsymbol{H}_{im}(k,k) + \text{ICI}_m^q(k) + \tilde{\boldsymbol{z}}_m^q(k) \tag{37}$$

The CPE estimate of the PHN process in the qth OFDM symbol at the mth receive branch is computed by averaging over all pilot sub-carriers as follows

$$\begin{aligned} \overline{\phi_m^q} &= \text{angle}\{\boldsymbol{p}_m^q(0)\} \\ &= \text{angle}\left\{ \sum_{k \in \text{pilots}} \sum_{i=1}^{n_T} \boldsymbol{y}_m^q(k) \boldsymbol{H}_{im}^*(k,k) \boldsymbol{x}_i^{q*}(k) \right\} \end{aligned} \tag{38}$$

After estimating the CPE term, $\overline{\phi_m^q}$ for $q = 1, \cdots, Q$ are assigned as the PHN estimates for the middle samples of their corresponding OFDM symbols. These points are then connected using lines which determine the estimate of the time-domain PHN process over all the intermediate OFDM samples between consecutive OFDM symbols. After PHN estimation and interpolation, the effect of $\boldsymbol{\mathcal{P}}$ is compensated for by multiplying $\boldsymbol{y}$ by $\widehat{\boldsymbol{\mathcal{P}}}^H$ as follows (see (35))

$$\boldsymbol{u} \triangleq \widehat{\boldsymbol{\mathcal{P}}}^H \boldsymbol{y} = \boldsymbol{\Omega} \boldsymbol{\mathcal{H}} \boldsymbol{G} \boldsymbol{s} + \widehat{\boldsymbol{\mathcal{P}}}^H \tilde{\boldsymbol{z}} \tag{39}$$

where $\boldsymbol{\Omega} \triangleq \widehat{\boldsymbol{\mathcal{P}}}^H \boldsymbol{\mathcal{P}} = \boldsymbol{I}_{n_R QN} + \boldsymbol{\Omega}_R$ and $\boldsymbol{\Omega}_R$ is the residual PHN after multiplying the received signal vector with $\widehat{\boldsymbol{\mathcal{P}}}^H$. The data information symbols are detected in every sub-carrier using the minimum mean square error (MMSE), zero-forcing (ZF) or ML receiver on the equivalent channel $\boldsymbol{\mathcal{H}}\boldsymbol{G}$, in each sub-carrier. Note that the achievable BER performance depends on the choice of $\boldsymbol{G}$ at the transmitter and the equalizer type at the receiver.

D. Residual Inter-carrier Interference and Noise (ICIN) Variance

The performance of our proposed scheme is analyzed in this subsection based on the ICIN variance. Assuming that the I/Q imbalance and PHN are present in the system and are ignored in the data detection stage, the noise and interference term in the mth receive antenna over the qth time slot can be written as

$$\text{ICIN}_m^q = (\mu_m \boldsymbol{P}_m^q - \boldsymbol{I}_N) \boldsymbol{x}_h + \nu_m \boldsymbol{P}_m^{qH} F F^T \boldsymbol{x}_h^* + \boldsymbol{z}_m^q \tag{40}$$

where $\boldsymbol{x}_h = \sum_{i=1}^{n_T} \boldsymbol{X}_i^q \boldsymbol{h}_{im}$ and $\boldsymbol{h}_{im}$ is the frequency-domain channel vector from the ith transmit to the mth receive antenna. Averaging over n_R receive antenna and Q time slots, the variance of the noise and the interference term in (40) can be computed as (see Appendix II for details)

$$\begin{aligned} \sigma_{\text{uncomp}}^2 \approx 2\mathcal{E}_{\boldsymbol{x}} n_T \sigma_{\boldsymbol{h}}^2 [1 - [1 - N\pi\beta T_s] \\ \times \cos(\theta_{\max}/2) + \sigma_\epsilon^2] + \sigma_{\boldsymbol{z}}^2 \end{aligned} \tag{41}$$

where $\mathsf{E}|\boldsymbol{h}_{im}\boldsymbol{h}_{im}^H| = \sigma_{\boldsymbol{h}}^2\boldsymbol{I}_N$, $\theta_{\max}$ is the maximum phase imbalance and σ_ϵ^2 is the variance of the amplitude imbalance averaged over all receive branches. If our proposed estimation and compensation method is used, the noise and interference terms can be written from (39) as

$$\text{ICIN}_{\text{proposed}} = \boldsymbol{\Omega}_R\boldsymbol{\mathcal{H}}\boldsymbol{G}\boldsymbol{s} + \widehat{\boldsymbol{\mathcal{P}}}^H\tilde{\boldsymbol{z}} \tag{42}$$

The variance of the expression in (42) is approximated by (see Appendix II for details)

$$\sigma^2_{\text{proposed}} \approx \frac{2}{3}\pi\mathcal{E}_{\boldsymbol{x}}n_T\sigma_{\boldsymbol{h}}^2\beta(N+N_c)T_s + \sigma_{\boldsymbol{z}}^2\sigma_\epsilon^2 + \sigma_{\boldsymbol{z}}^2 \tag{43}$$

where N_c is the cyclic prefix length. At high SNR, if $N_c \ll N$ and $\sigma_\epsilon^2 \to 0$, then $\sigma^2_{\text{proposed}} \approx \frac{1}{3}\sigma^2_{\text{uncomp}}$.

