Least-Squares Iterative PAR Reduction for Point-to-Point Large-Scale MIMO-OFDM Systems

Abdul Wakeel and Werner Henkel
Transmission Systems Group (TrSyS)
Jacobs University Bremen
Bremen 28759, Germany
Email: {a.wakeel and w.henkel}@jacobs-university.de

***Abstract*—This paper addresses peak-to-average ratio (PAR) reduction in orthogonal frequency division multiplexing (OFDM) based large-scale multiple-input multiple-output (MIMO) systems. We propose a new technique for PAR reduction in point-to-point scenarios. The last eigenchannels of a massive MIMO channel are often very weak. Not using them for data transmission will offer redundancy for PAR reduction. These eigenchannels are used to approximate the peaks which exceed a given target value in a least squares fashion. This approximate exceedence model is then subtracted from the original signal in time domain for PAR reduction. It has been shown through simulation results that a considerable gain can be obtained with the proposed method with marginal increase in the average power and almost negligible loss in the channel capacity.**

***Index Terms*—Large-scale multiple-input multiple-output (MIMO), singular value decomposition (SVD), precoding, orthogonal frequency division multiplexing (OFDM), peak-to-average ratio (PAR), least squares estimation.**

I. Introduction

With the advent of new applications, the demand for high data rates are increasing day by day. New concepts and new technologies have been introduced to meet the public demands. One of the promising technique which has gained a lot of popularity in modern day communication systems is orthogonal frequency division multiplexing (OFDM), which offers high data rates flexibly. Moreover, due to its simple implementation, robustness against frequency selective fading in its coded form, and easy treatment of inter-symbol interference, OFDM is making its way into all wireless and wireline communication systems.

To reap the benefits of spatial multiplexing and provide high through puts, massive or large-scale multiple input multiple output (MIMO) is a concept for future wireless communication systems. The concept was first proposed in [2] for multiuser MIMO (MU-MIMO) systems. The transmitter at the base station (BS) was supposed to have a large number of transmit antennas serving single antenna terminal units (TU). However, we adopt the term large-scale MIMO and extend it to a point-to-point MIMO system, where the transmitter and the receiver are supposed to be equipped with large number of antennas, in the range of hundred (e.g. communication between two BSs). Thus, OFDM-based large-scale MIMO will be a choice for future wireless communication technologies for point-to-point communications. However, a major drawback of OFDM is its high peak-to-average ratio (PAR), which is defined as

$$\text{PAR} = \frac{\max_k |x_k|^2}{E\{|x_k|^2\}} \,, \tag{1}$$

where $\max_k |x_k|^2$ represents the maximum instantaneous peak power and $E\{|x_k|^2\}$ denotes the average power of the OFDM signal. The high PAR of the OFDM signal is due to the statistical independence of the carriers, which results in an almost Gaussian distributed time-domain signal, which exhibits a high PAR. This high PAR drives the power amplifier (PA) to operate in its non-linear range. It has been mentioned in [2], that large-scale MIMO systems will be equipped with low-power low-cost radio-frequency (RF) components. For low-power RF components to operate in its linear range, it is necessary to lower down the dynamic range of the transmitted signal. Different techniques have been proposed for PAR reduction in MIMO-OFDM systems, equipped with a small number of antennas. The well-known techniques amongst them are Selected Mapping (SLM) [4], [5], Partial Transmit Sequences (PTS) [6], and Tone Reservation (TR) [8]. However, for large-scale MU-MIMO-OFDM systems, a first approach has been made in [9], [10].

In this paper, we propose a new method for PAR reduction in point-to-point (PtP) large-scale MIMO-OFDM systems using the weakest eigenchannel(s) to estimate and model the threshold excursion caused by the remaining dimensions. This model function is then subtracted from the original signal in time domain for PAR reduction. The procedure is iterated to reach the PAR target value. The notations used throughout this paper are capital and lowercase letters for DFT domain and time-domain representations, respectively. Boldface letters are used to represent matrices and vectors.

The rest of this paper is organized as follows. In Section II, we present the key idea. The system model is discussed in Section III. Section IV provides the core idea of estimating the exceedence values above a certain threshold. The simulation results for our approach are shown in Section V. Section VI finally concludes the paper.

