RECEIVER BASED PAPR REDUCTION IN OFDMA

Anum Ali[1], *Ali Al-Zahrani*[2], *Tareq Y. Al-Naffouri*[1,2] *and Ayman Naguib*[3]

[1]Department of Electrical Engineering,
King Fahd University of Petroleum and Minerals (KFUPM), Saudi Arabia.
[2]Department of Electrical Engineering,
King Abdullah University of Science and Technology (KAUST), Saudi Arabia.
[3]Qualcomm Research Silicon Valley, USA.
anumali@ieee.org, {ali.zahrani,tareq.alnaffouri}@kaust.edu.sa, anaguib@qti.qualcomm.com

ABSTRACT

High peak-to-average power ratio is one of the major drawbacks of orthogonal frequency division multiplexing (OFDM). Clipping is the simplest peak reduction scheme, however, it requires clipping mitigation at the receiver. Recently compressed sensing has been used for clipping mitigation (by exploiting the sparse nature of clipping signal). However, clipping estimation in multi-user scenario (i.e., OFDMA) is not straightforward as clipping distortions overlap in frequency domain and one cannot distinguish between distortions from different users. In this work, a collaborative clipping removal strategy is proposed based on joint estimation of the clipping distortions from all users. Further, an effective data aided channel estimation strategy for clipped OFDM is also outlined. Simulation results are presented to justify the effectiveness of the proposed schemes.

Index Terms— PAPR reduction, clipping, channel estimation, compressed sensing, OFDMA

1. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is a popular modulation scheme that finds applications in many current and emerging communication standards [1]. The main idea in OFDM communications is to establish a flat fading channel per sub-carrier so that one tap equalization can be performed. On the down side, one major drawback of OFDM is high peak-to-average power ratio (PAPR) of the time domain signal, which leads to inefficient use of high power amplifier (HPA). Operation in the nonlinear region of the HPA is power efficient but causes distortions. Though one option is to back off and let the HPA operate in the linear region, it results in power inefficiency. A high PAPR signal requires HPAs with linear response over a wide range, hence expensive transmitters.

Many transmitter-based schemes have been proposed for PAPR reduction e.g., coding, partial transmit sequence (PTS), selected mapping (SLM), interleaving, tone reservation (TR), tone injection (TI) and active constellation extension (ACE) [2–5]. However, these schemes are quite complex and hence are themselves a source of energy inefficiency at the transmitter. In applications with limited source of power (e.g., satellites and mobile phones), the complexity/power efficiency is a bottleneck, and hence there is a need for energy efficient low complexity alternatives to PAPR reduction.

Recently, compressive sensing (CS) techniques have been proposed for PAPR reduction [6,7]. These techniques require simple clipping at the transmitter and relegate any distortion mitigation to the receiver side. Clipping results in in-band and out-of-band distortions in the frequency domain. By partially observing these distortions at reserved carriers (in-band or out-of-band), one can recover the time-sparse clipping signal using some ℓ_1-optimization. However, these schemes are tailored for the single user scenario and are inadequate for orthogonal frequency division multiple access (OFDMA) systems where multiple users are subjected to clipping at their respective transmitters. This is because the clipping distortions from all users overlap in frequency domain and one cannot distinguish between distortions from different users. Here, noteworthy is the fact that [6–9] and the proposed scheme assume the channel impulse response (CIR) knowledge at the receiver. However, the task of channel estimation in clipped OFDM is challenging, as the the pilot tones used for estimation are also corrupted by the clipping distortions.

In this paper, we tackle the problem of multi-user clipping distortion mitigation. A two stage recovery procedure is proposed in which initially, the clipping distortion experienced by all users is jointly estimated. The joint estimation of the clipping vector can be performed using any sparse reconstruction scheme (e.g., ℓ_1-minimization [10], Bayesian methods [11–14] and matching pursuits [15,16]). These joint estimates are then used in conjunction with a technique proposed by Wu *et. al.* [17] to form the uncoupled systems. Henceforth, we refer to [17] by calling it the contaminated pilot approach (for the reasons that will become clear later). This kick-starts a decoupled distortion-estimation stage where the clipping from each user is estimated individually and subtracted from the OFDMA signal to progressively reduce the distortion in the received signal. In addition, we present a CIR estimation scheme for clipped OFDM that uses reliable carriers (RC) and the contaminated pilot approach [17], to aid the minimum mean square error (MMSE) estimate. This approach estimates the contaminated training symbols (and RCs) actually transmitted and uses them for enhanced channel estimation.

We adopt the following notations in the paper: Italic letters (e.g., N) for scalers, lower-case bold-face letters (e.g., $\mathbf{x}$) for time domain vectors, upper-case bold-face calligraphic letters (e.g., $\boldsymbol{\mathcal{X}}$) for frequency domain vectors and upper-case bold-face letters (e.g., $\mathbf{X}$) for matrices. The symbols $\hat{\mathbf{x}}$, $\mathbf{x}(l)$, $\mathbf{x}^{\mathsf{T}}$ and $\mathbf{x}^{\mathsf{H}}$ represent the estimate, l^{th} entry, transpose and hermitian (conjugate transpose) of the vector $\mathbf{x}$. Further, $\mathbb{E}[\cdot]$, $\mathbf{I}_N$ and $\mathbf{0}$ denote the expectation operator, identity matrix of size $N \times N$ and the zero vector respectively. The operator

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

$\text{diag}(\mathbf{X})$ forms a column vector $\mathbf{x}$ from the diagonal of $\mathbf{X}$ and $\text{diag}(\mathbf{x})$ construct a diagonal matrix $\mathbf{X}$ with $\mathbf{x}$ on its diagonal. Finally, $\underline{\boldsymbol{\mathcal{X}}}^i$ represents i^{th} portion of the vector $\boldsymbol{\mathcal{X}}$, where $\boldsymbol{\mathcal{X}}$ is partitioned in I segments.