VI. Simulation Results

In this section, the performance of our proposed joint estimation and compensation algorithm is studied for both SISO and MIMO scenarios. The channel impulse response taps are generated as complex Gaussian random variables with variance $\sum_{n=1}^{L_{\max}}|\bar{h}_{im}(n)|^2 = \sigma_{\boldsymbol{h}}^2$ with an exponential power delay profile. The channel memory is assumed to be less than the cyclic prefix length (i.e. $L_{\max} < N_c$) which is 16. The DFT size is 64, the OFDM signal bandwidth is 20MHz and the number of equally-spaced pilots in the data OFDM symbols is 4. These parameters are adopted from the IEEE 802.11a/n standards [1].

Fig. 1 shows the channel MSE versus SNR in a SISO configuration. When the receiver suffers from PHN only, the MSE of our proposed scheme is very close to the scheme in [11]. In the presence of both I/Q imbalance and PHN, the amount of performance loss when using our proposed scheme is much less than that of [11]. Note that our proposed scheme starts to work reliably after SNR of around 7 dB. This is the result of 180° phase ambiguity in the estimated channel at low SNR due to unreliable estimates of I/Q imbalance parameters. This phenomenon is well studied in estimation theory as the threshold effect [32]. The phase mismatch θ is modeled as a uniformly-distributed random variable over the range $\pm\pi/40$ radians and the maximum amplitude mismatch is $20\log_{10}(1+\epsilon) = 1$ dB. The 3 dB bandwidth of the PHN is $\beta = 1$kHz [7].

Fig. 2 shows the un-coded BER performance for the SISO case with 16-QAM and 64-QAM signal constellations. Significant performance improvement is observed when applying our proposed joint compensation scheme compared to the uncompensated case. The number of off-diagonal elements of the PHN matrix, d, is chosen to be 2 in this simulation. The coded BER performance for the SISO scenario is shown in Fig. 3. The coding rate is $1/2$ and the WLAN standard convolutional encoder $[133, 171]$ with constraint length of 7 is used. The Viterbi decoder uses the soft information at the output of the detector in (23) with a decoding depth equal to five times the convolutional encoder constraint length.

The un-coded BER performance of our algorithm is compared to that of [22] and [21] in Fig. 4 for 16- and 64-QAM signal constellations. As it can be seen from the figure, since [22] ignores PHN in the full-pilot stage and estimates CPE only during the data detection stage, it suffers significant performance loss. On the other hand, since [21] uses hard-decision data for iterations, loss of information is incurred and the algorithm suffers from error propagation. The coded BER performance of our scheme is also compared with [22] and [21] in Fig. 5. From the figure, after one iteration, the scheme proposed in [21] achieves similar performance to our scheme but the receiver of [21] has to wait until the data at the output of the Viterbi decoder is ready for feedback. This increases the system latency, which may not be acceptable in many packet-based systems.

Fig. 6 depicts the channel estimation MSE for the 2×2 MIMO case. During the full-pilot OFDM transmission stage, the pilots are chosen from a BPSK constellation and are designed such that they satisfy (29). As it can be seen, our proposed scheme performs close to the case of perfect I/Q parameter knowledge and estimated PHN after the threshold SNR. Similar to Fig. 1, the channel MSE of the scheme in [11] is presented. As it can be seen from the figure, this scheme has a significant loss in the presence of I/Q imbalance.

Fig. 7 shows the un-coded BER performance for the 2×2 system. The transmission rate is set to be 4 bits per channel use (bpcu) for the Alamouti and the Golden codes. The detector for the Golden code is linear MMSE to keep the decoding complexity comparable to the Alamouti code. The linear MMSE compensation scheme for the Golden code does not mitigate the RF impairments satisfactorily since it is highly sub-optimal (compared to the ML detector) for the Golden code. However, application of the proposed linear compensation algorithm to the Alamouti code improves the BER performance significantly. Moreover, as it can be seen from the figure, using [11] for estimation and compensation in the Alamouti case is not as effective as our proposed scheme since [11] ignores I/Q imbalance.

The BER performance of the SM and Golden codes with ML decoding and 4-QAM signal constellation are compared in Fig. 8. It can be seen that our proposed joint channel, I/Q and PHN estimator becomes very effective when combined with the ML detector for both SM and the Golden code. However, [11] is less effective than our proposed scheme in reducing the ICI and therefore, suffers from a large error floor.

The coded BER performance of a 4×4 system is depicted in Fig. 9 and Fig. 10 for different values of β. Following IEEE 802.11n, two parallel Alamouti codes (AA) or one Alamouti code and one SM (AS) are transmitted in two time slots which are shown in Fig. 9 and Fig. 10, respectively. Also shown in these figures, is the BER performance of the AA and AS codes when the compensation scheme in [21] is used with one iteration. As it can be seen from these figures, using the scheme in [21], more than one iteration is required to further reduce the error floor which corresponds to more latency and higher complexity.

The variances of the ICI and noise terms derived in Subsection V-D for the compensated and uncompensated cases in the MIMO scenario are plotted in Fig.11. It can be seen from the figure that for different values of β, $\sigma^2_{\text{proposed}} < \sigma^2_{\text{uncomp}}$ in the high SNR regime. Therefore, our proposed algorithm reduces the ICI term due to PHN and I/Q imbalance significantly compared to the uncompensated case.

The complexity of the proposed estimation scheme is presented in Fig. 12 in terms of the total number of real multiplication as a function of the DFT size N. In this figure, $d = 2$, $L = 4$ and the number iterations in [21] is $i = 1$. As it can be seen from the figure, our scheme has much lower computational complexity especially for large N.