II. Key idea

Let $\mathbf{H}(n)$ be an $M_r \times M_t$ channel gain matrix in DFT domain at carrier n known to both the transmitter and the receiver. Using singular value decomposition (SVD), the channel

matrix $\mathbf{H}(n)$ for a point-to-point large-scale MIMO can be rephrased as

$$\mathbf{H}(n) = \mathbf{U}(n) \cdot \mathbf{\Lambda}(n) \cdot \mathbf{V}^H(n) \ , \tag{2}$$

where $\mathbf{V}(n)$ and $\mathbf{U}(n)$ are unitary pre-processing and post-processing matrices, respectively, ($\mathbf{V}^H \cdot \mathbf{V} = \mathbf{U}^H \cdot \mathbf{U} = \mathbf{I}$) and $\mathbf{\Lambda}(n)$ is a diagonal matrix of the singular values of $\mathbf{H}(n)$. For $M_t = M_r$, $\mathbf{\Lambda}(n)$ can be written as

$$\mathbf{\Lambda}(n) = \begin{pmatrix} \sigma_{1,1}(n) & 0 & 0 & 0 \\ 0 & \sigma_{2,2}(n) & 0 & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma_{M_t,M_t}(n) \end{pmatrix} . \tag{3}$$

Usually, the singular values are sorted in a descending order along the diagonals by the SVD algorithm, i.e., $\sigma_{1,1} > \sigma_{2,2} > \ldots > \sigma_{M_t,M_t}$. For large-scale MIMO-OFDM systems, the last singular value is typically so small that the corresponding eigenchannels are hardly suited for data transmission. Not using them would offer redundancy for peak-to-average ratio (PAR) reduction without a lot of cost in data rate. Throughout this paper, it is assumed that the last eigenchannel is too weak and is hence utilized for PAR reduction.

III. System Model

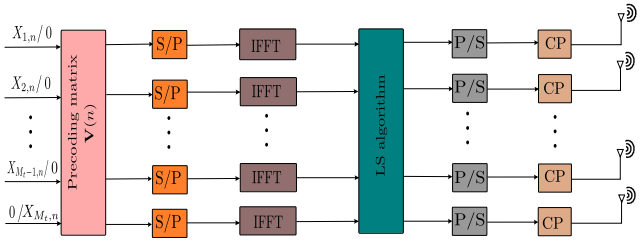

Fig. 1. Transmitter system model of MIMO-OFDM for LS algorithm

Figure 1 shows transmitter model with some basic components for a point-to-point MIMO-OFDM. For a large-scale MIMO, the number of transmit antennas can be in the range of hundred, however, for the purpose of this paper, we have considered a 40×40 MIMO-OFDM system, i.e., $M_t = M_r = 40$. As shown in Fig. 1, the input vector $\mathbf{X}(n) = (X_{1,n}, X_{2,n}, \cdots, X_{M_t,n})^T$, where $X_{\mu,n}$ is the input symbol at the μth ($\mu = 1, 2, \cdots, M_t$) spatial channel and the nth carrier is pre-processed using a preprocessing matrix and converted into a parallel data stream using a serial to parallel converter. The time-domain signal is obtained taking an IFFT. The time-domain signal is padded with a cyclic prefix (CP) to mitigate the effect of inter-symbol interference (ISI) after parallel-to-serial conversion. This signal is transmitted over a MIMO channel with a channel gain matrix $\mathbf{H}$. At the receiver, reverse processes are used to obtain an estimate of the original signal. For better understanding, we give the input-output relation in frequency domain here. Let $\mathbf{Y}(n) = (Y_{1,n}, Y_{2,n}, \cdots, Y_{M_r,n})^T$ be the output vector, where $Y_{\rho,n}$ is the output symbol of the ρth ($\rho = 1, 2, \cdots, M_r$) spatial channel. The transmit signal $\mathbf{X}(n)$ is pre-multiplied by $\mathbf{V}(n)$, whereas the signal at the receiver is post-multiplied by $\mathbf{U}^H(n)$ to obtain the output $\mathbf{Y}(n)$ as shown in Fig. 2.