The remainder of the paper is organized as follows: Section 2 introduces data model for OFDMA signals subjected to an amplitude limiting operation. In Section 3 the proposed multi-user clipping reconstruction scheme is presented. Section 4 presents the proposed channel estimation strategy. Simulation results are presented in Section 5 and Section 6 concludes the paper.

2. DATA MODEL

In an OFDM system, serially incoming bits are divided into N low rate parallel streams. An N dimensional data vector $\boldsymbol{\mathcal{X}} = [\mathcal{X}(0), \mathcal{X}(1), \cdots, \mathcal{X}(N-1)]^{\mathsf{T}}$ results when these parallel streams are mapped to an M-ary QAM alphabet $\{\mathcal{A}_0, \mathcal{A}_1, \cdots, \mathcal{A}_{M-1}\}$. The time domain signal vector is obtained by an inverse discrete Fourier transform (IDFT) operation, so that $\mathbf{x} = \mathbf{F}^{\mathsf{H}}\boldsymbol{\mathcal{X}}$. Here $\mathbf{F}$ denotes the unitary DFT matrix with (n_1, n_2) element

$$\mathrm{f}_{n_1,n_2} = 1/\sqrt{N}e^{-j2\pi n_1 n_2/N}, \quad n_1, n_2 \in 0, 1, \cdots, N-1$$

In OFDMA, each user is assigned a subset of sub-carriers, and each carrier is assigned exclusively to one user [18]. For simplicity we confine our attention to the two user OFDMA scenario, although our approach naturally extends to the general I user case. Hence at the user equipment, the incoming stream of data is divided into $K = N/2$ parallel streams followed by modulation. In the context of a complete OFDMA symbol, the frequency domain signal corresponding to the 1[st] user is given as

$$\boldsymbol{\mathcal{X}}^1 = [\mathcal{X}^1(0), \mathcal{X}^1(1), \cdots, \mathcal{X}^1(K-1), \mathbf{0}^{\mathsf{T}}]^{\mathsf{T}} = [\underline{\boldsymbol{\mathcal{X}}}^{1\mathsf{T}}, \mathbf{0}^{\mathsf{T}}]^{\mathsf{T}} \quad (1)$$

where $\underline{\boldsymbol{\mathcal{X}}}^1 = [\mathcal{X}^1(0), \mathcal{X}^1(1), \cdots, \mathcal{X}^1(K-1)]^{\mathsf{T}}$ is a length K vector corresponding to the data from 1[st] user (similarly we have $\boldsymbol{\mathcal{X}}^2 = [\mathbf{0}^{\mathsf{T}}, \underline{\boldsymbol{\mathcal{X}}}^{2\mathsf{T}}]^{\mathsf{T}}$). The time domain signal for the i^{th} user (i.e., $\mathbf{x}^i$ for $i = 1, 2$) is obtained by taking the IDFT of $\boldsymbol{\mathcal{X}}^i$. To reduce the PAPR, the signals $\mathbf{x}^i$ are subjected to clipping, such that the maximum amplitude is limited to the threshold γ^i, i.e.,

$$x_p^i(n) = \begin{cases} \gamma^i e^{\jmath\angle x^i(n)} & \text{if } |x^i(n)| > \gamma^i \\ x^i(n) & \text{otherwise} \end{cases} \quad \text{for } i = 1, 2 \quad (2)$$

where $x_p^i(n)$ is the n^{th} element of the signal from i^{th} user after clipping, γ^i is the limiting threshold for i^{th} user and $\angle x^i(n)$ is the phase of $x^i(n)$. The clipping ratio (CR) of the i^{th} user and threshold γ^i are related by $\text{CR}^i = \gamma^i/\sigma^i$, where σ^i is the root mean squared power of the OFDM signal transmitted by the i^{th} user.

The hard clipping in (2) is equivalent to adding a *sparse* signal $\mathbf{c}^i$ to the original time domain signal $\mathbf{x}^i$ with active elements only where clipping has occurred, i.e.,

$$\mathbf{x}_p^i = \mathbf{x}^i + \mathbf{c}^i \quad (3)$$

where $\mathbf{c}^i$ represents the clipping signal for the i^{th} user. The clipped signal is transmitted through a channel with the impulse response $\mathbf{h}^i = [h^i(0), h^i(1), \cdots, h^i(N_c-1)]^{\mathsf{T}}$. The time domain signal from i^{th} user at the output of the channel can be written as $\mathbf{y}^i = \mathbf{H}^i\mathbf{x}_p^i$, where $\mathbf{H}^i$ is the channel matrix which is circulated by virtue of the cyclic prefix. Thus, we can diagonalize $\mathbf{H}^i$ using the DFT matrix and write $\mathbf{H}^i = \mathbf{F}^{\mathsf{H}}\mathbf{D}^i\mathbf{F}$, where $\mathbf{D}$ is a diagonal matrix containing the frequency response coefficients on its diagonal. The channel frequency response is given as $\boldsymbol{\mathcal{D}}^i = \text{diag}(\mathbf{D}^i) = \sqrt{N}\bar{\mathbf{F}}\mathbf{h}^i$ (where, $\bar{\mathbf{F}}$ represents the partial DFT matrix). In the development of the proposed scheme, we assume the knowledge of the CIR; hence the frequency response matrix $\mathbf{D}^i$ and the circulant channel matrix $\mathbf{H}^i$ are perfectly known. The procedure for obtaining the CIR is discussed later in Section 4.