VII. Conclusion

We proposed a novel joint ML I/Q imbalance, PHN and channel estimator for OFDM systems. We reduced the complexity of our proposed estimator by approximating its cost functions under small PHN and I/Q imbalance assumptions. In addition, we generalized our estimator to the MIMO scenario and investigated its un-coded and coded BER performance for different MIMO transmission schemes. Furthermore, we derived approximate expressions for the variance of the residual ICI and noise with

and without compensation. Our proposed compensation scheme is non-iterative resulting in smaller decoding delays compared to iterative receivers, yet achieves significant performance gains over uncompensated schemes at practical complexity levels.

APPENDIX I
PROOF OF (2)

The PDF of the random vector $\bar{r}$ given μ, ν, $\bar{D}_1$ and $\bar{D}_2$ according to (1) can be written as

$$p(\bar{r}|\mu,\nu,\bar{D}_1,\bar{D}_2) \propto \exp[-(\bar{r}-\mu\bar{D}_1\bar{x}-\nu\bar{D}_2\bar{x}^*)^H K_z^{-1} \times (\bar{r}-\mu\bar{D}_1\bar{x}-\nu\bar{D}_2\bar{x}^*)] \quad (44)$$

Defining the log-likelihood function for $\bar{x}$, we have

$$\mathcal{L}(\bar{x}) = -\log\left(p(\bar{r}|\mu,\nu,\bar{D}_1,\bar{D}_2)\right) \quad (45)$$
$$= (\bar{r}-\mu\bar{D}_1\bar{x}-\nu\bar{D}_2\bar{x}^*)^H K_z^{-1}(\bar{r}-\mu\bar{D}_1\bar{x}-\nu\bar{D}_2\bar{x}^*)$$

The ML estimator of $\bar{x}$ can be written as

$$\widehat{\bar{x}} = \arg\min_{\bar{x}}\{\mathcal{L}(\bar{x})\} = \arg\min_{\bar{x}}\{(\bar{r}-\mu\bar{D}_1\bar{x}-\nu\bar{D}_2\bar{x}^*)^H \times K_z^{-1}(\bar{r}-\mu\bar{D}_1\bar{x}-\nu\bar{D}_2\bar{x}^*)\} \quad (46)$$

For the special case of uncorrelated noise i.e. $K_z = \sigma_z^2\boldsymbol{I}$, differentiating (46) with respect to $\bar{x}$ and $\bar{x}^*$ and setting the results to zero, the system of equations can be arranged as

$$\begin{bmatrix} \bar{\Upsilon} & \bar{\Sigma} \\ \bar{\Sigma}^* & \bar{\Upsilon}^* \end{bmatrix}\begin{bmatrix} \bar{x} \\ \bar{x}^* \end{bmatrix} = \begin{bmatrix} \mu^*\bar{D}_1^H & \nu\bar{D}_2^T \\ \nu^*\bar{D}_2^H & \mu\bar{D}_1^T \end{bmatrix}\begin{bmatrix} \bar{r} \\ \bar{r}^* \end{bmatrix} \quad (47)$$

where $\bar{\Upsilon} = |\mu|^2\bar{D}_1^H\bar{D}_1 + |\nu|^2\bar{D}_2^T\bar{D}_2^*$ and $\bar{\Sigma} = \mu^*\nu\left(\bar{D}_1^H\bar{D}_2 + \bar{D}_2^T\bar{D}_1^*\right)$. Therefore, the estimate of $\bar{x}$ is

$$\hat{\bar{x}} = (\bar{\Upsilon}\bar{\Upsilon}^* - \bar{\Sigma}\bar{\Sigma}^*)^{-1} \times (\bar{\Upsilon}^*\mu^*\bar{D}_1^H\bar{r} - \bar{\Sigma}\nu^*\bar{D}_2^H\bar{r} + \bar{\Upsilon}^*\nu\bar{D}_2^T\bar{r}^* - \bar{\Sigma}\mu\bar{D}_1^T\bar{r}^*) \quad (48)$$

Substituting for $\bar{\Upsilon}$ and $\bar{\Sigma}$ into (48) and simplifying (2) is derived.

APPENDIX II
DERIVATION OF (41) AND (43)

A. Derivation of (41)

Taking the trace of the expected value of (40) multiplied by its complex conjugate we get

$$\sigma^2_{\text{uncomp}} = \frac{1}{N}\text{Tr}(\mathsf{E}[\text{ICIN}_m^{qH} \times \text{ICIN}_m^q]) \approx \frac{\mathcal{E}_{\boldsymbol{x}} n_T \sigma^2_{\boldsymbol{h}}}{N} \times \quad (49)$$

$$(\text{Tr}(\mathsf{E}[(\mu_m\boldsymbol{P}_m^q - \boldsymbol{I}_N)^H(\mu_m\boldsymbol{P}_m^q - \boldsymbol{I}_N)]) + N|\nu_m|^2) + \sigma^2_{\boldsymbol{z}}$$
$$= \mathcal{E}_{\boldsymbol{x}} n_T\sigma^2_{\boldsymbol{h}}[1 + |\mu_m|^2 + |\nu_m|^2 - \frac{2}{N}\text{Tr}(\mathsf{E}\Re\{\mu_m\boldsymbol{P}_m^q\})] + \sigma^2_{\boldsymbol{z}}$$

