Compressed Sensing Techniques for Receiver based Post-Compensation of Transmitter's Nonlinear Distortions in OFDM Systems[☆]

Damilola S. Owodunni[a], Anum Ali[a], Ahmed A. Quadeer[b], Ebrahim B. Al-Safadi[c], Oualid Hammi[a,*], Tareq Y. Al-Naffouri[a,d]

[a] *King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia.*
[b] *The Hong Kong University of Science & Technology, Clear Water Bay, Hong Kong.*
[c] *University of Southern California, Los Angeles, CA.*
[d] *King Abdullah University of Science & Technology, Thuwal, Saudi Arabia.*

Abstract

In this paper, compressed sensing techniques are proposed to linearize commercial power amplifiers driven by orthogonal frequency division multiplexing signals. The nonlinear distortion is considered as a sparse phenomenon in the time-domain, and three compressed sensing based algorithms are presented to estimate and compensate for these distortions at the receiver using a few and, at times, even no frequency-domain free carriers (i.e. pilot carriers). The first technique is a conventional compressed sensing approach, while the second incorporates a priori information about the distortions to enhance the estimation. Finally, the third technique involves an iterative data-aided algorithm that does not require any pilot carriers and hence allows the system to work at maximum bandwidth efficiency. The performances of all the proposed techniques are evaluated on a commercial power amplifier and compared. The error vector

[☆]This work was supported by King Abdulaziz City for Science and Technology (KACST) through the Science & Technology Unit at King Fahd University of Petroleum & Minerals through Project No. 11-ELE1651-04 as part of the National Science, Technology and Innovation Plan.

[*]Corresponding author. Tel.: +966-3-860-7394
Email addresses: dksadiq@kfupm.edu.sa (Damilola S. Owodunni), anumali@ieee.org (Anum Ali), aaquadeer@ust.hk (Ahmed A. Quadeer), alsafadi@usc.edu (Ebrahim B. Al-Safadi), ohammi@ieee.org (Oualid Hammi), tareq.alnaffouri@kaust.edu.sa (Tareq Y. Al-Naffouri)

magnitude and symbol error rate results show the ability of compressed sensing to compensate for the amplifier's nonlinear distortions.
Keywords: compressed sensing, data-aided estimation, nonlinear distortion, orthogonal frequency division multiplexing, power amplifier.

1. Introduction

Emerging communication systems intensively use orthogonal frequency division multiplexing (OFDM) due to its numerous advantages such as high spectral efficiency, robustness to frequency selective fading, etc, which make it very attractive for the majority of communication systems. Due to its wide applicability, it remains one of the most explored topics in communication engineering [1–3]. However, OFDM signals often result in time-domain waveforms that have a high peak to average power ratio (PAPR) of up to 10 dB. These amplitude modulated signals are sensitive to the nonlinear distortions caused by the radio frequency (RF) power amplifier (PA) of the RF front-end. Indeed, the PA needs to linearly amplify the amplitude-modulated signals to avoid a high error vector magnitude (EVM) and symbol error rate (SER) which will translate into a loss of information. Simultaneously, the power efficiency of the PA needs to be maximized since the amplifier consumes most of the power in the RF front-end. However, power amplifiers have a low power efficiency when they are operated in their linear region, and their efficiency increases as they are driven into the nonlinear region close to saturation. Practically, power amplifiers are operated in their nonlinear region for power efficiency considerations. Then, the linearity is restored by means of system level architectures and mainly linearization techniques such as digital predistortion and feedforward implemented at the transmitter side [4–8].

Linearization techniques have been widely used to compensate for the PA's nonlinear distortions at the transmitter side. Among the various linearization

techniques, digital predistortion is commonly used. It consists of applying a complementary nonlinearity (predistorter) before the nonlinear PA such that the cascade of the predistorter and the amplifier behaves as a linear amplification system [4]. Such a technique requires an accurate knowledge of the amplifier's nonlinearity and thus its implementation calls for the use of extra circuitry which makes it unattractive for low power transmitters (up to a few watts) where power consumption and computational complexity are primary concerns.

Besides predistortion, there has been quite a number of other techniques proposed to combat the PA distortions at the transmitter through crest factor reduction, including, coding schemes [9–11], companding transforms [12, 13], tone reservation [14, 15], and constellation expansion [16, 17]. However, like predistortion, these methods add a varying degree of complexity to the transmitter which makes them inappropriate for applications where the transmitter needs to be kept quite simple.

One of the suggested solutions, clipping [18], is of relatively low complexity but introduces clipping noise that leads to performance degradation. Accordingly, some receiver based techniques to estimate and correct these distortions have been proposed [19–22]. Further, some receiver based techniques to mitigate undeliberate clipping caused by analog-to-digital converters were also proposed recently [23, 24]. In this work, a commercial PA is considered to be the source of nonlinear distortion in contrast to the self introduced clipping. Thus, not only the amplitude of the resulting signal is considered distorted but also its phase. As will be shown later in this paper, the power amplifier's nonlinear distortions can be considered as a sparse phenomenon in the time-domain. Hence, in this case, the sparsity of the resulting clips is not under the system designer's control, as was the case in previous correspondences on receiver based clipping estimation.

Recently, there has been an increased interest in the recovery of sparse signals using compressed sensing (CS) [26–28]. The significance of CS lies in the fact that it can reconstruct a sparse signal by utilizing a few linear projections over a basis that is incoherent with the basis in which the signal is sparse [29–32]. Thus, CS can be applied to recover and then compensate for these distortions using a few frequency-domain data-free or pilot carriers. The use of a data-aided technique along with CS can further improve the bandwidth efficiency by alleviating the need for frequency-domain free carriers. In such a case, the amplifier's distortions can be mitigated without using any frequency-domain free carriers. In this work, a CS based approach is used to develop a novel linearization technique that compensates for the power amplifier's nonlinear distortions. It is worth mentioning that, in [22], a CS based technique was employed at the receiver to compensate for the clipping of the transmitted signal that occurred at the transmitter. The knowledge of the clipping level was used in estimating the clips at the receiver. In contrast, the present work reports a blind nonlinearity estimation and cancellation technique which does not require any knowledge of the transmitter's nonlinearity parameters. Furthermore, both CS and weighted CS algorithms are used to estimate and compensate for the PA nonlinearities. Also, a novel approach using data-aided CS is proposed. This technique is shown to outperform conventional CS approaches as it can recover the nonlinear distortions with no need for free carriers. This will circumvent the bandwidth limitation of conventional CS techniques that require free carriers in order to estimate the amplifier's distortions. The employment of CS at the receiver, distinguishes the proposed approach from the work that attempts to linearize the PA at the receiver using an iterative procedure requiring the estimate of the nonlinear model characterizing the PA [25].

In this paper, measurement derived model of a commercial handset power

amplifier is used to validate the proposed CS based distortion mitigation technique at the receiver. It is worth mentioning that the nonlinear operation of the PA also causes out-of band emissions (also known as spectrum regrowth). This power leakage is a concern for transmitters which are required to abide to spectrum emission mask requirements. However, this work focuses primarily on the compensation of in-band distortions. Further, as this work solely focuses on the effects of amplifier's distortions, the additional impairments and imperfections of the transmitter/receiver such as carrier frequency offset (CFO) or timing phase offset (TPO) are not considered, as they can be easily estimated and compensated for using preamble signals.

1.1. Paper Organization

This paper is organized as follows. Section 2 describes the device under test (i.e. the PA) and the experimental setup used to carry out the measurements in characterizing the PA. In Section 3, the transceiver model is presented. The CS-based algorithm developed for the post-compensation of the PA's nonlinear distortions is proposed in Section 4. An enhanced data-aided CS algorithm is reported in Section 5 to further improve the linearization performance and bandwidth efficiency. The problem of channel estimation in presence of amplifier's nonlinear distortions is addressed in Section 6. A comparison between the various proposed algorithms appears in Section 7. Conclusions are presented in Section 8 and finally an appendix comparing CS with another sparse recovery algorithm comes in Section 9.

1.2. Notation

Scalars are represented by lower-case letters (e.g. x), vectors by bold-face lower-case letters (e.g., $\mathbf{x}$), and matrices by bold-face upper-case letters (e.g. $\mathbf{F}$). The i^{th} entry of a vector $\mathbf{x}$ is denoted by $\mathrm{x}(i)$ while a hat over a variable (e.g., $\hat{\mathbf{x}}$) is reserved for its estimate. The frequency-domain variables are represented in

upper-case calligraphic font (e.g. $\mathcal{X}$) to differentiate them from the time-domain ones. The transpose and complex conjugate transpose are represented by T and H (e.g. $\mathbf{x}^{\mathsf{T}}$ and $\mathbf{x}^{\mathsf{H}}$), respectively.

2. Experimental Setup and Device Under Test

The device under test (DUT) used in this work is a commercial PA from RF Micro Devices, Greensboro, NC. The amplifier is designed for handset applications in the 1920MHz to 1980MHz frequency band. First, the AM/AM and AM/PM characteristics of the DUT were measured using the experimental setup presented in Figure 1. A 5MHz wide OFDM based input signal waveform is downloaded into the signal generator that feeds the PA with the corresponding RF signal centered around the 1950MHz carrier frequency. The output of the DUT is then attenuated, and demodulated using a vector signal analyzer. The baseband complex input and output waveforms are then used to derive a behavioral model of the device under test. Since the DUT exhibits a memoryless behavior, a look-up table (LUT) behavioral model is adopted. Indeed, the proposed technique is intended for handset PAs which commonly exhibit a memoryless behavior due to the bandwidth of the signals being transmitted. The LUT model is derived from the measured AM/AM and AM/PM characteristics using the exponential weighted moving average algorithm proposed in [33]. The LUT model structure and its equations are described in [4]. This model will be used for the simulations that will be discussed in the next sections. The measured AM/AM and AM/PM characteristics as well as the model characteristics are reported in Figure 2. The low dispersion of the AM/AM and AM/PM characteristics confirms the memoryless behavior of the DUT and the suitability of the LUT behavioral model.

3. Transceiver Model

In this section, the details of the transmitter, the PA, and the receiver models used in the paper are presented.