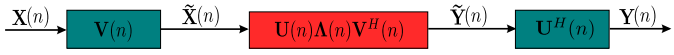

Fig. 2. SVD MIMO diagonalization

$$\mathbf{Y}(n) = \overbrace{\mathbf{U}(n)^H \cdot \underbrace{\mathbf{U}(n)}_{\mathbf{I}} \cdot \mathbf{\Lambda}(n) \cdot \underbrace{\mathbf{V}^H(n)}_{\mathbf{I}} \cdot \mathbf{V}(n)}^{\mathbf{H}(n)} \cdot \underbrace{\mathbf{X}(n)}_{\tilde{\mathbf{X}}(n)} , \tag{4}$$

which can also be written as

$$\mathbf{Y}(n) = \mathbf{\Lambda}(n) \cdot \mathbf{X}(n) \ ,$$

For the $M_t \times M_r = 40 \times 40$ case, Eq. (4) is rephrased as

$$\mathbf{Y}(n) = \begin{pmatrix} \sigma_{1,1}(n) & 0 & 0 & 0 \\ 0 & \sigma_{2,2}(n) & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sigma_{M_t,M_t}(n) \end{pmatrix} \begin{pmatrix} X_{1,n} \\ X_{2,n} \\ \vdots \\ X_{M_t,n} \end{pmatrix} . \tag{5}$$

It is assumed that only the last spatial dimension is reserved for PAR reduction while the other dimensions are used for data transmission. Let $\mathbf{S}(n)$ denote the first dimensions used for data transmission, i.e.,

$$\mathbf{S}(n) = \begin{pmatrix} X_{1,n} & X_{2,n} & \ldots & X_{M_t-1,n} & 0 \end{pmatrix}^T .$$

For estimating and modeling the peak values which exceed a given threshold, we use $\mathbf{R}(n)$, defined as,

$$\mathbf{R}(n) = \begin{pmatrix} 0 & 0 & \ldots & 0 & X_{M_t,n} \end{pmatrix}^T ,$$

whereas $X_{M_t,n}$ should not be confused with the input data, it contains the modeled corrective signal that is used for PAR reduction. As shown in Fig. 1, the input data vector $\mathbf{S}(n)$ and corrective signal $\mathbf{R}(n)$ are pre-processed as

$$\tilde{\mathbf{S}}(n) = \mathbf{V}(n) \cdot \mathbf{S}(n) = \mathbf{V}(n) \cdot \begin{pmatrix} X_{1,n} & \ldots & X_{M_t-1,n} & 0 \end{pmatrix}^T \tag{6}$$

and

$$\tilde{\mathbf{R}}(n) = \mathbf{V}(n) \cdot \mathbf{R}(n) = \mathbf{V}(n) \cdot \begin{pmatrix} 0 & \ldots & 0 & X_{M_t,n} \end{pmatrix}^T . \tag{7}$$

The two are then transformed into time domain and subtracted, i.e.,

$$\tilde{\mathbf{x}} = \tilde{\mathbf{s}} - \tilde{\mathbf{r}} = \mathbf{F}^{-1}\mathbf{V}\mathbf{X} = \mathbf{F}^{-1}\mathbf{V}(\mathbf{S} - \mathbf{R}) \ , \tag{8}$$

using a block-IDFT matrix $\mathbf{F}^{-1}$ with blocks of diagonal submatrices with $M_t = M_r (= 40)$ identical diagonal elements $w_{n,k} = \frac{1}{\sqrt{N}} e^{j2\pi kn/N}$. All vectors $\mathbf{X}, \tilde{\mathbf{X}}, \ldots$ are obtained by concatenating the vectors $\mathbf{X}(n), \tilde{\mathbf{X}}(n), \ldots$. Now, the goal is to estimate and model the exceedence values by $\mathbf{R}$ in DFT domain to finally subtract the model.