At the base station we have $\mathbf{y} = \sum_{i=1}^2 \mathbf{H}^i(\mathbf{x}^i + \mathbf{c}^i) + \mathbf{z}$, which is the combined channel output of both users plus additive white Gaussian noise (AWGN), $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \sigma_z^2\mathbf{I}_N)$. The frequency domain received signal can be obtained by the DFT operation as $\boldsymbol{\mathcal{Y}} = \sum_{i=1}^2 \mathbf{D}^i\boldsymbol{\mathcal{X}}^i + \sum_{i=1}^2 \mathbf{D}^i\boldsymbol{\mathcal{C}}^i + \boldsymbol{\mathcal{Z}}$. The frequency domain white Gaussian vector $\boldsymbol{\mathcal{Z}}$ has the same statistics as $\mathbf{z}$ i.e., $\boldsymbol{\mathcal{Z}} \sim \mathcal{CN}(\mathbf{0}, \sigma_z^2\mathbf{I}_{N\times N})$. The product column vector $\mathbf{D}^i\boldsymbol{\mathcal{X}}^i$ will have nonzero elements only at the locations corresponding to the sub-carriers that are allocated to user i for transmission. Hence $\sum_{i=1}^2 \mathbf{D}^i\boldsymbol{\mathcal{X}}^i$ can be written as $\mathbf{D}\boldsymbol{\mathcal{X}}$, where

$$\boldsymbol{\mathcal{X}} = \begin{bmatrix} \underline{\boldsymbol{\mathcal{X}}}^1 \\ \underline{\boldsymbol{\mathcal{X}}}^2 \end{bmatrix} \quad \text{and} \quad \mathbf{D} = \begin{bmatrix} \underline{\mathbf{D}}^1 & \\ & \underline{\mathbf{D}}^2 \end{bmatrix}$$

Note that although the channel frequency responses $\mathbf{D}^i$ for $i = 1, 2$ are diagonal matrices of size $N \times N$ and hence are overlapping, the matrix $\mathbf{D}$ comprises of only the portions of $\mathbf{D}^i$, belonging to the i^{th} user band, which is denoted by $\underline{\mathbf{D}}^i$. Now we can write

$$\boldsymbol{\mathcal{Y}} = \mathbf{D}\boldsymbol{\mathcal{X}} + \sum_{i=1}^{2} \mathbf{D}^i\boldsymbol{\mathcal{C}}^i + \boldsymbol{\mathcal{Z}} \quad (4)$$

Unlike $\sum_{i=1}^2 \mathbf{D}^i\boldsymbol{\mathcal{X}}^i$ whose component do not overlap, the distortions $\boldsymbol{\mathcal{C}}^i$ spill over both users and $\sum_{i=1}^2 \mathbf{D}^i\boldsymbol{\mathcal{C}}^i$ cannot be simplified into the product of a matrix and a column vector and needs further investigation.

In the absence of distortion (i.e., when the second term of (4) is zero), the receiver could easily separate the users (as they occupy different carriers) and equalize the users' channels to recover the transmitted data. Mathematically, we can write $\underline{\boldsymbol{\mathcal{Y}}}^i = \underline{\mathbf{D}}^i\underline{\boldsymbol{\mathcal{X}}}^i + \underline{\boldsymbol{\mathcal{Z}}}^i$, where $\underline{\boldsymbol{\mathcal{Y}}}^i$ is the portion of $\boldsymbol{\mathcal{Y}}$ confined to the carriers of the i^{th} user (a similar definition applies to $\underline{\mathbf{D}}^i$,$\underline{\boldsymbol{\mathcal{X}}}^i$ and $\underline{\boldsymbol{\mathcal{Z}}}^i$). Upon equalizing, we obtain

$$\widehat{\underline{\boldsymbol{\mathcal{X}}}^i} = (\underline{\mathbf{D}}^i)^{-1}\underline{\boldsymbol{\mathcal{Y}}}^i = \underline{\boldsymbol{\mathcal{X}}}^i + (\underline{\mathbf{D}}^i)^{-1}\underline{\boldsymbol{\mathcal{Z}}}^i \quad (5)$$

The noisy estimate $\widehat{\underline{\boldsymbol{\mathcal{X}}}^i}$ is then rounded to the nearest constellation point, which we denote by $\lfloor\widehat{\underline{\boldsymbol{\mathcal{X}}}^i}\rfloor$. However, in presence of the distortions, clipping needs to be estimated and cancelled before the equalization step of (5).