3.1. Transmitter

In OFDM, the serial stream of data $\mathbf{d}$ to be transmitted is divided into N parallel streams that are modulated using either phase-shift keying (PSK) or quadrature amplitude modulation (QAM) to obtain a set of N data symbols, $\boldsymbol{\mathcal{X}} = [\mathcal{X}(0)\, \mathcal{X}(1) \cdots \mathcal{X}(N-1)]$. The time-domain signal that serves as an input to the PA is obtained by performing an inverse discrete Fourier transform (IDFT) operation on $\boldsymbol{\mathcal{X}}$

$$\mathbf{x} = \mathbf{F}^{\mathsf{H}} \boldsymbol{\mathcal{X}} \tag{1}$$

where $\mathbf{F}$ denotes the unitary discrete Fourier transform (DFT) matrix with $(a, b)^{th}$ element

$$F_{a,b} = \frac{1}{\sqrt{N}} e^{-j2\pi ab/N}, \quad a, b \in \{0, 1, \cdots, N-1\},$$

Furthermore in OFDM systems, a cyclic prefix is appended to $\mathbf{x}$ to avoid inter-symbol interference. This signal then passes through the PA before transmission.

3.2. Power Amplifier

The model used for the PA has been described in Section 2.

Based on Bussgang's theorem [34], a generic formula relating the power amplifier's output signal to its input is given by:

$$\mathbf{x}_p = \alpha \mathbf{x} + \mathbf{x}_c \tag{2}$$

where α is a constant amplification value that corresponds to the PA's small signal gain (which can be obtained from the measured AM/AM and AM/PM

characteristics of the amplifier) and $\mathbf{x}_c$ represents the distortion added to the amplified signal. The input-output relationship presented in equation (2) is applicable to any amplifier and can be easily derived from the standard equation typically used to describe the LUT model. However, the generic relation presented by equation (2) was preferred since it is more suitable for the calculations required by the compressive sensing technique.

Figure 3(a) shows a sample snapshot of the amplifier's scaled input, $\alpha\mathbf{x}$, and output, $\mathbf{x}_p$, while the distortion, $\mathbf{x}_c = \mathbf{x}_p - \alpha\mathbf{x}$ is shown in Figure 3(b). From both figures, the largest and hence, most problematic distortions are observed at instances when the input signal is of high amplitude but some small nonlinearities also appear even at low amplitudes of the input signal thus making the added distortion signal, $\mathbf{x}_c$ only approximately sparse — or, in proper terms, compressible. This is one difference between the simple clipping model and the more general PA distortion. In the former, hard-clipping is involved and the signal is clipped at a particular threshold value such that the resulting distortion (i.e. the clipped signal) is sparse (i.e. only relatively very few non-zero values occur at the clipping locations). Due to this, the location of the clipping could be estimated, with high accuracy, from the clipped signal. However, in the case of amplifier distortions, nonlinearities of varying degrees are introduced at all signal levels due the amplitude-variant distortion caused by the amplifier itself. The work in [22] shows that CS performs well for the simpler case of hard-clipping. Looking at the amplifier characteristics plotted in Figure 2, it is expected that only the highest amplitudes of input signals will be severely distorted.

A block diagram showing fundamental building blocks of the transmitter is presented in Figure 4. Free Carriers are inserted and data is modulated before performing the IDFT operation, which results in the time domain signal $\mathbf{x}$. This

time domain signal is then passed through the PA to obtain the signal ready for transmission.

3.3. Receiver

The time-domain received signal is modeled as

$$\mathbf{y}_{\mathrm{h}} = \mathbf{H}\mathbf{x}_p + \mathbf{z}_{\mathrm{h}} \tag{3}$$

where $\mathbf{y}_{\mathrm{h}} \in \mathbb{C}^N$ is the time-domain received OFDM symbol (after removing the cyclic prefix) and $\mathbf{z}_{\mathrm{h}}$ is the circular complex additive white Gaussian noise (AWGN), $\mathbf{z}_{\mathrm{h}} \sim \mathcal{CN}(\mathbf{0}, \sigma^2_{z_h}\mathbf{I})$, where $\sigma^2_{z_h}$ is the variance of noise samples. In OFDM systems, the linear convolution between the transmitted data, $\mathbf{x}_p$, and the channel impulse response (CIR), $\mathbf{h} = [\mathrm{h}(0), \mathrm{h}(1), \cdots, \mathrm{h}(L-1)]^{\mathsf{T}}$, is converted into a cyclic convolution due to the presence of the cyclic prefix. The cyclic prefix length is assumed to be greater than L to avoid inter-symbol interference. Thus, $\mathbf{H}$ denotes the circulant channel matrix in (3) that can be decomposed as, $\mathbf{H} = \mathbf{F}^{\mathsf{H}}\mathbf{\Lambda}\mathbf{F}$,where $\mathbf{\Lambda} = \mathrm{diag}(\mathcal{H})$, and $\mathcal{H} = \sqrt{N}\mathbf{F}\mathbf{h}$ is the DFT of the CIR. Using this fact and taking the DFT of both sides of (3) yields

$$\mathcal{Y}_{\mathrm{h}} = \mathbf{\Lambda}\mathbf{F}\mathbf{x}_p + \mathcal{Z}_{\mathrm{h}} \tag{4}$$

where $\mathcal{Y}_{\mathrm{h}}$ and $\mathcal{Z}_{\mathrm{h}}$ are the DFT's of $\mathbf{y}_{\mathrm{h}}$ and $\mathbf{z}_{\mathrm{h}}$ respectively.

Let us focus on the development of the proposed scheme with the assumption that the CIR is perfectly known at the receiver. How this information can be obtained is addressed in Section 6. The frequency-domain received signal (after equalization) is given by

$$\mathcal{Y} = \mathbf{F}\mathbf{x}_p + \mathcal{Z} \tag{5}$$

where $\mathcal{Y} = \mathbf{\Lambda}^{-1}\mathcal{Y}_{\mathrm{h}}$ and $\mathcal{Z} = \mathbf{\Lambda}^{-1}\mathcal{Z}_{\mathrm{h}}$. It is worth mentioning that the minimum mean square error (MMSE) equalization may yield better results than zero forcing (ZF) or least square (LS) but in this work LS equalizer is implemented.

Substituting the value of $\mathbf{x}_p$ from (2) in (5) yields

$$\mathcal{Y} = \alpha\mathbf{F}\mathbf{x} + \mathbf{F}\mathbf{x}_c + \mathcal{Z} \tag{6}$$

The time-domain equivalent of the received signal can thus be written as

$$\mathbf{y} = \alpha\mathbf{x} + \mathbf{x}_c + \mathbf{z} \tag{7}$$

where $\mathbf{y}$ is the IDFT of $\mathcal{Y}$ and $\mathbf{z}$ is the IDFT of $\mathcal{Z}$ and has the same distribution of $\mathcal{Z}$ since $\mathbf{F}$ is unitary.

4. Post-Compensation using Compressed Sensing

In this paper, the PA nonlinear distortions, $\mathbf{x}_c$, are estimated using CS based techniques by exploiting the free carriers inserted in the OFDM symbol. Let $\boldsymbol{\omega}$ of cardinality $|\boldsymbol{\omega}| = N$ be the set of all carriers available in the OFDM symbol and $\boldsymbol{\omega}_P \subset \boldsymbol{\omega}$ of cardinality $|\boldsymbol{\omega}_P| = P$ with $P < N$ denote the set of free or pilot carriers that will be used to estimate $\mathbf{x}_c$. As we use CS-based techniques to estimate $\mathbf{x}_c$, it is desirable for the P free carriers to be randomly placed [29] and known to the receiver. Let D be the number of active tones used for data transmission with $D = N - P$ and define $\mathbf{S}_D$ as a binary selection matrix (of size $N \times D$) with only one non-zero element equal to 1 per row and column that selects the data carriers and all zero rows with indices belonging to $\boldsymbol{\omega}_P$. Then, the time-domain OFDM signal in (1) can be re-defined as

$$\mathbf{x} = \mathbf{F}^{\mathsf{H}}\mathbf{S}_D\mathcal{X}_D \tag{8}$$

where $\mathcal{X}_D$ is the $D \times 1$ frequency-domain modulated data vector. The frequency-domain received signal (6) is thus modified as

$$\mathcal{Y} = \alpha\mathbf{S}_D\mathcal{X}_D + \mathbf{F}\mathbf{x}_c + \mathcal{Z} \tag{9}$$

Let us denote by $\mathbf{S}_P$ the selection matrix (of size $N \times P$) that spans the orthogonal complement of the columns of $\mathbf{S}_D$ (i.e. $\mathbf{S}_P$ is a binary matrix of size $N \times P$ with only one non-zero element equal to 1 per row and column and all zero rows with indices belonging to $(\boldsymbol{\omega} - \boldsymbol{\omega}_P)$. The distortion $\mathbf{x}_c$ is estimated by projecting $\mathcal{Y}$ on $\mathbf{S}_P^\mathsf{T}$ as follows

$$\mathcal{Y}_P = \mathbf{S}_P^\mathsf{T}\mathcal{Y} = \mathbf{S}_P^\mathsf{T}\mathbf{F}\mathbf{x}_c + \mathcal{Z}_P \tag{10}$$

where $\mathcal{Z}_P = \mathbf{S}_P^\mathsf{T}\mathcal{Z}$ is a Gaussian vector of length P. For notational convenience, we re-write the above equation as

$$\mathcal{Y}_P = \boldsymbol{\Psi}_P\mathbf{x}_c + \mathcal{Z}_P \tag{11}$$

where $\boldsymbol{\Psi}_P \triangleq \mathbf{S}_P^\mathsf{T}\mathbf{F}$ is a measurement matrix of size $P \times N$.

4.1. Compressive Sensing (CS)

Note that (11) forms an under-determined system of linear equations as $\mathbf{x}_c \in \mathbb{C}^N$ and $\mathcal{Y}_P \in \mathbb{C}^P$ with $P < N$ and hence cannot be solved by using the conventional linear techniques. This is in fact a typical CS problem when it is known a priori that the signal of interest $\mathbf{x}_c$ is sparse [29–32]. This problem can be solved by using the convex relaxation approach that solves an ℓ_1-norm minimization problem using linear programming [29]. Following the notation used in the paper, the problem can be casted as

$$\begin{aligned} &\text{minimize} \quad \|\mathbf{x}_c\|_1, \\ &\text{subject to} \quad \|\mathcal{Y}_P - \boldsymbol{\Psi}_P\mathbf{x}_c\|_2 \leq \epsilon \end{aligned} \tag{12}$$

where $\epsilon = \sqrt{\sigma_z^2(P + \sqrt{2P})}$ [35]. It is important to mention here that the above convex relaxation approach is used to estimate $\mathbf{x}_c$ from (11) in this paper but any other CS-based technique (for example, Bayesian methods [36–38], and matching pursuits [39–41]) can be utilized.