The approximation in (49) follows from the assumption that the I/Q imbalance parameters are uncorrelated i.e. $\mathsf{E}[\mu_m\nu_m^*] = 0$. Since $|\mu_m|^2 + |\nu_m|^2 = 1 + |\epsilon_m|^2$ (see (4)) and by averaging over all receive antennas and all transmission time slots we get $\sigma^2_{\text{uncomp}} \approx \mathcal{E}_{\boldsymbol{x}} n_T\sigma^2_{\boldsymbol{h}}[2 - \frac{2}{Nn_R}\sum_{m=1}^{n_R}\text{Tr}(\mathsf{E}\Re\{\mu_m\boldsymbol{P}_m^q\}) + \sigma_\epsilon^2] + \sigma^2_{\boldsymbol{z}}$ We can approximate the trace operation as

$$\text{Tr}(\mathsf{E}\Re\{\mu_m\boldsymbol{P}_m^q\}) \approx \cos(\theta_{\max}/2)\sum_{n=0}^{N-1}\mathsf{E}\Re\{e^{j\phi_m^q(n)}\} = N\cos(\theta_{\max}/2)[1-(N-1)\pi\beta T_s] \quad (50)$$

The approximation in (50) follows from the fact that $\Re\{\mu_m\boldsymbol{P}_m^q\} \approx \Re\{\mu_m\}\Re\{\boldsymbol{P}_m^q\}$. Substituting (50) into the expression for σ^2_{uncomp}, results in (41).

B. Derivation of (43)

Following the same steps to derive (41) and assuming that $|\nu_m|^2$ is close to zero, we can write

$$\sigma^2_{\text{proposed}} = \frac{1}{n_R NQ}\text{Tr}(\mathsf{E}[(\Xi)^H(\Xi)]) \quad (51)$$
$$\stackrel{a}{=} \frac{\mathcal{E}_{\boldsymbol{x}} n_T}{n_R NQ}\text{Tr}(\mathsf{E}[\boldsymbol{\mathcal{H}}\boldsymbol{G}\boldsymbol{G}^H\boldsymbol{\mathcal{H}}^H]\mathsf{E}[\boldsymbol{\Omega}_R^H\boldsymbol{\Omega}_R]) + \sigma^2_{\boldsymbol{z}}(1+\sigma^2_\epsilon)$$
$$\stackrel{b}{=} \frac{\mathcal{E}_{\boldsymbol{x}} n_T\sigma^2_{\boldsymbol{h}}}{n_R NQ}\text{Tr}(\mathsf{E}[\boldsymbol{\Omega}_R^H\boldsymbol{\Omega}_R]) + \sigma^2_{\boldsymbol{z}}(1+\sigma^2_\epsilon)$$

where $\Xi = \boldsymbol{\Omega}_R\boldsymbol{\mathcal{H}}\boldsymbol{G}\boldsymbol{s} + \widehat{\boldsymbol{\mathcal{P}}}^H\tilde{\boldsymbol{z}}$. Note that (a) holds because the channel and PHN vectors are independent and (b) follows from the fact that $\mathsf{E}[\boldsymbol{\mathcal{H}}\boldsymbol{G}\boldsymbol{G}^H\boldsymbol{\mathcal{H}}^H]$ is diagonal because the channel taps are assumed spatially uncorrelated and the channel is time-invariant over each packet. The trace of the residual PHN in (51) can be computed as follows

$$\text{Tr}(\mathsf{E}[\boldsymbol{\Omega}_R^H\boldsymbol{\Omega}_R]) = n_R Q\text{Tr}(\mathsf{E}|\boldsymbol{p}_m^q - \widehat{\boldsymbol{p}}_m^q|^2) = n_R Q\text{Tr}(\mathsf{E}|\bar{e}_m^q - \bar{\mathcal{I}}\boldsymbol{u}|^2) \quad (52)$$

where $\bar{e}_m^q = [e^{j\phi_m^q(0)}\cdots e^{j\phi_m^q(N+N_c-1)}]^T$ is the time-domain PHN vector, $\bar{\mathcal{I}}$ is the linear interpolation matrix of length $N+N_c$ and $\boldsymbol{u}$ is a vector containing CPE estimates. Assuming that the CPE estimation error is negligible compared to the interpolation error, we can simplify (52) as

$$\text{Tr}(\mathsf{E}[\boldsymbol{\Omega}_R^H\boldsymbol{\Omega}_R]) = n_R Q\text{Tr}(\mathsf{E}|\bar{e}_m^q - \bar{\mathcal{I}}\boldsymbol{u}_g|^2) + n_R Q\text{Tr}(\mathsf{E}|\bar{\mathcal{I}}\boldsymbol{u}_g - \bar{\mathcal{I}}\boldsymbol{u}|^2) \approx n_R Q\text{Tr}(\mathsf{E}|\bar{e}_m^q - \bar{\mathcal{I}}\boldsymbol{u}_g|^2) = \frac{2}{3}n_R Q(N+N_c)\pi\beta T_s \quad (53)$$

where $\boldsymbol{u}_g$ is the vector containing the exact CPE values. The derivation in (53) is carried out following the results in [7]. Substituting (53) into (51) results in (43).