3. COLLABORATIVE CLIPPING RECOVERY FOR OFDMA

In this section, we demonstrate how clipping distortion can be estimated. To this end, let us re-write (4) as

$$\boldsymbol{\mathcal{Y}} = \mathbf{D}\boldsymbol{\mathcal{X}} + [\mathbf{D}^1\mathbf{D}^2]\begin{bmatrix} \boldsymbol{\mathcal{C}}^1 \\ \boldsymbol{\mathcal{C}}^2 \end{bmatrix} + \boldsymbol{\mathcal{Z}} = \mathbf{D}\boldsymbol{\mathcal{X}} + [\mathbf{D}^1\mathbf{D}^2]\mathbf{F}\begin{bmatrix} \mathbf{c}^1 \\ \mathbf{c}^2 \end{bmatrix} + \boldsymbol{\mathcal{Z}} \quad (6)$$

where we have made the substitution $\boldsymbol{\mathcal{C}}^i = \mathbf{F}\mathbf{c}^i$. Note that the vector $\mathbf{c} = [\mathbf{c}^{1\mathsf{T}}\mathbf{c}^{2\mathsf{T}}]^{\mathsf{T}}$ is sparse in time and hence a few measurements in the dual (frequency) domain are enough to estimate $\mathbf{c}$ [19,20]. Let, $\mathbf{S}_c$ denote an $N \times N$ diagonal binary selection matrix with J number of 1s along its diagonal at the locations of the reserved (data free) carriers. This selection matrix extracts the element of any column vector corresponding to the reserved sub-carriers. Now proceed by projecting $\boldsymbol{\mathcal{Y}}$ onto the subspace spanned by the reserved carriers. This yields

$$\mathbf{S}_c\boldsymbol{\mathcal{Y}}=\mathbf{S}_c(\mathbf{D}\boldsymbol{\mathcal{X}}+[\mathbf{D}^1\mathbf{D}^2]\mathbf{F}\mathbf{c}+\boldsymbol{\mathcal{Z}}) \text{ i.e., } \boldsymbol{\mathcal{Y}}'=\boldsymbol{\Psi}\mathbf{c}+\boldsymbol{\mathcal{Z}}' \tag{7}$$

where the first term in (7) is zero because the carriers corresponding to reserved tones are data free and hence zero. While (7) is underdetermined, we can recover $\mathbf{c}$ using many available sparse-reconstruction schemes e.g., [10–13, 15, 16, 21]. Consider as an example using least absolute shrinkage and selection operator (LASSO) [22], then we have

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}}\{\|\boldsymbol{\Psi}\mathbf{c}-\boldsymbol{\mathcal{Y}}'\|_2^2+\lambda\|\mathbf{c}\|_1\} \tag{8}$$

This was pursued in [6] for the single user case by reserving J data free carriers ($J << N$) and using them to estimate the time-domain distortion and cancel it. This strategy is not as effective in the multiuser case especially as the number of users gets larger. For example the in the two user scenario of (4), the sparse vector is twice as large and could have twice the number of active elements. As such to maintain the quality of the estimate in 2-user scenario, we need to double the number of free carriers, which will reduce the throughput. Alternatively in this paper, we get by with the estimate obtained from (8) and once these estimates are available we proceed in a decoupled manner to improve these estimates.

Once the clipping signals are initially reconstructed using (8) (i.e., the joint estimation), it is possible to setup two uncoupled systems of equations for user 1 and 2 respectively. Once the isolated systems are formed, the sparse clipping reconstruction can be performed for each user for enhanced recovery. Therefore, the crux of the proposed reconstruction scheme can be summarized in the following two steps: 1) Estimate $\mathbf{c} = [\mathbf{c}^{1\mathsf{T}}\mathbf{c}^{2\mathsf{T}}]^\mathsf{T}$ via joint sparse reconstruction using (8) and 2) Decouple the two systems of linear equations corresponding to user 1 and user 2 and perform clipping reconstruction for each user.

To obtain the decoupled systems, we modify the contaminated pilot approach initially proposed for channel estimation [17]. It was noted that as the clipped signal is transmitted (transmitted pilots are also clipped) hence it is not optimal to use ideal pilot sequence at the receiver as a reference for channel estimation. Instead, the clipped pilot sequence was first estimated at the receiver and then used for enhanced channel estimation. As the clipped pilots are used in [17] instead of clean pilot signals, we call this scheme the contaminated pilot approach. In this work, we use the idea of reconstructing the clipped version of the transmitted signal at the receiver to form the decoupled systems. To do that, the initial estimate of $\mathbf{c}$ obtained using (8) is subtracted from (4) to get

$$\boldsymbol{\mathcal{Y}}_{\text{cs}} = \boldsymbol{\mathcal{Y}} - [\mathbf{D}^1\mathbf{D}^2]\mathbf{F}\hat{\mathbf{c}} = \mathbf{D}\boldsymbol{\mathcal{X}}+\boldsymbol{\mathcal{Z}}' \tag{9}$$

We proceed by extracting the carriers allocated to user i and get $\underline{\boldsymbol{\mathcal{Y}}}^i_{\text{cs}}$, which is then equalized using (5) to obtain $\widehat{\underline{\boldsymbol{\mathcal{X}}}^i} = (\underline{\mathbf{D}}^i)^{-1}\underline{\boldsymbol{\mathcal{Y}}}^i_{\text{cs}}$. Now, we estimate the transmitted frequency domain signal by making the maximum likelihood (ML) decisions $\lfloor\widehat{\underline{\boldsymbol{\mathcal{X}}}^i}\rfloor$ (the operation $\lfloor\cdot\rfloor$ represents rounding to the nearest constellation point). The time domain signal is obtained by IDFT as $\widehat{\mathbf{x}^i} = \mathbf{F}^\mathsf{H}\lfloor\widehat{\underline{\boldsymbol{\mathcal{X}}}^i}\rfloor$. This time domain signal is then clipped using (1) to get $\widehat{\mathbf{x}^i_p}$. Now the difference between the clipped and unclipped versions of $\widehat{\mathbf{x}}^i$ i.e., $\widehat{\mathbf{c}}^i = (\widehat{\mathbf{x}}^i_p - \widehat{\mathbf{x}}^i)$ is entrusted as the improved estimate of the clipping distortion and is subtracted from (4) to form the decoupled systems. The stepwise procedure for formulation of the decoupled system of user 2 is outlined below:

1. Do the joint sparse signal reconstruction based on (8).
2. Subtract the estimated distortion $\hat{\mathbf{c}}$ from (4) to obtain $\boldsymbol{\mathcal{Y}}_{\text{cs}} = \boldsymbol{\mathcal{Y}} - [\mathbf{D}^1\mathbf{D}^2]\mathbf{F}\hat{\mathbf{c}} = \mathbf{D}\boldsymbol{\mathcal{X}}+\boldsymbol{\mathcal{Z}}$
3. Get $\underline{\boldsymbol{\mathcal{Y}}}^1_{\text{cs}} = \underline{\mathbf{D}}^1\boldsymbol{\mathcal{X}}^1+\underline{\boldsymbol{\mathcal{Z}}}^1$ by extracting user 1's carriers.
4. Equalize $\underline{\boldsymbol{\mathcal{Y}}}^1_{\text{cs}}$ using (5) and obtain $(\lfloor\widehat{\underline{\boldsymbol{\mathcal{X}}}^1}\rfloor)$.
5. Using pilots and $\lfloor\widehat{\underline{\boldsymbol{\mathcal{X}}}^1}\rfloor$, form a time domain signal $\widehat{\mathbf{x}^1}$.
6. Obtain $\widehat{\mathbf{x}^1_p}$ from $\widehat{\mathbf{x}^1}$ using (1) and obtain $\widehat{\boldsymbol{\mathcal{C}}^1} = \widehat{\boldsymbol{\mathcal{X}}}^1_p - \widehat{\boldsymbol{\mathcal{X}}^1}$.
7. Obtain $\boldsymbol{\mathcal{Y}}^2=\boldsymbol{\mathcal{Y}}-\mathbf{D}^1\widehat{\boldsymbol{\mathcal{C}}^1}=\mathbf{D}\boldsymbol{\mathcal{X}}+\mathbf{D}^2\boldsymbol{\mathcal{C}}^2+\boldsymbol{\mathcal{Z}}$ based on (4).

Note that $\boldsymbol{\mathcal{Y}}^2$ is decoupled from user 1's clipping. Now with this decoupled system for user 2, we can extract subcarriers allocated to user 2 to form $\underline{\boldsymbol{\mathcal{Y}}}^2 = \underline{\mathbf{D}}^2\underline{\boldsymbol{\mathcal{X}}}^2+\underline{\mathbf{D}}^2\boldsymbol{\mathcal{C}}^2+\underline{\boldsymbol{\mathcal{Z}}}^2$ and reconstruct $\mathbf{c}^2$ using sparse recovery. The formulation and reconstruction of $\mathbf{c}^1$ can be done in a similar manner.

4. DATA AIDED CHANNEL ESTIMATION STRATEGY

Let us discuss the channel estimation problem for clipped OFDM[1]. The received OFDM signal can be written as

$$\boldsymbol{\mathcal{Y}} = \mathbf{D}\boldsymbol{\mathcal{X}}+\mathbf{D}\boldsymbol{\mathcal{C}}+\boldsymbol{\mathcal{Z}} = \mathbf{D}\boldsymbol{\mathcal{X}}+\boldsymbol{\mathcal{Z}}' \tag{10}$$

where $\boldsymbol{\mathcal{Z}}' = \mathbf{D}\boldsymbol{\mathcal{C}}+\boldsymbol{\mathcal{Z}}$ is the combined AWGN noise and clipping distortion. Note that, $\mathbf{D}\boldsymbol{\mathcal{X}}$ is a product of a diagonal matrix and a column vector, and hence we can exchange the roles of $\mathbf{D}$ and $\boldsymbol{\mathcal{X}}$ by rewriting (10) as

$$\begin{aligned}\boldsymbol{\mathcal{Y}} &= \text{diag}(\boldsymbol{\mathcal{X}})\text{diag}(\mathbf{D})+\boldsymbol{\mathcal{Z}}' = \text{diag}(\boldsymbol{\mathcal{X}})\boldsymbol{\mathcal{D}}+\boldsymbol{\mathcal{Z}}'\\ &= \sqrt{N}\text{diag}(\boldsymbol{\mathcal{X}})\bar{\mathbf{F}}\mathbf{h}+\boldsymbol{\mathcal{Z}}' = \mathbf{X}\mathbf{h}+\boldsymbol{\mathcal{Z}}'\end{aligned} \tag{11}$$