After obtaining an estimate of the distortion, $\hat{\mathbf{x}}_c$, using CS, an estimate of the distortion-free signal can be obtained as follows

$$\alpha\hat{\mathbf{x}} = \mathbf{y} - \hat{\mathbf{x}}_c \tag{13}$$

The signal, $\alpha\hat{\mathbf{x}}$ is then scaled down by a factor of α and then transformed by a DFT operation to the frequency-domain data signal, $\hat{\mathfrak{X}}$. Finally, this is demodulated to obtain an estimate of the transmitted data, $\hat{\mathbf{d}}$. A block diagram of a typical receiver based on CS for nonlinear distortion estimation is shown in Figure 5.

4.2. Weighted CS (WCS)

If one has some a priori information related to the sparse signal $\mathbf{x}_c$, an alternative approach to (12) can be pursued by penalizing the less probable locations of $\mathbf{x}_c$ as follows [42]

$$\begin{aligned} &\text{minimize} \quad \|\mathbf{w}^{\mathsf{T}}\mathbf{x}_c\|_1, \\ &\text{subject to} \quad \|\mathcal{Y}_P - \mathbf{\Psi}_P\mathbf{x}_c\|_2 \leq \epsilon \end{aligned} \tag{14}$$

where $\mathbf{w}$ is a vector consisting of weights for each location in $\mathbf{x}_c$.

As discussed in Section 3.2, the nonlinearity of any PA increases with an increase in amplitudes of its input signal. Thus the major distortions caused by the PA occur at the locations where the input amplitude is large. Accordingly, we can define $\mathbf{w}$ to be the inverse of the magnitude of the received signal $\mathbf{y}$, i.e.

$$\mathrm{w}(n) = \begin{cases} \frac{1}{|\mathrm{y}(n)|}, & \mathrm{y}(n) \neq 0, \\ \infty, & \mathrm{y}(n) = 0. \end{cases} \tag{15}$$

where $n = 1, 2, ..., N$. This way, the small entries in $\mathbf{w}$ correspond to the most probable locations where the PA might have distorted the signal and thus, this forces (14) to concentrate on them.

4.3. Simulation Setup and Results

In all the simulations presented in the paper, the number of subcarriers is fixed at $N = 256$ and a 64QAM modulation scheme is employed. Unless otherwise stated, the number of free carriers P is set to 10% of N. The PA is represented by its measured LUT based model as described in Section 2.

The following two performance measures are used for comparing the proposed techniques:

- Error Vector Magnitude (EVM) [43, 44]:

$$\text{EVM} = \sqrt{\frac{\frac{1}{N}\sum_{r=1}^{N}\left|\mathcal{X}(r) - \hat{\mathcal{X}}(r)\right|^2}{\frac{1}{N}\sum_{r=1}^{N}\left|\mathcal{X}(r)\right|^2}} \tag{16}$$

  where $\mathcal{X}(r)$ and $\hat{\mathcal{X}}(r)$ are the original and estimated data symbols respectively.

- Symbol Error Rate (SER):

$$\text{SER} = \frac{\sum \textit{symbol errors (comparing } \mathbf{d} \textit{ and } \hat{\mathbf{d}}\textit{)}}{\textit{total number of symbols in } \mathbf{d}} \tag{17}$$

Both performance measures are plotted as functions of the SNR ranging from 15 dB to 35 dB. The SNR is given by:

$$\text{SNR} = \frac{\sigma_z^2}{\sigma_{\mathbf{x}_p}^2} \tag{18}$$

where σ_z^2 and $\sigma_{\mathbf{x}_p}^2$ are the variances of the noise and PA output signal, respectively. In all cases, the performance is upper bounded by the case when no nonlinear distortion estimation procedure is used at the receiver and is lower bounded by the case when the PA is assumed to be perfectly linear, i.e. when $\mathbf{x}_p = \alpha\mathbf{x}$. It is worth mentioning that, the performance in the case of linear PA is limited by the presence of the AWGN noise.

The EVM and SER performances of CS (12) and WCS (14) are compared in Figure 6 for free carriers ranging from 10% to 20% of N. As expected, it can be observed that increasing the number of free carriers enhances the performance of both CS and WCS. It can also be easily seen that WCS outperforms CS for the same number of free carriers. In fact, the performance of WCS with 15% free carriers is even better than CS with 20% free carriers. This shows the advantage of incorporating the a priori information of the power amplifier's behavior in estimating the nonlinear distortions[1].

5. Data-Aided CS Algorithm

One disadvantage of the previous algorithm is that a few carriers need to be reserved and be used for estimating the distortion. This causes a reduction in the available bandwidth. In this section, a data-aided algorithm is presented. This algorithm utilizes reliable data to aid in CS estimation. The advantages of the proposed iterative approach, labelled "Data-Aided CS (DACS) algorithm", are twofold:

- It enhances the performance of the CS/WCS algorithms (presented in the previous section).
- It helps increase the bandwidth efficiency of the system (by reducing the number of free or pilot carriers required) with a nominal increase in the receiver complexity.

5.1. Algorithm Description

This algorithm is based on the assumption that even after the nonlinear distortions caused by the PA, a part of the data samples still remains within its corresponding decision regions. Let $\boldsymbol{\omega}_R \subset \boldsymbol{\omega}$ of cardinality $|\boldsymbol{\omega}_R| = R$ denote the

[1]The advantage of using CS algorithm (especially in the case of WCS) over other sparse signal reconstruction schemes is highlighted in Appendix.

set of these carriers in which the perturbations are not severe i.e. the carriers are reliable. In other words, the noisy and perturbed data samples would remain in the decision regions of their respective constellation points, so that the following would hold with high probability

$$\hat{\mathfrak{X}}_R = \mathfrak{X}_R \tag{19}$$

where $\hat{\mathfrak{X}}_R$ is the estimated data at the reliable carriers. The equation of the received signal given in (6) can be written as

$$\mathcal{Y} = \alpha\mathfrak{X} + \mathbf{F}\mathbf{x}_c + \mathcal{Z} \tag{20}$$

where $\mathfrak{X}$ is the FFT of $\mathbf{x}$.

Let $\mathbf{S}_R$ be a binary selection matrix (of size $N \times R$) with only one non-zero element equal to 1 per column that selects the reliable carriers. Multiplying both sides of (20) by $\mathbf{S}_R^\mathsf{T}$ yields

$$\mathbf{S}_R^\mathsf{T}\mathcal{Y} = \alpha\mathbf{S}_R^\mathsf{T}\mathfrak{X} + \mathbf{S}_R^\mathsf{T}\mathbf{F}\mathbf{x}_c + \mathbf{S}_R^\mathsf{T}\mathcal{Z} \tag{21}$$

which, following the convention used in equations (10) and (11), can be written as

$$\mathcal{Y}_R = \alpha\mathfrak{X}_R + \mathbf{\Psi}_R\mathbf{x}_c + \mathcal{Z}_R \tag{22}$$

where $\mathcal{Y}_R = \mathbf{S}_R^\mathsf{T}\mathcal{Y}$, $\mathfrak{X}_R = \mathbf{S}_R^\mathsf{T}\mathfrak{X}$, and $\mathcal{Z}_R = \mathbf{S}_R^\mathsf{T}\mathcal{Z}$.

The perturbations, $\mathbf{\Psi}_R\mathbf{x}_c + \mathcal{Z}_R$, at the reliable carriers do not push the data outside the reliable regions. In other words, if we divide $\mathcal{Y}_R$ by α and round to the nearest decision region, we get $\mathfrak{X}_R$, i.e.

$$\begin{aligned} \left\lfloor \frac{1}{\alpha}\mathcal{Y}_R \right\rceil &= \mathfrak{X}_R \\ \text{or } \alpha\left\lfloor \frac{1}{\alpha}\mathcal{Y}_R \right\rceil &= \alpha\mathfrak{X}_R \end{aligned} \tag{23}$$

where the $\lfloor\cdot\rceil$ operator denotes rounding to the nearest neighbor. Thus, we can

write (22) as

$$\mathcal{Y}_R - \alpha\mathcal{X}_R = \mathbf{\Psi}_R\mathbf{x}_c + \mathcal{Z}_R$$

or

$$\mathcal{U}_R = \mathbf{\Psi}_R\mathbf{x}_c + \mathcal{Z}_R \tag{24}$$

where $\mathcal{U}_R = \mathcal{Y}_R - \alpha\lfloor\frac{1}{\alpha}\mathcal{Y}_R\rfloor$. It is important to note that it is not needed to determine all reliable carriers, $\boldsymbol{\omega}_R$, rather, it is sufficient to determine a subset of these carriers, $\boldsymbol{\omega}_{R'} \subset \boldsymbol{\omega}_R$ and use them (Here onwards, R' is used to distinguish the variables corresponding to the subset $\boldsymbol{\omega}_{R'}$, from the variables corresponding to the set $\boldsymbol{\omega}_R$). The system of equations (24) can be solved using a CS-based approach similar to (12) as follows

$$\begin{aligned} &\text{minimize} \quad \|\mathbf{x}_c\|_1, \\ &\text{subject to} \quad \|\mathcal{U}_{R'} - \mathbf{\Psi}_{R'}\mathbf{x}_c\|_2 \leq \epsilon \end{aligned} \tag{25}$$

The above procedure can be repeated $j_{\max}$ times to further enhance the performance as follows.

1. Let $\hat{\mathcal{X}}^j$ be the j^{th} estimated modulated data vector (obtained by taking the DFT of $\hat{\mathbf{x}}$).
2. Compare $\hat{\mathcal{X}}^j$ with the P size constellation points and obtain the R' reliable carriers.
3. Find $\mathbf{S}_{R'}$, $\mathcal{U}_{R'}$ and $\mathbf{\Psi}_{R'}$ based on R'.
4. Evaluate (25) to obtain $\hat{\mathbf{x}}_c^j$.
5. Obtain $\hat{\mathbf{x}}^{j+1} = \hat{\mathbf{x}}^j - \hat{\mathbf{x}}_c^j$.
6. Repeat step 2-5 till $j = j_{\max}$.

Figure 7 illustrates the receiver design based on DACS algorithm. Note that the above procedure can also be applied to the case when no free carriers are

used for CS estimation in the first iteration. In this case, the algorithm will highly rely on the set of reliable carriers available.

5.2. Constructing $\boldsymbol{\omega}_{R'}$

In the following, we pursue a geometrical approach to select the most reliable set of carriers from the observed data.