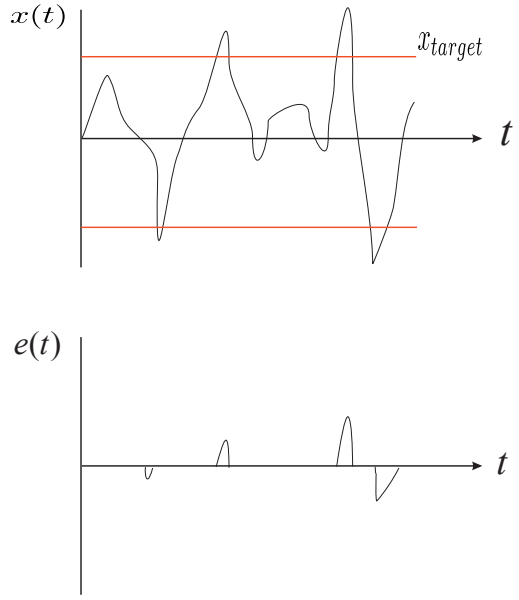

Fig. 3. Representation of exceedence values

IV. Modeling the peak excursions by the last spatial dimension

Let $\tilde{\mathbf{s}}$ be the transmitted signal in time domain with no data transmission on the reserved eigenchannels. The proposed algorithm searches for any peak value that exceeds a PAR target value. These exceedence excursions are then represented and modeled by the last spatial dimension in frequency domain. Let $\mathbf{e}$ be a vector representing the exceeding excursions in time domain on all spatial dimensions of a large-scale MIMO-OFDM transmitted signal, i.e.,

$$\mathbf{e} = ((e_{1,1}e_{2,1}\cdots e_{M_t,1}),\cdots,(e_{1,n}e_{2,n}\cdots e_{M_t,n}))^T, \quad (9)$$

where

$$e_{\mu,k} = \begin{cases} 0 & \text{for } x_{\mu,k} \leq x_{target}\ , \\ x_{\mu,k} - e^{j\angle(x_{\mu,k})}\cdot x_{target} & \text{for } x_{\mu,k} > x_{target}\ . \end{cases} \quad (10)$$

Let $\mathbf{E}$ be the DFT-domain counterpart of $\mathbf{e}$, i.e., $\mathbf{E} = \mathbf{F}\cdot\mathbf{e}$, where $\mathbf{F}$ is a block DFT matrix. We will use two equivalent approaches to estimate the peak values exceeding the given threshold.

1) 1st approach: In order to estimate and model the exceedence excursions by the reserved spatial dimension, Eq. (7) can be reformulated as

$$\begin{pmatrix} \mathbf{V}(0) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}(1) & \dots & \mathbf{0} \\ \vdots & \dots & \ddots & \vdots \\ \mathbf{0} & \cdots & \cdots & \mathbf{V}(N-1) \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ X_{M_t,0} \\ \vdots \\ 0 \\ \vdots \\ X_{M_t,N-1} \end{pmatrix} = \boldsymbol{\varphi}\ , \quad (11)$$

for $n = 0,...,N-1$, which in more compact form can be written as

$$\boldsymbol{\varphi} = \mathbf{V}\cdot\mathbf{R}\ , \quad (12)$$

where $\mathbf{V}$ is a block-diagonal unitary matrix. The column vector $\boldsymbol{\varphi}$ shall approximate $\mathbf{E}$ in a least-squares sense at all frequencies representing the determined exceedence peaks, i.e.,

$$\min_{\mathbf{R}} ||\boldsymbol{\varphi} - \mathbf{E}||^2\ . \quad (13)$$

Eq. (13) can be solved for $\mathbf{R}$ as a least squares estimation,

$$\mathbf{R} = \mathbf{V}^{-1}\cdot\mathbf{E}\ . \quad (14)$$

Inside $\mathbf{R}$, only the components at positions $M_t(i)$, $i = 1,...,N$ will be kept, eliminating (zeroing) all others. The modified $\mathbf{R}$ is then pre-coded, transformed into time domain using the IFFT, and is subtracted from the original signal for PAR reduction.