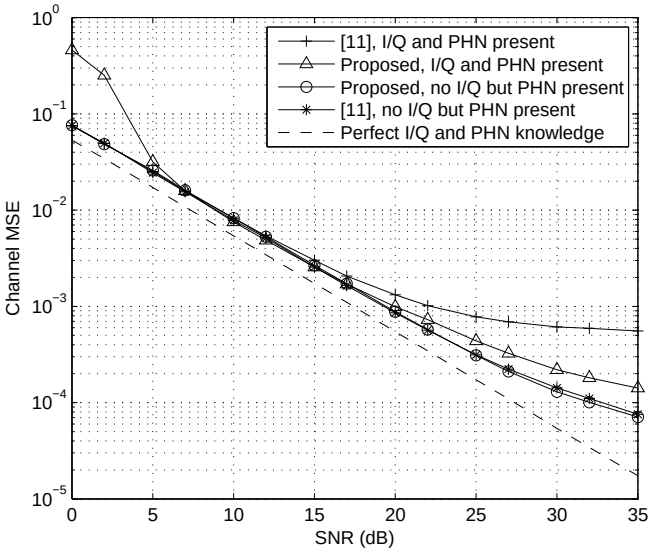

Fig. 1. SISO channel estimation MSE with $\beta = 1$kHz for PHN process. The dashed line is the channel MSE in the ideal case where there is no I/Q imbalance and PHN in the system. Note that our scheme performs much better than [11] after the threshold SNR when both I/Q imbalance and PHN are present

REFERENCES

[1] "Supplement to IEEE standard for information technology-telecommunications and information exchange between systems-local and metropolitan area networks-specific requirements-part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: High-speed physical layer in the 2.4 GHz band," *IEEE*, April 1999.

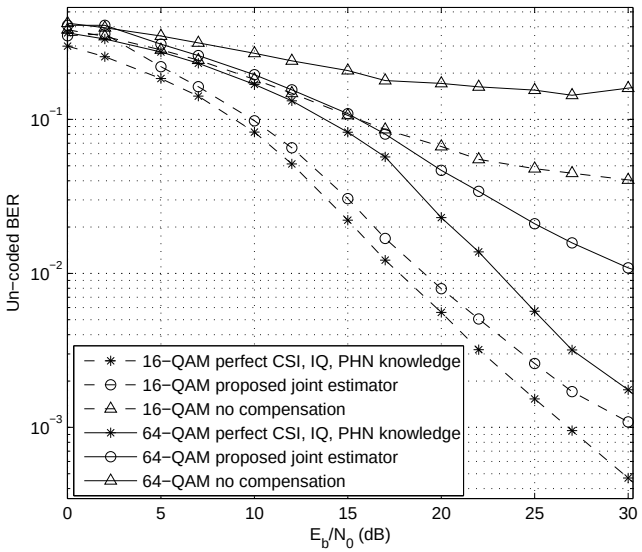

Fig. 2. Un-coded BER performance for the SISO scenario with $\beta = 1$kHz. This figure illustrates the effectiveness of our proposed estimation and compensation technique compared to the uncompensated case

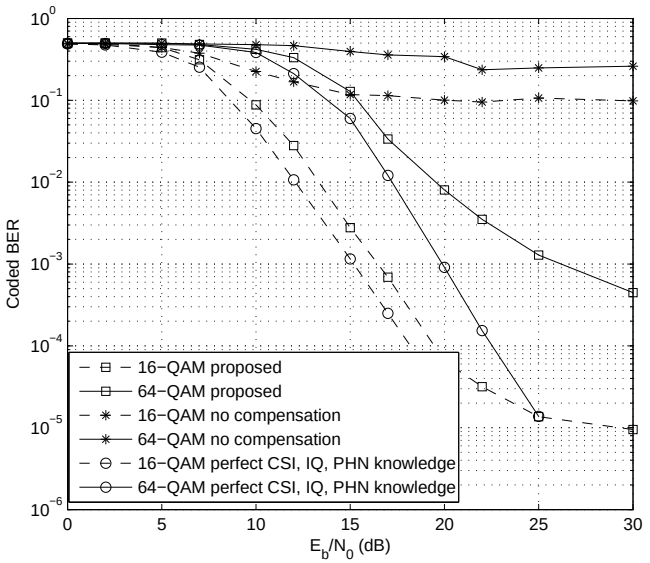

Fig. 3. Coded BER performance for the SISO case. The code rate is $1/2$ with the constraint length 7 convolutional encoder $[133, 171]$. Our compensation scheme results in significant performance improvements compared to the uncompensated case for both 16-QAM and 64-QAM. $\beta = 1$kHz

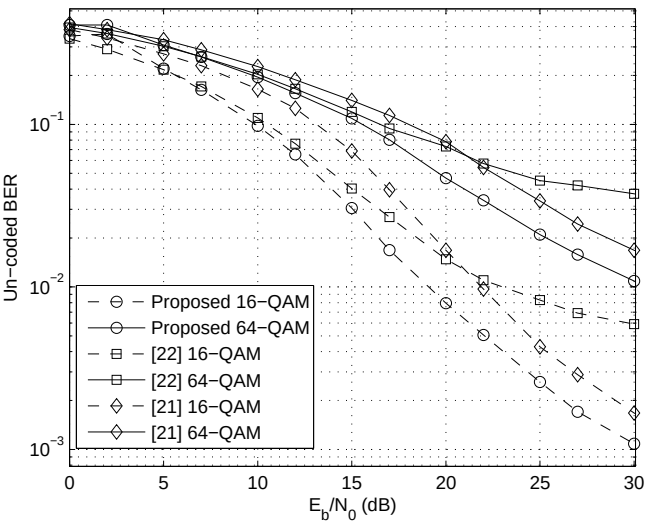

Fig. 4. Un-coded BER performance comparison for the SISO scenario with $\beta = 1$kHz. The number of iterations is one for the scheme in [21]