In order to explain the adopted approach, we consider as a motivating example the constellation shown in Figure 8. Here $\hat{\mathcal{X}}_1$ and $\hat{\mathcal{X}}_2$ are two equalized data samples which are equidistant from the closest constellation point, $\mathcal{X}$. However, in spite of being equidistant from $\mathcal{X}$, $\hat{\mathcal{X}}_1$ and $\hat{\mathcal{X}}_2$ have different reliability values. This is because the distances of these two points from their respective next nearest neighbors are different. Specifically, note that $\mathcal{X}_a$ is next nearest neighbor of $\hat{\mathcal{X}}_1$ and $\mathcal{X}_c$ is next nearest neighbor of $\hat{\mathcal{X}}_2$ respectively. Note also that, given that $\hat{\mathcal{X}}_1$ and $\hat{\mathcal{X}}_2$ are equidistant from $\mathcal{X}$, it is clear that $\hat{\mathcal{X}}_2$ is more reliable than $\hat{\mathcal{X}}_1$ and in relative terms we have

$$\frac{|\hat{\mathcal{X}}_2 - \mathcal{X}|}{|\hat{\mathcal{X}}_2 - \mathcal{X}_c|} < \frac{|\hat{\mathcal{X}}_1 - \mathcal{X}|}{|\hat{\mathcal{X}}_1 - \mathcal{X}_a|} \tag{26}$$

This motivates the following reliability matrix $\mathfrak{R}(n)$,

$$\mathfrak{R}(n) = -\log\left(\frac{|\hat{\mathcal{X}} - \lfloor\hat{\mathcal{X}}\rfloor|}{|\hat{\mathcal{X}} - \lfloor\hat{\mathcal{X}}\rfloor_{NN}|}\right) \tag{27}$$

where, as defined before, $\lfloor\hat{\mathcal{X}}\rfloor$ denotes rounding to the nearest constellation point while $\lfloor\hat{\mathcal{X}}\rfloor_{NN}$ denotes rounding to the next nearest constellation point.

Thus, it is possible to calculate the reliability of all $N - P$ carriers (or N carriers in the case when no free/pilot carriers are used), sort the reliabilities in descending order $\mathfrak{R}(n_1) \geq \mathfrak{R}(n_2) \geq \cdots \geq \mathfrak{R}(n_{N-P})$ and choose the R' carriers with the highest reliability $\boldsymbol{\omega}_{R'} = \{n_1, n_2, \ldots, n_R\}$.

5.3. Simulation results

Figure 9 compares the performance of DACS algorithm with WCS for varying number of free carriers, reliable carriers, and iterations. In the cases studied, whenever free carriers are reserved, WCS (14) instead of plain CS is used for distortion estimation before implementing the DACS algorithm. The number of reliable carriers R' for DACS is either 30% or 40% of $N - P$ and a maximum of $j = 2$ DACS iterations is used. Here also, the performance is lower bounded by the linear PA case, in which AWGN is the limiting factor. It can be observed that DACS with 2 iterations easily outperforms WCS. Moreover, the performance of DACS with 40% reliable carriers is better than the one with 30% reliable carriers. It is important to note here that the performance can be improved further by using more reliable carriers at the expense of increased receiver complexity. The performance of the bandwidth efficient case (when no free carriers are used, i.e., $P = 0$) is also presented in Figure 9. Thus in this case, the estimate in the first iteration also depends on the credibility of the reliable carriers. It can be observed from the EVM plot that with 2 iterations, DACS with 40% reliable carriers outperforms both WCS and DACS (with 10% free carriers and 1 iteration). This shows the trade-off between bandwidth efficiency and computational complexity at the receiver end.

In Figure 10, the EVM performances of all the algorithms presented in the paper are compared as a function of the number of free carriers. The target EVM requirement is set to 8%. Indeed, this value is consistent with the standard requirements for the same modulation type (64QAM) in LTE standard [45].

The SNR is fixed at 25 dB (Note that all the algorithms meet the minimum EVM requirement if the SNR is increased to 30 dB). It can be observed that CS and WCS do not meet the minimum EVM requirement even with 20% free carriers. In Figure 10, there are no data points for 0 carriers for CS and WCS as

these two algorithms cannot work without free carriers. The trend of the plot indicates that they can meet the requirement if more free carriers are utilized. It also demonstrates that DACS algorithms with 40% reliable carriers meet the minimum EVM requirement even when no free carriers are used.

6. Channel Estimation in Presence of PA's Nonlinear Distortions

In OFDM communication systems, the channel can be estimated at the receiver with the help of pilot signals. As the equi-spaced formulation of the pilot carriers is found to be optimal, uniformly spaced pilot sub-carriers are chosen for channel estimation [46–48]. Based on transmitted pilot sequence LS or MMSE estimation is utilized to estimate the channel at the receiver [49]. However, the use of simple pilot based MMSE for channel estimation yields sub-optimal results when the transmitted signal contains nonlinear distortions. This is because the pilots used as a reference for LS or MMSE estimation are all corrupted by the nonlinear distortions.

The problem of channel estimation in the presence of deliberate clipping is addressed in [50]. The idea is to utilize the knowledge of clipping threshold at the receiver to reconstruct the clipped transmitted signal and use clipped pilots as a reference for MMSE to improve the channel estimation. However, doing so for the nonlinear distortion problem on hand will require the exact knowledge of the characteristics of the PA (i.e. the LUT of the PA) at the receiver, which is generally not available. In this situation the following strategy based on MMSE estimation using the reliable carriers can be used to improve channel estimation.

In the proposed channel estimation scheme, initially the MMSE channel estimate $\hat{\mathbf{h}}_{\text{mmse}}$ is obtained based on the transmitted pilot sequence. Ignoring the nonlinear distortions in equation (4), we can write the received signal as

$$\mathcal{Y}_{\text{h}} = \alpha \mathbf{\Lambda F x} + \mathcal{Z}_{\text{h}} \tag{28}$$

We can reduce this system of equation to the one containing only the pilot sub-carriers, and hence we write

$$\mathcal{Y}_{\mathrm{p,h}} = \alpha \mathbf{\Lambda}_{\mathrm{p}} \mathbf{F}_{\mathrm{p}} \mathbf{x}_{\mathrm{p}} + \mathcal{Z}_{\mathrm{p,h}} \tag{29}$$

Here onwards for notational simplicity, we drop the subscript h and write equation (31) simply as $\mathcal{Y}_{\mathrm{p}} = \mathbf{\Lambda}_{\mathrm{p}} \mathbf{F}_{\mathrm{p}} \mathbf{x}_{\mathrm{p}} + \mathcal{Z}_{\mathrm{p}}$. For MMSE estimation, we need to know the autocorrelation $R_{\mathcal{Y}_{\mathrm{p}}\mathcal{Y}_{\mathrm{p}}}$ and the cross-correlation $R_{\mathcal{Y}_{\mathrm{p}}\mathbf{h}}$. These can be found if we rewrite $\mathcal{Y}_{\mathrm{p}}$ as

$$\mathcal{Y}_{\mathrm{p}} = \mathbf{X}_{\mathrm{p}} \mathcal{H}_{\mathrm{p}} + \mathcal{Z}_{\mathrm{p}} \tag{30}$$

The last equation results when we replace the channel diagonal matrix $\mathbf{\Lambda}_{\mathrm{p}}$ with a column vector $\mathcal{H}_{\mathrm{p}}$ ($\mathbf{\Lambda}_{\mathrm{p}}$ contained $\mathcal{H}_{\mathrm{p}}$ on its diagonal). We also replaced the column vector $\mathbf{F}_{\mathrm{p}}\mathbf{x}_{\mathrm{p}}$ with a diagonal matrix $\mathbf{X}_{\mathrm{p}}$ ($\mathbf{X}_{\mathrm{p}}$ contains $\mathbf{F}_{\mathrm{p}}\mathbf{x}_{\mathrm{p}}$ on its diagonal). Further we can write $\mathcal{H}_{\mathrm{p}} = \sqrt{N}\mathbf{F}_{\mathrm{p}}\mathbf{h}$ to finally get

$$\mathcal{Y}_{\mathrm{p}} = \alpha\sqrt{N}\mathbf{X}_{\mathrm{p}}\mathbf{F}_{\mathrm{p}}\mathbf{h} + \mathcal{Z}_{\mathrm{p}} \tag{31}$$

Now the autocorrelation $R_{\mathcal{Y}_{\mathrm{p}}\mathcal{Y}_{\mathrm{p}}}$ is found to be

$$R_{\mathcal{Y}_{\mathrm{p}}\mathcal{Y}_{\mathrm{p}}} = \alpha^2 N \sigma_h^2 \mathbf{X}_{\mathrm{p}} \mathbf{F}_{\mathrm{p}} \mathbf{F}_{\mathrm{p}}^{\mathsf{H}} \mathbf{X}_{\mathrm{p}}^{\mathsf{H}} + \sigma_z^2 \mathbf{I} \tag{32}$$

here the noise $\mathbf{z}$ and the channel $\mathbf{h}$ are assumed uncorrelated. Further, all taps of the channel vector $\mathbf{h}$ are assumed independent and identically distributed (i.i.d) with variance σ_h^2. The cross correlation is found as

$$R_{\mathcal{Y}_{\mathrm{p}}\mathbf{h}} = \alpha\sqrt{N}\sigma_h^2 \mathbf{X}_{\mathrm{p}}\mathbf{F}_{\mathrm{p}} \tag{33}$$

Here, using the MMSE estimate, we can find the $\hat{\mathbf{h}}_{\mathrm{mmse}}$ as given below

$$\hat{\mathbf{h}}_{\mathrm{mmse}} = \alpha\sqrt{N} R_{\mathcal{Y}_{\mathrm{p}}\mathbf{h}} R^{-1}_{\mathcal{Y}_{\mathrm{p}}\mathcal{Y}_{\mathrm{p}}} \mathcal{Y}_{\mathrm{p}} \tag{34}$$

The received data $\mathcal{Y}_{\mathrm{h}}$ can now be equalized using $\hat{\mathbf{h}}_{\mathrm{mmse}}$, and reliable carriers

are found based on the procedure given in Section 5. The maximum likelihood decisions, as given in equation (23), on the reliable sub-carriers can now be used as additional pilots for MMSE estimation. Using the additional pilots will result in improved MMSE estimate of the channel which can again be utilized for equalization, and this procedure can be iterated. It is worth mentioning that the reliable carriers are playing a dual role, as they are used both for enhancing the channel estimation as well as enhancing the recovery of nonlinear distortions. The proposed channel estimation scheme is summarized below:

1. Obtain the MMSE channel estimate $\hat{\mathbf{h}}_{\text{mmse}}$ based on transmitted pilots as given in equation (34).
2. Equalize the received data using $\hat{\mathbf{h}}_{\text{mmse}}$.
3. Find reliable carriers by the procedure given in Section 5.
4. Obtain the improved MMSE estimate based on the transmitted pilots and the reliable data carriers as additional measurements.
5. Repeat 2-4.