where $\mathbf{X} \triangleq \sqrt{N}\text{diag}(\boldsymbol{\mathcal{X}})\bar{\mathbf{F}}$. For channel estimation in OFDM, P equally spaced pilot signals are inserted at the transmitter [23–25]. Based on this known pilot sequence, the receiver finds the MMSE esitmate of the channel [26–28]. Let $\mathcal{I}_p$ denote the index set of the pilot locations, then we can write $\boldsymbol{\mathcal{Y}}_{\mathcal{I}_p} = \mathbf{X}_{\mathcal{I}_p}\mathbf{h}+\boldsymbol{\mathcal{Z}}'_{\mathcal{I}_p}$, where $\boldsymbol{\mathcal{U}}_{\mathcal{I}_p}$ prunes $\boldsymbol{\mathcal{U}}$ of all rows except for the rows belonging to $\mathcal{I}_p$. Now the MMSE estimate of $\mathbf{h}$ can be obtained by solving the regularized least squares (LS) problem $\hat{\mathbf{h}} = \arg\max_{\mathbf{h}}\{\|\boldsymbol{\mathcal{Y}}_{\mathcal{I}_p}-\mathbf{X}_{\mathcal{I}_p}\mathbf{h}\|^2_{\mathbf{R}^{-1}_{\boldsymbol{\mathcal{Z}}'}}+\|\mathbf{h}\|^2_{\mathbf{R}^{-1}_h}\}$ where $\mathbf{R}_h = \mathbb{E}[\mathbf{h}\mathbf{h}^\mathsf{H}] = \sigma_h^2\mathbf{I}_{N_c}$. Further, ignoring the clipping noise component of $\boldsymbol{\mathcal{Z}}'$ we can write $\mathbf{R}_{\boldsymbol{\mathcal{Z}}'} = \mathbb{E}[\boldsymbol{\mathcal{Z}}'\boldsymbol{\mathcal{Z}}'^\mathsf{H}] = \sigma_z^2\mathbf{I}_P$ (the subscript $\mathcal{I}_p$ of $\boldsymbol{\mathcal{Z}}'$ is dropped here for notational convenience). Solving this LS problem yields [29]

$$\hat{\mathbf{h}} = \mathbf{X}^\mathsf{H}_{\mathcal{I}_p}(\mathbf{X}_{\mathcal{I}_p}\mathbf{X}^\mathsf{H}_{\mathcal{I}_p}+(\sigma_z^2/\sigma_h^2)\mathbf{I}_P)^{-1}\boldsymbol{\mathcal{Y}}_{\mathcal{I}_p} \tag{12}$$

Increasing the number of pilot tones for CIR estimation results in improved estimation accuracy. However, generally it is not feasible to spare additional pilots as it reduces the data rate. In this work, we increase the number of measurements without increasing the number of reserved pilots by using the concept of RCs. Specifically speaking, at the receiver, given $\boldsymbol{\mathcal{Y}}$ and the initial estimate of the channel (found using (12)) we expect the following: for some sub-carriers,

[1]For clarity of exposition we consider the single user case for CIR estimation. The proposed scheme extends easily to the multiuser case.

the combined AWGN noise and clipping perturbation $\mathcal{Z}'(i)$ is strong enough to take $\mathcal{Y}(i)/\hat{\mathcal{D}}(i)$ out of its correct decision region i.e., $\lfloor \mathcal{Y}(i)/\hat{\mathcal{D}}(i) \rfloor = \lfloor \hat{\mathcal{X}}(i) \rfloor \neq \mathcal{X}(i)$, while for others $\lfloor \hat{\mathcal{X}}(i) \rfloor = \mathcal{X}(i)$. The subset of data carries which satisfy $\lfloor \hat{\mathcal{X}}(i) \rfloor = \mathcal{X}(i)$ are called reliable data carriers and fortunately constitute a major part of all sub-carriers. We can write the reliability function for i^{th} sub-carrier in terms of $\mathcal{Z}'(i)$ as

$$\mathfrak{R}(i) = \frac{p(\mathcal{Z}'(i) = \mathcal{X}(i) - \lfloor \hat{\mathcal{X}}(i) \rfloor)}{\sum_{k=0, \mathcal{A}_k \neq \lfloor \hat{\mathcal{X}}(i) \rfloor}^{M-1} p(\mathcal{Z}(i) = \mathcal{X}(i) - \mathcal{A}_k)} \tag{13}$$

where $p(\cdot)$ represents the pdf of $\boldsymbol{\mathcal{Z}}'$, which is assumed to be zero mean Gaussian with variance $\sigma^2_{\mathcal{Z}'}$ (see [30] for details). Hence, the numerator is the probability that $\mathcal{Z}'(i)$ does not take $\mathcal{X}(i)$ beyond its correct decision region and the denominator sums the probabilities of all possible incorrect decisions that $\mathcal{Z}'(i)$ can cause (i.e., $\mathcal{Z}'(i)$ takes $\lfloor \hat{\mathcal{X}}(i) \rfloor$ to a QAM constellation point $\mathcal{A}_k$ different from $\mathcal{X}(i)$). From (13), it is obvious that the higher the value of $\mathfrak{R}$ the higher the probability of staying in correct decision region and hence the higher the reliability of the carrier. In other words, the ML decisions on the carriers with higher values of $\mathfrak{R}$ will not be erroneous with a high probability. The detailed investigation of this reliability criteria is reported in [30]. Once the reliability of all data carriers is known, we select R carriers with highest reliability values. Let $\mathcal{I}_r$ denote the index set of the RCs and the pilot carriers. We can now retain these carriers in estimating $\mathbf{h}$ and prune all other sub-carriers from (11). This yields

$$\boldsymbol{\mathcal{Y}}_{\mathcal{I}_r} = \mathbf{X}_{\mathcal{I}_r}\mathbf{h} + \boldsymbol{\mathcal{Z}}'_{\mathcal{I}_r} \tag{14}$$

Now we can obtain the refined estimate of $\mathbf{h}$ based on (12) by replacing the pilot index set $\mathcal{I}_p$ with enhanced set $\mathcal{I}_r$ consisting of the pilots and RCs. The enhanced MMSE estimation procedure based on reliable data carriers can be summarized in the following three steps: 1) Find the initial MMSE estimate of the CIR using (12), 2) Find reliability $\mathfrak{R}$ for all sub-carriers and select R sub-carriers with highest reliability index as RCs and 3) Use RCs as additional measurements (by using (14)) and find MMSE estimate using (12).