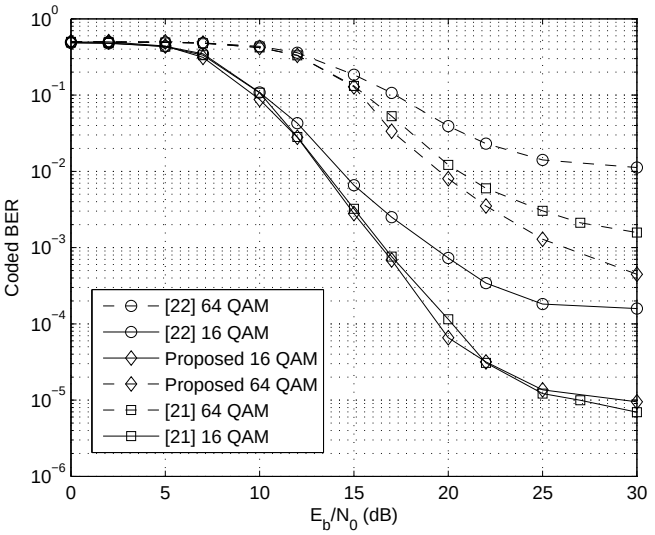

Fig. 5. Coded BER performance comparison for the SISO case. The code rate is $1/2$ with the constraint length 7 convolutional encoder $[133, 171]$. $\beta = 1$kHz

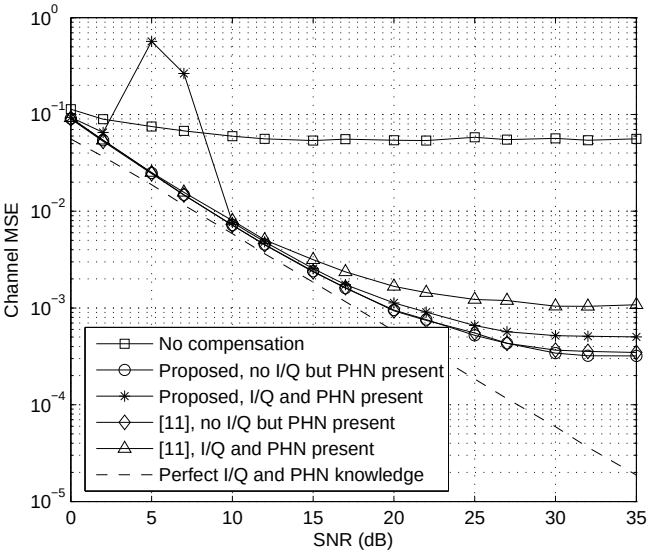

Fig. 6. Channel estimation MSE in the 2×2 MIMO scenario. The MSE spike in the proposed algorithm is the threshold effect in our estimator. The presented MSE result is averaged over all the paths from the ith transmit to the mth receive antennas. $\beta = 1$kHz

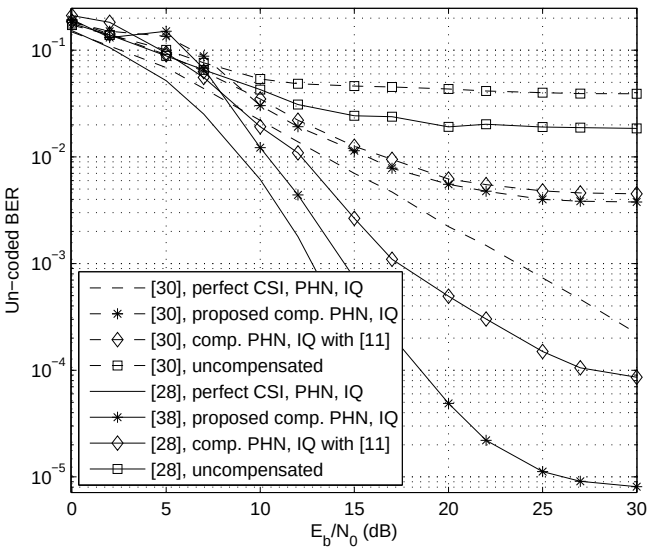

Fig. 7. Un-coded BER performance in the 2×2 MIMO scenario. The transmission rate for both cases is 4 bpcu. Linear MMSE detection is used to decode the Golden code and simple matched filtering is used to decode the Alamouti code. $\beta = 2$kHz.

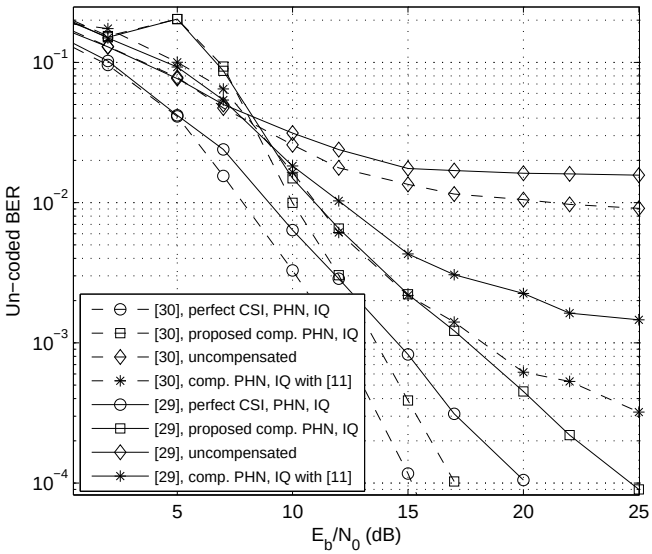

Fig. 8. Comparison of un-coded BER performance in the 2×2 MIMO case for SM and the Golden code. Maximum likelihood is used for decoding. Constellation size is 4-QAM for both cases. $\beta = 2$kHz.