Note that the proposed channel estimation scheme does not rely on iterative channel estimation and distortion mitigation. Rather, the improved channel estimation accuracy by the proposed scheme is expected due to the increased number of measurements by virtue of reliable carriers.

For the experiments, a 4-tap rayleigh fading channel is used, where all taps are complex i.i.d and come from a Gaussian distribution. To estimate the channel 16 equi-spaced sub-carriers are pre-allocated as pilot signals. The channel estimation results for the proposed scheme are presented in Figure 11. The mean square error (MSE) results for the simple MMSE and the MMSE aided by the reliable carriers are compared in Figure 11(a). It can be observed from the results that the proposed scheme improves the channel estimation accuracy by more than 3dB throughout the range of interest. Further the SER results

are shown in Figure 11(b). It can be observed that the first iteration of the proposed scheme significantly enhances the performance of the channel estimation in comparison with MMSE estimation based on the pilots. However, in the current scenario the channel estimation accuracy of the proposed scheme saturates on the second iteration and no further enhancement is achieved. Also, it is observed that the proposed distortion cancellation scheme based on CS works effectively even when the channel is estimated in the presence of inband nonlinear distortions. However, there is some loss in performance due to channel estimation error.

7. Algorithms Benchmarking

In Table 1, the performance expressed in terms of EVM values are compared, at an SNR of 35dB, for the various algorithms presented in the paper alongside the cost, in terms of execution time and bandwidth, of each. The EVM values were calculated as described in Section IV.B, while the bandwidth efficiencies and execution times of the algorithms were calculated with respect to those of the DACS with 40% reliable carriers.

In Table 1, it is worth mentioning that using WCS with 20% pilot carriers performs better than WCS with just 10% carriers. This is because the more free pilots are available, the more measurements will be available for use in the CS/WCS algorithm and thus, the better the estimate of the distortion becomes. However, this also implies a reduction in the bandwidth left for the actual data transmission and an increase in the time needed to run the algorithm as reported in Table I. To improve bandwidth efficiency, DACS was used, which, unlike WCS, does not require free carriers and thus allows a full use of the system bandwidth. But in order to achieve a reasonable performance improvement using DACS, about 40% of the data carriers needed to be used for measurements and this leads to an increase in the execution time as it can be noticed in rows 3

and 4 of Table 1. Also, it is important to note that the performance of the first iteration of the DACS algorithm (without a previous implementation of WCS) is relatively poor. This is a result of the low amplitude of the measurements in the first set of data carriers selected. However, the second iteration gives a significant performance improvement since the second set of data carriers used offer measurements of larger amplitudes. From a designer's point of view, it is important to highlight that looking at Table 1 and choosing the most suitable algorithm from the proposed schemes, depends on the limitations imposed by the system on hand and the design requirements. If bandwidth efficiency is of utmost importance, DACS with 2 iterations provides the best possible results. However, if computational costs are to be kept to the minimum WCS with 20% free pilots is a good choice.

8. Conclusion

In this paper, the nonlinear distortion caused by the PA, considered as one of the main problems in any OFDM-based system, has been addressed. As the nonlinear distortion is a sparse signal in time-domain, it has been shown that CS-based techniques can be employed to estimate it at the receiver. Specifically, three algorithms have been presented to mitigate the nonlinear distortions. The paper started with the basic CS approach that utilizes frequency-domain free carriers which was followed by weighted CS method that uses the a priori information according to which the high amplitude signals are more probable to be distorted by the PA. An iterative data-aided CS-based method that utilizes the reliable data carriers to improve the estimation process has also been proposed. It has been shown that this approach can provide good estimation of the nonlinear distortion without even wasting any bandwidth for free carriers. The results demonstrate the favorable performance of the proposed algorithms for commercial power amplifiers. The benchmarking of proposed schemes investi-

gated in this work, provides valuable guidelines to the designer on choosing the most suitable algorithm depending on requirement.

9. Appendix

In order to justify the use of CS instead of other sparse-recovery algorithms, we compare in this section the performance of WCS against that of the Fast Bayesian Matching Pursuit (FBMP) algorithm [38] using the EVM performance criterion.

As it can be seen in Figure 12, WCS performs better than FBMP at both high and low SNR values. As a matter of fact, FBMP does worse when the SNR is low.

Acknowledgment

The authors thank Agilent Technologies, Palo Alto, CA, for the donation of the Advanced Design System Software used in this work.

References

[1] S. Bin Hong and H.-M. Kim, "Pilot signal design algorithm for efficient symbol time offset estimation in an OFDM system," *Signal Processing*, vol. 87, no. 3, pp. 489-499, Mar. 2007.

[2] P. Salvo Rossi, R. R. Mller, and O. Edfors, "Linear MMSE estimation of timefrequency variant channels for MIMO-OFDM systems," *Signal Processing*, vol. 91, no. 5, pp. 1157-1167, May 2011.

[3] A. Bouzegzi, P. Ciblat, and P. Jallon, "New algorithms for blind recognition of OFDM based systems," *Signal Processing*, vol. 90, no. 3, pp. 900-913, Mar. 2010.

[4] F. M. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," *IEEE Microw. Mag.*, vol. 10, no. 7, pp. 52-64, Dec. 2009.

[5] O. Hammi, J. Sirois, S. Boumaiza, and F. M. Ghannouchi, "Design and performance analysis of mismatched Doherty amplifiers using accurate load-pull based model," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 8, pp. 3246-3254, Aug. 2006.

[6] H. H. Chen, C. H. Lin, P. C. Huang, and J. T. Chen, "Joint polynomial and look-up-table predistortion power amplifier linearization," *IEEE Trans. Circuits Syst. I, Exp. Briefs*, vol. 53, no. 8, pp. 612-616, Aug. 2006.

[7] K. Rawat. M. Rawat, and F. M. Ghannouchi, "Compensating I-Q imperfections in hyvrid RF/Digital predistortion with an adapted lookup table implemented in an FPGA," *IEEE Trans. Circuits Syst. I, Exp. Briefs*, vol. 57, no. 5, pp. 389-393, May 2010.

[8] A. Ghadam, S. Burglechner, A. Gokceoglu, M. Valkama, and A. Springer, "Implementation and performance of DSP-oriented feedforward power amplifier linearizer," *IEEE Trans. Circuits Syst. II, Reg. Papers*, vol. 59, no. 2, pp. 409-425, Feb. 2012.

[9] T. Jiang and G. X. Zhu, "Complement Block Coding For Reduction in Peak-To-Average Power Ratio of OFDM Signals," *IEEE Commun. Mag.*, vol. 43, no. 9, Sep. 2005.

[10] P. Y. Fan and X. G. Xia, "Block coded modulation for the reduction of the peak to average power ratio in OFDM systems," *IEEE Trans. Consumer Electronics*, vol. 45, no. 4, pp. 1025-1029, Nov. 1999.

[11] K. Sathananthan and C. Tellambura, "Coding to reduce both PAR and

PICR of an OFDM signal," *IEEE Commun. Lett.*, vol.6, no.8, pp. 316-318, Aug 2002.

[12] T. Jiang, W. D. Xiang, P. C. Richardson, D. M. Qu, and G. X. Zhu, "On the Nonlinear Companding Transform for Reduction in PAPR of MCM signals," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2017-2021, Jun. 2007.

[13] T. Jiang, Y. Yang and Y. Song, "Exponential companding technique for PAPR reduction in OFDM Systems," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 244-248, Jun. 2005.

[14] B. S. Krongold and D.L. Jones, "An Active-Set Approach for OFDM PAR Reduction via Tone Reservation," *IEEE Trans. Signal Process.*, vol.52, no.2, pp. 495-509, Feb. 2004.

[15] N. Andgart, P. Odling, A. Johansson, and P. O. Borjesson, "Designing Tone Reservation PAR Reduction," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp 82-82, 2006.

[16] B. Crongold and D. Jones, "PAR Reduction in OFDM via Active Constellation Extension," *IEEE Trans. Broadcast.*, vol.49, no. 3, Sep. 2003.

[17] Y. J. Kou, W. S. Lu, and A. Antoniou, "A New Peak-To-Average Power-Ratio Reduction Algorithm For OFDM Systems via Constellation Extension," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1823-1832, May 2007.

[18] H. Ochiai and H. Imai, "Performance of the deliberate clipping with adaptive symbol selection for strictly band-limited OFDM systems," *IEEE J. on Selected Areas in Commun.*, vol. 18, no. 11, pp. 2270-2277, Nov. 2000.

[19] 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, Jan. 1999.

[20] H. Chen and A. Haimovich, "Iterative estimation and cancellation of clipping noise for OFDM signals," *IEEE Commun. Lett.*, vol. 7, no. 7, pp. 305-307, Jul. 2003.

[21] H. Saeedi, M. Sharif, and F. Marvasti, "Clipping noise cancellation in OFDM systems using oversampled signal reconstruction," *IEEE Commun. Lett.*, vol. 6, no. 2, pp. 73-75, Feb. 2002.

[22] E. B. Al-Safadi and T. Y. Al-Naffouri, "On reducing the complexity of tone-reservation based PAPR reduction schemes by compressive sensing," *in Proc. IEEE Globecom*, pp. 1-6., USA, 2009.

[23] M. Allén, L. Toni, J. Marttila and M. Valkama,"Iterative signal processing for mitigation of analog-to-digital converter clipping distortion in multiband OFDMA receivers," *Journal of Electrical and Computer Engineering*, 2012.

[24] M. Allén, J. Marttila and M. Valkama, "Digital post-processing for reducing A/D converter nonlinear distortion in wideband radio receivers," *Signals, Systems and Computers, 2009 Conference Record of the Forty-Third Asilomar Conference on*, pp. 1111-1114, 2009.

[25] F. H. Gregorio, S. Werner, J. Cousseau, J. Figueroa, and R. Wichman, "Receiver-side nonlinearities mitigation using an extended iterative decision-based technique," *Signal Processing*, vol. 91, no. 8, pp. 2042-2056, Aug. 2011.

[26] S. Yu, A. Shaharyar Khwaja, and J. Ma, "Compressed sensing of complex-valued data," *Signal Processing*, vol. 92, no. 2, pp. 357-362, Feb. 2012.