It is important to note that however many pilots and RCs we use to enhance the channel estimate, we are bottle-necked by the clipping distortions. Another way to look at this is to notice that what passes through the channel is not the pure signal or pilots but their clipped versions. As such, motivated by the work of [17], we first estimate the contaminated pilots and the contaminated RCs and use them for enhanced MMSE estimation. The proposed data aided CIR estimation scheme can be summarized as:

1. Obtain the initial MMSE estimate by using (12).
2. Equalize the received data and make ML decisions on the equalized data i.e., $\lfloor \mathcal{Y}(i)/\hat{\mathcal{D}}(i) \rfloor = \lfloor \hat{\mathcal{X}}(i) \rfloor$.
3. Find reliability $\mathfrak{R}$ for all sub-carriers and select R sub-carriers with highest reliability index as RCs.
4. Construct time domain signal $\hat{\mathbf{x}} = \mathbf{F}^{\mathsf{H}} \lfloor \hat{\mathcal{X}}(i) \rfloor$.
5. Find $\hat{\mathbf{x}}_p$ by clipping $\hat{\mathbf{x}}$ using (2) and obtain $\hat{\boldsymbol{\mathcal{X}}}_p = \mathbf{F}\hat{\mathbf{x}}_p$.
6. Project $\hat{\boldsymbol{\mathcal{X}}}_p$ on $\mathbf{S}_r$ to obtain (clipped pilot sequence and reliable carriers) $\boldsymbol{\mathcal{X}}_{p_{\mathcal{I}_r}}$ and $\mathbf{X}_{p_{\mathcal{I}_r}} = \text{diag}(\boldsymbol{\mathcal{X}}_{p_{\mathcal{I}_r}})$.
7. Use $\mathbf{X}_{p_{\mathcal{I}_r}}$ in (12) to obtain the improved CIR estimate.

5. SIMULATION RESULTS

For clipping recovery an OFDMA system with 512 subcarriers is used and each user is assigned 256 consecutive carriers. 64-QAM and 16-QAM alphabets are used for modulation. A 10 tap rayleigh fading channel is used where each tap comes from standard normal distribution and results are averaged over 5×10^3 bit errors. Random pilots are allocated for CS measurements and the fast Bayesian matching pursuit (FBMP) [12] is used for sparse signal reconstruction. Perfect CIR knowledge is assumed and CR$=$ 1.61. Fig. 1 presents results for the proposed clipping reconstruction scheme as a function of Eb/No (number of measurements fixed at 80) and number of CS measurement tones (Eb/No fixed at 27dB). It is evident from results that proposed clipping mitigation provides good results for both 16-QAM and 64-QAM and approaches the no clipping case for high Eb/No values. For

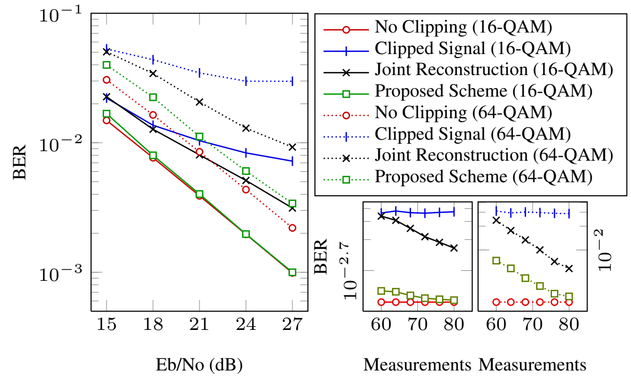

Fig. 1. BER of the proposed clipping mitigation scheme.

channel estimation 256 sub-carrier OFDM and 64-QAM modulation is used. A total of 16 equispaced pilots are inserted for channel estimation and the number of RCs is chosen to be 16. Fig. 2 shows the mean squared error (MSE) results of simple MMSE estimation (MMSE), the RCs approach (RC), the contaminated pilot approach (CPA), the proposed scheme (RC+CPA) and the MMSE for unclipped OFDM (No clipping). The MSE as a function of Eb/No results are generated by keeping CR=1.73. The results show that for high Eb/No the proposed scheme provides upto 8 dB advantage on simple MMSE estimation. Fig. 2 also shows MSE for varying CR and fixed Eb/No=15dB. It is clear from the results that for any value of CR the proposed estimation scheme (RC+CPA) outperforms other reconstruction schemes.

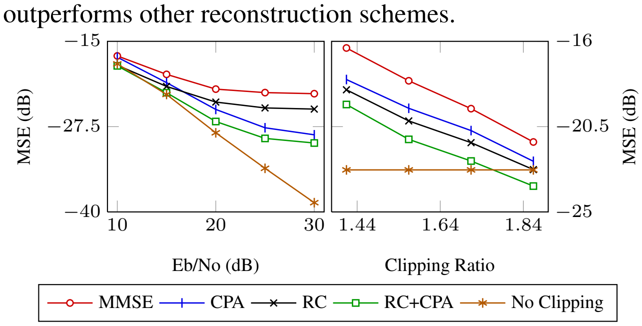

Fig. 2. MSE (dB) of the proposed CIR estimation scheme.

6. CONCLUSION

A multi-user PAPR reduction scheme is proposed that uses joint-estimation for initial clipping reconstruction and the contaminated pilot approach to form the uncoupled systems. Further a data aided channel estimation strategy is proposed using RCs and the contaminated pilot approach. The results demonstrate the effectiveness of the proposed distortion cancellation and channel estimation schemes.