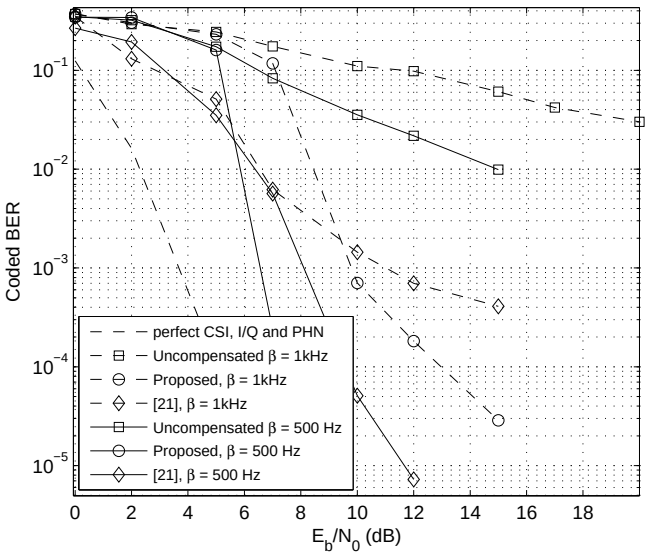

Fig. 9. Coded BER performance of a 4×4 AA system with $\beta = 1$kHz and 500Hz. Note that a spectral efficiency of 8 bpcu which corresponds to 16 QAM for all information symbols in the AA scheme is used

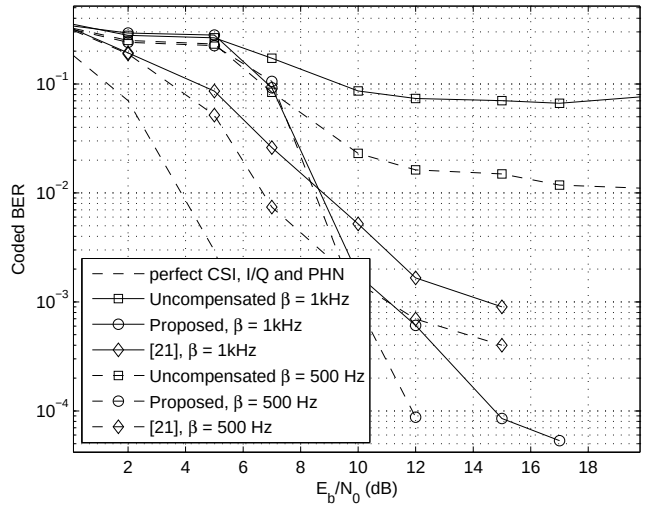

Fig. 10. Coded BER performance of a 4×4 AS system with $\beta = 1$kHz and 500Hz. Note that a spectral efficiency of 8 bpcu corresponds to 16 QAM for Alamouti code and 4 QAM for SM in the AS case, respectively

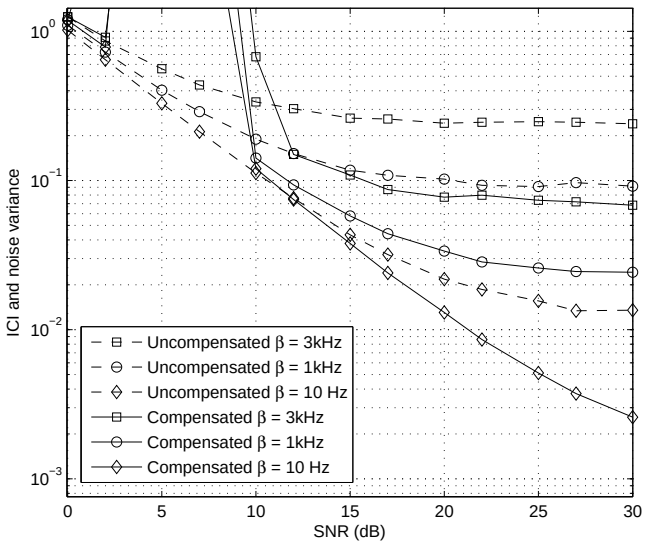

Fig. 11. ICI and noise variance comparison for compensated and uncompensated cases in 2×2 MIMO scenario for different values of β

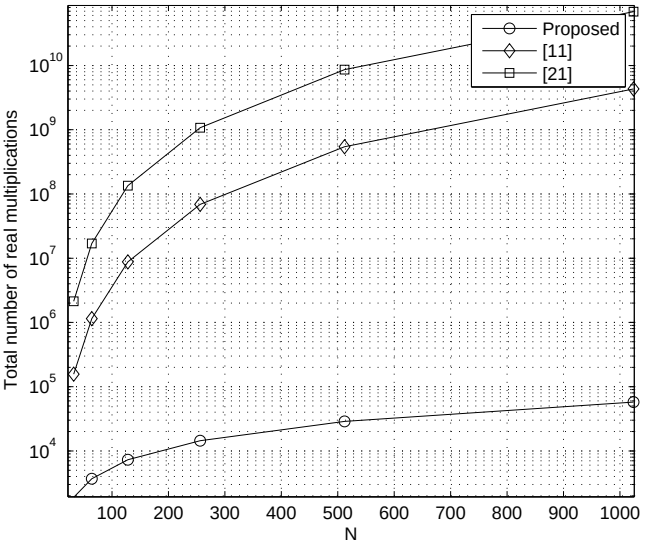

Fig. 12. Complexity of the proposed scheme compared to that of schemes in [11] and [21]. The number of iterations in [21] is one.