[27] T. Ince, A. Nacaroglu, and N. Watsuji, "Nonconvex compressed sensing with partially known signal support," *Signal Processing*, vol. 93, no. 1, pp. 338-344, Jan. 2013.

[28] N. Kalouptsidis, G. Mileounis, B. Babadi, and V. Tarokh, "Adaptive algorithms for sparse system identification," *Signal Processing*, vol. 91, no. 8, pp. 1910-1919, Aug. 2011.

[29] E. Candes, J. Romberg and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math,* vol. 59, no. 8, pp. 1207-1223, Aug. 2006.

[30] E. J. Candes and P. Randall, "Highly robust error correction by convex programming," *IEEE Trans. Inform. Theory*, vol. 54, no. 7, pp. 2829–2840, Jun. 2008.

[31] E. Candes and T. Tao, "The Dantzig selector: statistical estimation when p is much larger than n," *Annals of Statistics*, vol. 35, no. 6, 2313–2351, 2007.

[32] J. Tropp, "Just relax: Convex programming methods for identifying sparse signals in noise," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 1030–1051, 2006.

[33] L. Taijun, S. Boumaiza, and F. M. Ghannouchi,"Deembedding static nonlinearities and accurately identifying and modeling memory effects in wideband RF transmitters," *IEEE Trans. Microw. Theory Tech.*, vol. 53, no. 11, pp. 3578-3587 , 2005.

[34] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," *Research Laboratory of Electronics, Massachusetts Institute of Technology*, 2012.

[35] A. A. Quadeer and T. Y. Al-Naffouri, "Structure-Based Bayesian Sparse Reconstruction," *IEEE Trans. Signal Proc.*, vol. 60, pp. 6354-6367, 2012.

[36] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Proc.*, vol. 56, pp. 2346-2356, 2008.

[37] A. A. Quadeer, S. F. Ahmed, and T. Y. Al-Naffouri, "Structure based Bayesian sparse reconstruction using non-Gaussian prior," *Allerton Conf. Commun., Control, and Comp.*, USA, Dec. 2011.

[38] P. Schniter, L. C. Potter, and J. Ziniel, "Fast Bayesian Matching Pursuit," *Proc. Workshop on Information Theory and Applications (ITA)*, (La Jolla, CA), Jan. 2008.

[39] T. Blumensath and M. E. Davies, "Gradient Pursuits," *IEEE Trans. on Signal Proc.*, vol. 56, no. 6, pp. 2370-2382, Jun. 2008.

[40] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4655-4666, Dec. 2007.

[41] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comp. Harmonic Analysis*, 2008.

[42] E. J. Candes, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 877-905, 2008.

[43] M. D. McKinley, K. A. Remley, M. Myslinski, J. S. Kenney, D. Schreurs, and B. Nauwelaers, "EVM Calculation for Broadband Modulated Signals," *64th ARFTG Conf. Dig., Orlando, FL*, pp. 45–52, Dec. 2004.

[44] S. Forestier, P. Bouysse, R. Quere, A. Mallet, J.-M. Nebus, and L. Lapierre, "Joint optimization of the power-added efficiency and the error-vector measurement of 20-GHz pHEMT amplifier through a new dynamic bias-control method," *IEEE Trans. Microwave Theory and Techniques*, vol. 52, pp. 1132-1141, April 2004.

[45] 3GPP TS 36.104 V9.4.0 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) radio transmission and reception (Release 9), 2010-06.

[46] 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.

[47] K. M. 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.

[48] T. Y. Al-Naffouri, K. M. 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.

[49] J. J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson and P. O. Borjesson, "On channel estimation in OFDM systems", *Vehicular Technology Conference, 1995 IEEE 45th*, pp. 815-819.

[50] B. Wu, S. Cheng and H. Wang,"Iterative channel estimation and signal detection in clipped OFDM",*Global Telecommunications Conference, 2005. GLOBECOM'05. IEEE.*

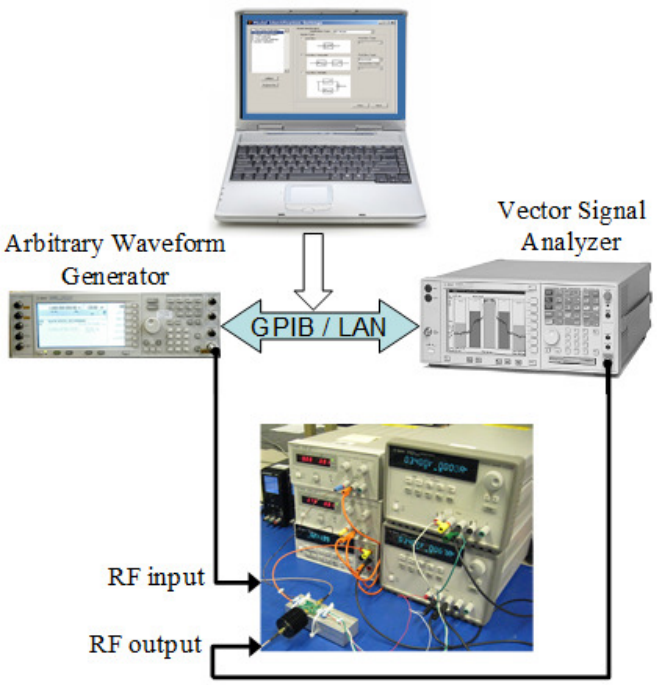
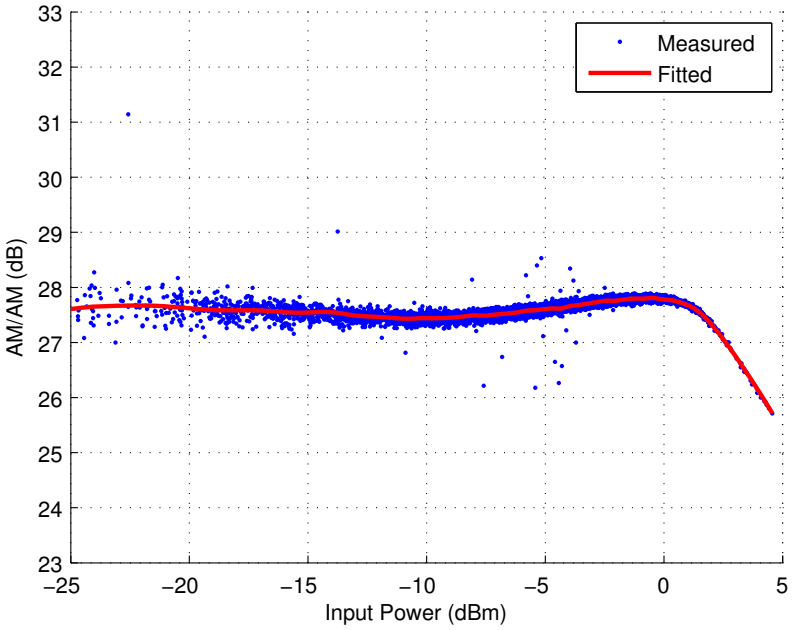
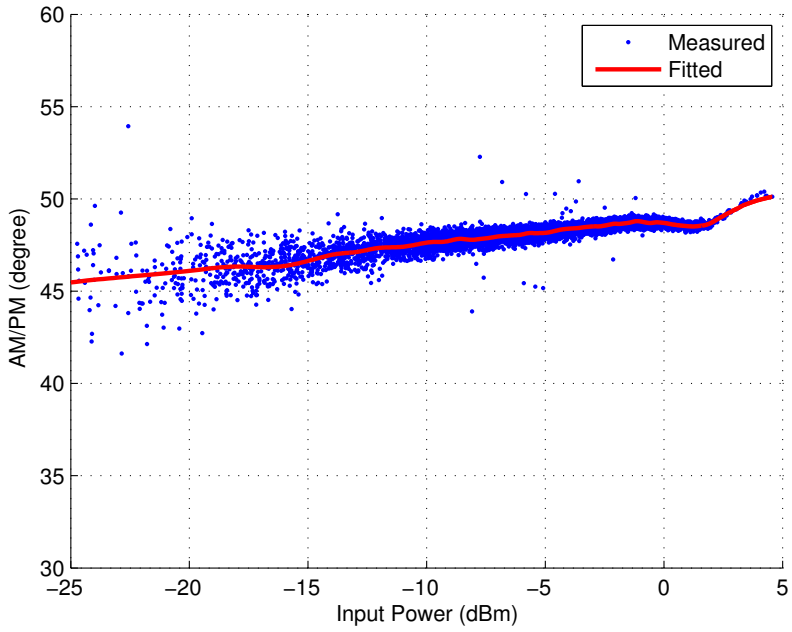


Figure 1: Experimental setup for DUT characterization.

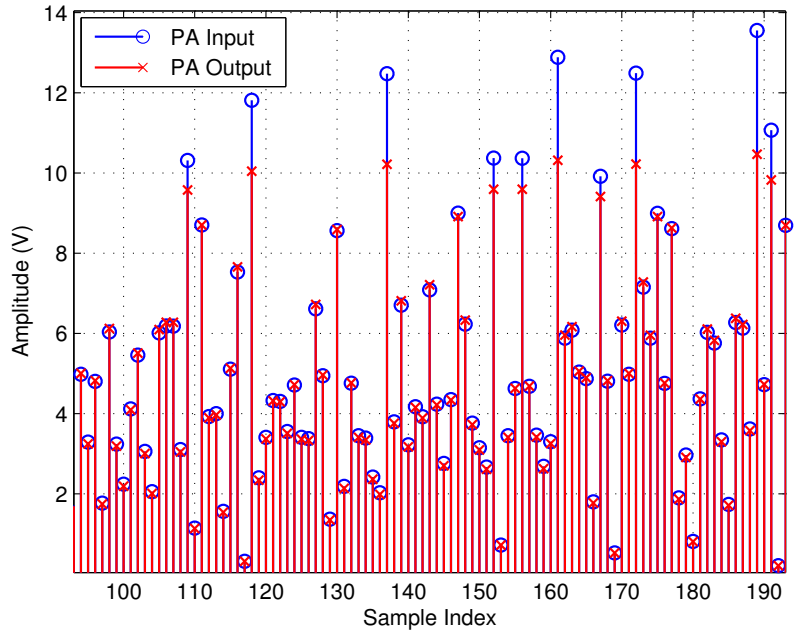


(a)