7. REFERENCES

[1] A. Stamoulis, S. N. Diggavi, and N. Al-Dhahir, “Intercarrier interference in MIMO-OFDM,” *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2451–2464, 2002.

[2] S. H. Han and J. H. Lee, “An overview of peak-to-average power ratio reduction techniques for multicarrier transmission,” *IEEE Wireless Commun.*, vol. 12, no. 2, pp. 56–65, 2005.

[3] L. Wang and C. Tellambura, “An Overview of peak-to-average power ratio reduction techniques for OFDM systems,” in *Proc. IEEE ISSPIT*, 2006, pp. 840–845.

[4] T. Jiang and Y. Wu, “An overview: peak-to-average power ratio reduction techniques for OFDM signals,” *IEEE Trans. Broadcast.*, vol. 54, no. 2, pp. 257–268, 2008.

[5] D. W. Lim, S. J. Heo, and J. S. No, “An overview of peak-to-average power ratio reduction schemes for OFDM signals,” *J. Commun. Netw.*, vol. 11, no. 3, pp. 229–239, 2009.

[6] E. B. Al-Safadi and T. Y. Al-Naffouri, “Peak reduction and clipping mitigation in OFDM by augmented compressive sensing,” *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3834–3839, 2012.

[7] A. Al-Rabah, M. Masood, A. Ali, and T. Y. Al-Naffouri, “Receiver-based Bayesian PAPR reduction in OFDM,” in *Proc. Eur. Signal Process. Conf.*, 2013.

[8] J. Tellado, *Multicarrier modulation with low PAR: applications to DSL and wireless*, 2000: Kluwer Academic Press.

[9] D. Kim and G. L. Stuber, “Clipping noise mitigation for ofdm by decision-aided reconstruction,” *IEEE Commun. Lett.*, vol. 3, no. 1, pp. 4–6, 1999.

[10] E. J. Candès and M. B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, 2008.

[11] S. Ji, Y. Xue, and L. Carin, “Bayesian compressive sensing,” *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346–2356, 2008.

[12] P. Schniter, L. C. Potter, and J. Ziniel, “Fast Bayesian matching pursuit,” in *Proc. Inform. Theory & Appl. Workshop, Feb. 2008*.

[13] A. A. Quadeer and T. Y. Al-Naffouri, “Structure-Based Bayesian Sparse Reconstruction,” *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6354–6367, 2012.

[14] M. Masood and T. Y. Al-Naffouri, “Sparse Reconstruction Using Distribution Agnostic Bayesian Matching Pursuit,” *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5298–5309, 2013.

[15] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[16] D. Needell and J. A. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” *Appl. Computat. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, 2009.

[17] B. Wu, S. Cheng, and H. Wang, “Iterative channel estimation and signal detection in clipped OFDM,” in *Proc. IEEE GLOBECOM*, 2005.

[18] D. Kivanc, G. Li, and H. Liu, “Computationally efficient bandwidth allocation and power control for OFDMA,” *IEEE Trans. Wireless Commun.*, vol. 2, no. 6, pp. 1150–1158, 2003.

[19] E. Candes and J. Romberg, “Sparsity and incoherence in compressive sampling,” *Inverse prob.*, vol. 23, no. 3, p. 969, 2007.

[20] E. J. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[21] T. Blumensath and M. E. Davies, “Gradient pursuits,” *IEEE Trans. Signal Process*, vol. 56, no. 6, pp. 2370–2382, 2008.

[22] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *J. Roy. Stat. Soc*, pp. 267–288, 1996.

[23] X. Cai and G. B. Giannakis, “Error probability minimizing pilots for OFDM with M-PSK modulation over Rayleigh-fading channels,” *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 146–155, 2004.

[24] K. Z. Islam, T. Y. Al-Naffouri, and N. Al-Dhahir, “On optimum pilot design for comb-type OFDM transmission over doubly-selective channels,” *IEEE Trans. Commun.*, vol. 59, no. 4, pp. 930–935, 2011.

[25] T. Y. Al-Naffouri, K. Z. Islam, N. Al-Dhahir, and S. Lu, “A model reduction approach for OFDM channel estimation under high mobility conditions,” *IEEE Trans. Signal Process*, vol. 58, no. 4, pp. 2181–2193, 2010.

[26] T. Y. Al-Naffouri, A. Bahai, and A. Paulraj, “Semi-blind channel identification and equalization in OFDM: An expectation-maximization approach,” in *Proc. IEEE Vehicular Tech. Conf.*, vol. 1, 2002, pp. 13–17.

[27] T. Y. Al-Naffouri, O. Awoniyi, O. Oteri, and A. Paulraj, “Receiver design for MIMO-OFDM transmission over time variant channels,” in *Proc. IEEE Glob. Telecom. Conf.*, vol. 4, 2004, pp. 2487–2492.

[28] T. Y. Al-Naffouri and A. A. Quadeer, “A forward-backward Kalman filter-based STBC MIMO OFDM receiver,” *EURASIP J. Adv.Signal Process.*, vol. 2008.

[29] A. H. Sayed, *Fundamentals of adaptive filtering*, 2003: Wiley.

[30] E. B. Al-Safadi, T. Y. Al-Naffouri, M. Masood, and A. Ali, “Nonlinear Distortion Reduction in OFDM from Reliable Perturbations in Data Carriers,” *submitted to IEEE Tans. Commun.*