[2] B. Razavi, *RF Microelectronics*. Prentice-Hall: Upper Saddle River, 1998.

[3] T. Pollet, M. Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Transactions on Communications*, vol. 43, pp. 191–193, February 1995.

[4] S. Wu and Y. Bar-Ness, "OFDM systems in the presence of phase noise: Consequences and solutions," *IEEE Transactions on Communications*, vol. 52, no. 11, pp. 1988–1996, November 2004.

[5] D. Petrovic, W. Rave, and G. Fettweis, "Effects of phase noise on OFDM systems with and without PLL: Characterization and compensation," *IEEE Transactions on Communications*, vol. 55, no. 8, pp. 1607–1615, August 2007.

[6] A. Tarighat, R. Bagheri, and A. H. Sayed, "Compensation schemes and performance analysis of IQ imbalances in OFDM receivers," *IEEE Transactions on Signal Processing*, vol. 53, pp. 3257–3268, August 2005.

[7] E. Costa and S. Pupolin, "M-QAM-OFDM system performance in the presence of nonlinear amplifier and phase noise," *IEEE Transactions on Communications*, vol. 50, no. 3, pp. 462–472, March 2002.

[8] G. Fettweis, M. Lohning, D. Petrovic, M. Windisch, P. Zillmann, and W. Rave, "Dirty RF: A new paradigm," *IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 4, pp. 2347–2355, September 2005.

[9] A. G. Armada and M. Calvo, "Phase noise and sub-carrier spacing effects on the performance of an OFDM communication system," *IEEE Communications Letters*, vol. 2, pp. 11–13, Jan 1998.

[10] Q. Zou, A. Tarighat, and A. H. Sayed, "Compensation of phase noise in OFDM wireless systems," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5407–5424, November 2007.

[11] D. D. Lin, R. A. Pacheco, T. J. Lim, and D. Hatzinakos, "Joint estimation of channel response, frequency offset, and phase noise in OFDM,"

IEEE Transactions on Signal Processing, vol. 54, no. 9, pp. 3542–3554, September 2006.

[12] F. Munier, T. Eriksson, and A. Svensson, "An ICI reduction scheme for OFDM system with phase noise over fading channels," *IEEE Transactions on Communications*, vol. 56, no. 7, p. 11191126, July 2008.

[13] Y. Zhang and H. Liu, "MIMO-OFDM systems in the presense of phase noise and doubly selective fading," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 2277–2285, July 2007.

[14] R. Corvaja and A. G. Armada, "Joint channel and phase noise compensation for OFDM in fast-fading multipath applications," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 2, pp. 636–643, February 2009.

[15] A. Tarighat and A. H. Sayed, "Joint compensation of transmitter and receiver impairments in OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 6, no. 1, p. 240247, January 2007.

[16] D. Mattera, L. Paura, and F. Sterle, "MMSE WL equalizer in presence of receiver IQ imbalance," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1735–1740, April 2008.

[17] A. Tarighat and A. H. Sayed, "MIMO OFDM receivers for systems with IQ imbalances," *IEEE Transactions on Signal Processing*, vol. 53, no. 9, pp. 3583–3596, September 2005.

[18] K. Y. Sung and C. Chao, "Estimation and compensation of I/Q imbalance in OFDM direct-conversion receivers," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 3, pp. 438–453, June 2009.

[19] C. J. Tsai, C. H. Liao, and T. D. Chiueh, "IQ imbalance and phase noise mitigation for wireless OFDM systems," *IEEE International Symposium on Circuits and Systems*, pp. 2482–2485, May 2008.

[20] Q. Zou, A. Tarighat, and A. H. Sayed, "On the joint compensation of IQ imbalances and phase noise in MIMO-OFDM systems," *IEEE International Symposium on Circuits and Systems*, pp. 37–40, May 2007.

[21] ——, "Joint compensation of IQ imbalance and phase noise in OFDM wireless systems," *IEEE Transactions on Communications*, vol. 57, pp. 404–414, February 2009.

[22] J. Tubbax, B. Come, L. V. der Perre, S. Donnay, M. Engels, M. D. Hugo, and M. Moonen, "Compensation of IQ imbalance and phase noise in OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 872–877, May 2005.

[23] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Transactions on Communications*, vol. 42, pp. 2908–2914, October 1994.

[24] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Transactions on Communications*, vol. 45, no. 12, pp. 1613–1621, December 1997.

[25] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Transactions on Wireless Communications*, vol. 5, no. 5, pp. 1158–1168, May 2006.

[26] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: A unifying theory and numerical methods for characterization," *IEEE Transactions on Circuits and Systems I*, vol. 47, no. 5, pp. 655–674, May 2000.

[27] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE transactions on information theory*, vol. 44, no. 2, pp. 744–765, March 1998.

[28] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, October 1998.

[29] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, pp. 41–59, Autumn 1996.

[30] J. C. Belfiore, G. Rakaya, and E. Viterbo, "High-rate codes that are linear in space and time," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1432–1436, April 2005.

[31] Y. Mostofi and D. C. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Transactions on Wireless Communications*, vol. 4, no. 2, pp. 765–774, March 2005.

[32] D. C. Rife and R. R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE transactions on information theory*, vol. 20, no. 5, pp. 591–598, September 1974.