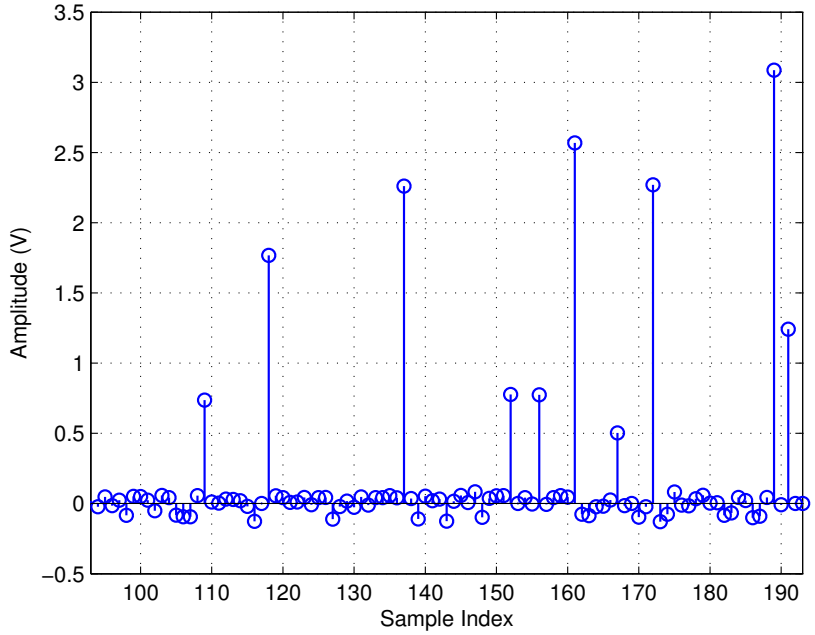


(b)

Figure 2: Measured and fitted characteristics of the amplifier. (a) AM/AM characteristics. (b) AM/PM characteristics.



(a)



(b)

Figure 3: Sample snapshots. (a) PA input (scaled) and output, showing nonlinear distortion particularly large at high amplitudes. (b) Distortion $\mathbf{x}_c$, of signal in (a).

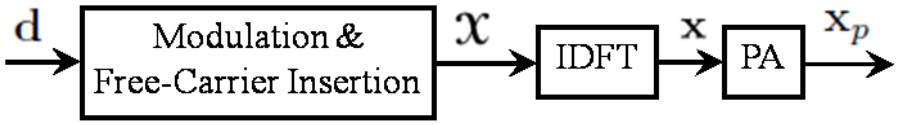


Figure 4: Simplified block diagram of the transmitter.

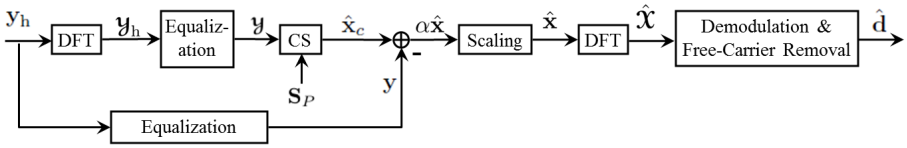
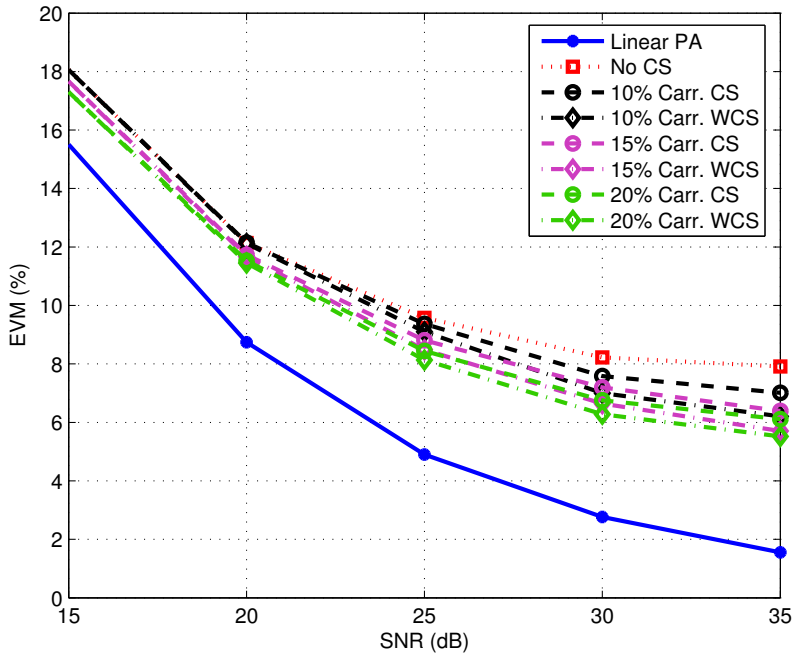
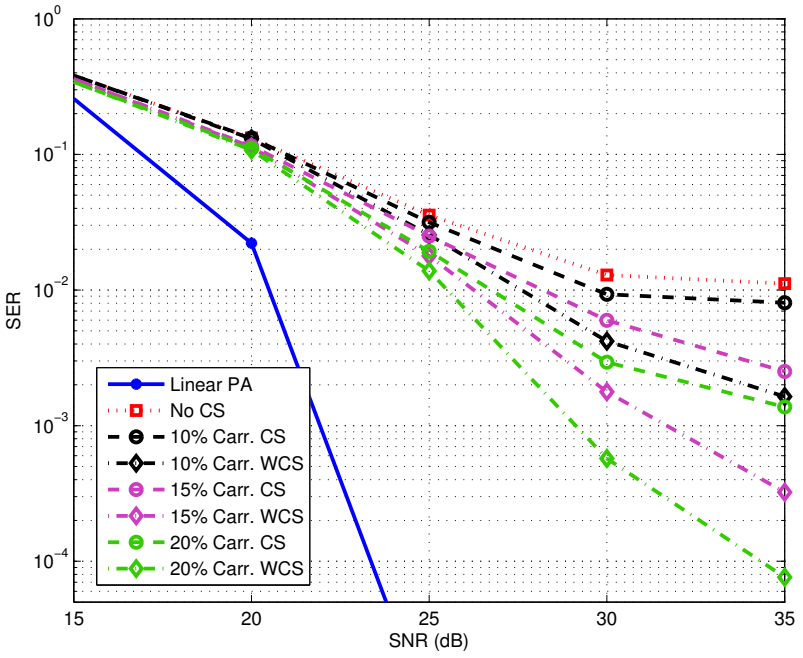


Figure 5: Compressive sensing implementation at the receiver.



(a)



(b)

Figure 6: Performance comparison of CS and WCS for different number of free carriers. (a) EVM as a function of SNR. (b) SER as a function of SNR.

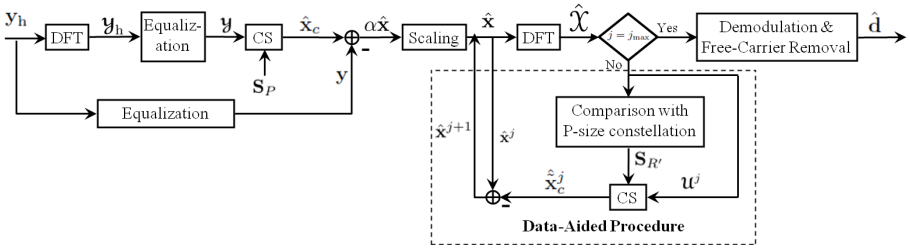


Figure 7: Data-aided compressive sensing implementation at the receiver.

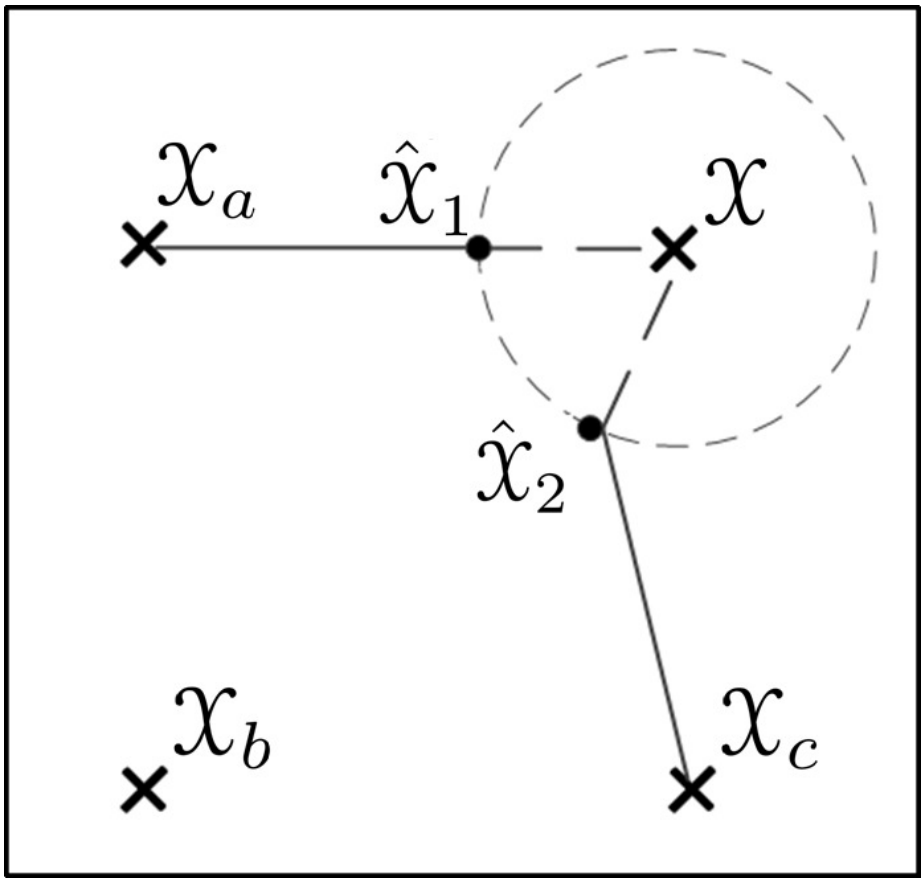
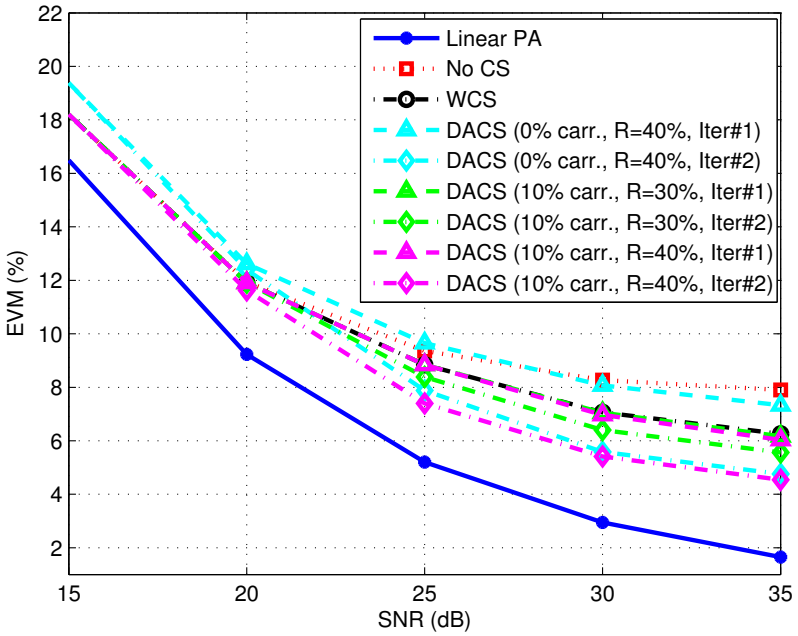
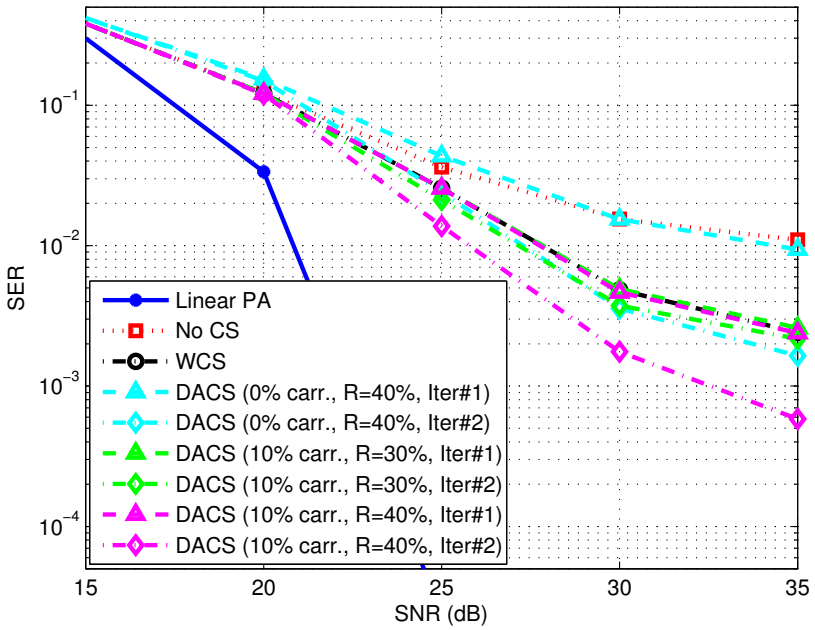


Figure 8: Graphical explanation of the adopted reliability criterion.



(a)



(b)

Figure 9: Performance comparison of DACS for different number of free carriers, reliable carriers, and iterations. (a) EVM as a function of SNR. (b) SER as a function of SNR.

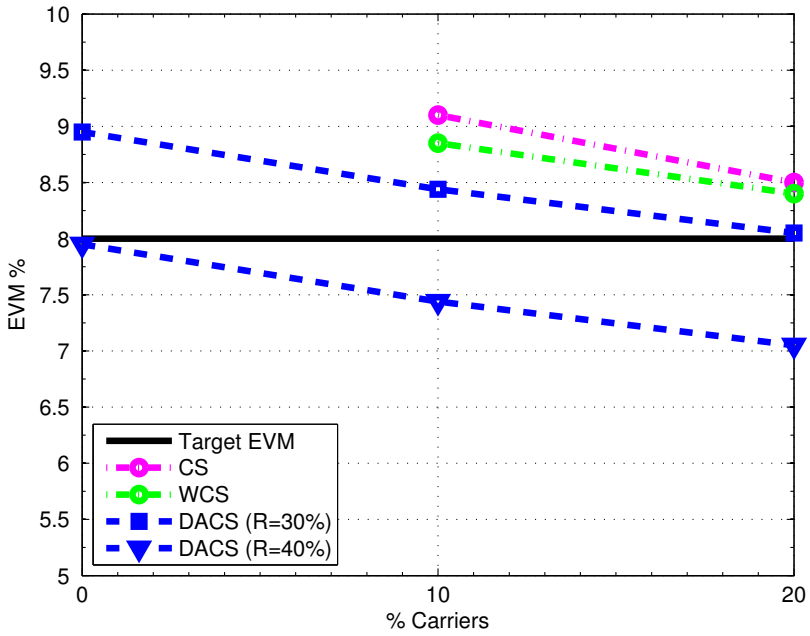
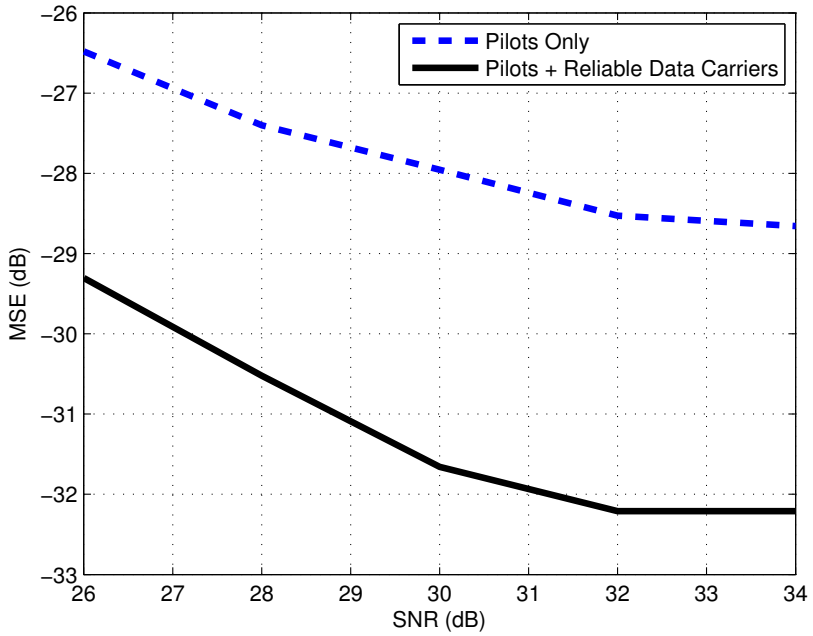
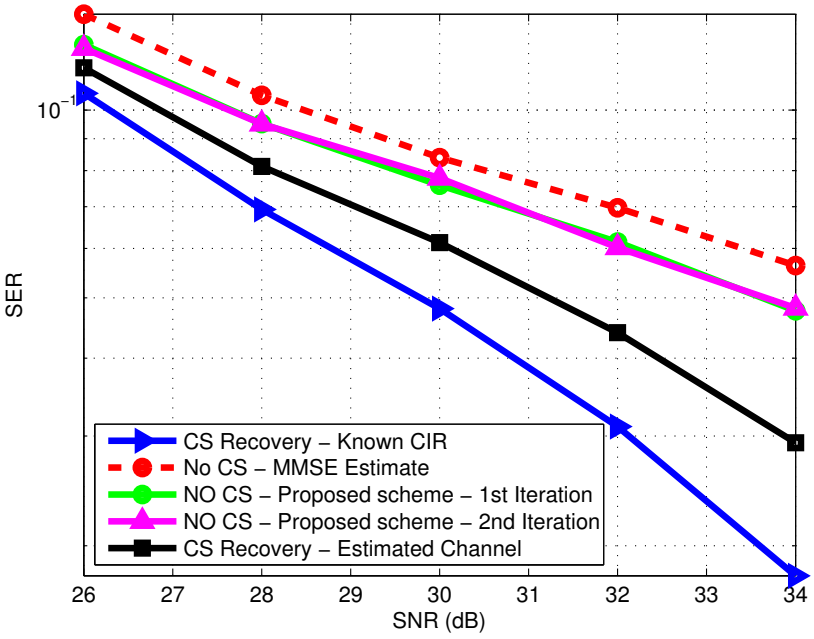


Figure 10: Performance comparison of the proposed algorithms w.r.t EVM as a function of number of free carriers at SNR = 25 dB.



(a)



(b)

Figure 11: Channel Estimation Performance (a) MSE Performance enhancement using proposed iterative scheme (b) SER performance comparison.

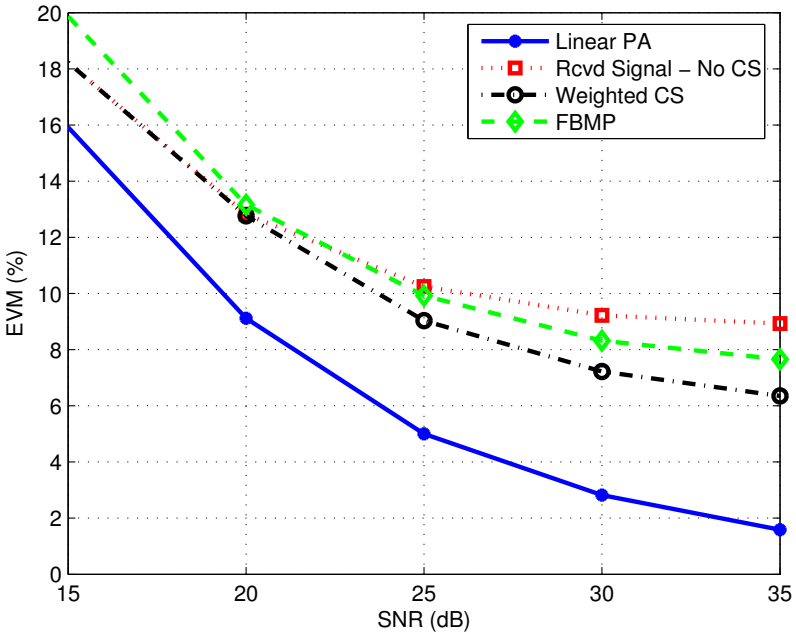


Figure 12: EVM performance comparison of WCS vs. FBMP

Table 1: Comparison of CS algorithms

Algorithm Implemented	**Bandwidth Efficiency (%)**	**Execution Time (%)**	EVM (%)
WCS (10% pilots)	90	32.0	7.3014
WCS (20% pilots)	80	53.5	5.9608
WCS (10% pilots) + DACS (40% carriers)	90	74.0	6.9705
DACS (1 iter. 40% carr.) [No WCS]	100	74.0	8.2075
DACS (2 iter. 40% carr.) [No WCS]	100	100	5.3392