2) 2nd approach: As shown, Eq. (7) essentially only addresses the last columns of the blocks $\mathbf{V}(i)$ inside the $\mathbf{V}$ matrix. Thus, Eq. (7) is rephrased as

$$\begin{pmatrix} \mathbf{V}_{:M_t}(0) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{:M_t}(1) & \dots & \mathbf{0} \\ \vdots & \dots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{V}_{:M_t}(N-1) \end{pmatrix} \begin{pmatrix} X_{M_t,0} \\ X_{M_t,1} \\ \vdots \\ X_{M_t,N-1} \end{pmatrix} = \boldsymbol{\varphi}\ , \quad (15)$$

where $\mathbf{V}_{:M_t}(n)$ represents the M_tth column (: stands for all rows) of the nth $\mathbf{V}$ matrix, which can be represented in a compact form as

$$\boldsymbol{\varphi} = \mathbf{V}_{:M_t}\cdot\mathbf{R}_{M_t}\ , \quad (16)$$

where $\mathbf{V}_{:M_t}$ is a block diagonal matrix with the M_tth column of $\mathbf{V}(n)$ at the nth diagonal, and $\mathbf{R}_{M_t}$ contains only the last spatial components. In a least-squares sense, the column vector $\boldsymbol{\varphi}$ shall approximate $\mathbf{E}$ according to

$$\min_{\mathbf{R}_{M_t}} ||\boldsymbol{\varphi} - \mathbf{E}||^2\ . \quad (17)$$

In Eq. (17), the number of equations are more than the number of unknowns. Thus, this leads to a pseudo-inverse solution. Solving Eq. (17) for $\mathbf{R}_{M_t}$ in least-squares sense results in

$$\mathbf{R}_{M_t} = (\mathbf{V}_{:M_t}^H\cdot\mathbf{V}_{:M_t})^{-1}\cdot\mathbf{V}_{:M_t}^H\cdot\mathbf{E}\ , \quad (18)$$

where H stands for Hermitian (conjugate transpose). Both approaches are equivalent and, of course, deliver the same result. The first one, however, is the less complex, since no pseudo-inverse needs to be computed. The procedure is iterated, as the peaks exceeding the target value are not estimated exactly by the least-squares approach in one iteration. The least squares algorithm is described as follows.

Least Squares algorithm

1) Initialize $\mathbf{S}$ to be the DFT-domain information vector, with the reserved dimension set to zero. Precode it using the pre-processing matrix as shown in Eq. 6. Set $i = 0$.
2) Initialize the time-domain solution $\tilde{\mathbf{s}}$, i.e., $\tilde{\mathbf{s}} = \text{IFFT}(\tilde{\mathbf{S}})$.

3) Initialize $\mathbf{e}$ to represent the exceedence excursion according to Eq. (10) and transform it into DFT domain, i.e., $\mathbf{E} = \text{FFT}(\mathbf{e})$.
4) Approximate $\mathbf{E}$ by the last spatial dimension using equations (14) or (18). Precode it with the first three dimensions set to zero as shown in Eq. (7) and transform it into time domain applying the IFFT modulator, i.e., $\tilde{\mathbf{r}} = \mathbf{F^{-1}}\tilde{\mathbf{R}}$.
5) Update the time-domain vector

$$\tilde{\mathbf{s}}^{(i+1)} = \tilde{\mathbf{s}}^{(i)} - \tilde{\mathbf{r}}\,, \tag{19}$$

$i = i + 1$ and go to Step 3.

i is the iteration counter and $\tilde{\mathbf{s}}$ denotes the $M_t \cdot N$ long time-domain vector, where M_t is the number of transmit antennas and N is the IFFT length. The structure follows (9).

It is very unlikely for all antenna channels to encounter a value exceeding the threshold at the same time. Hence, the least-squares approach will not approximate the peak, but yield a result reduced by a factor of M, since it tries, at the same time, to approximate the zeros (no exceedences) in the other antenna channels.[1] This is the drawback of an l_2 norm instead of l_∞. We introduce a weighting factor γ in the algorithm, which leads to optimum performance for $\gamma = M$. The weighting can be realized by modifying steps 3 or 5 of the presented LS algorithm. Modifying Step 3 means weighting the peak limit excursions before mapping them onto the last spatial dimension. Equation (10) becomes

$$e_{\mu,k} = \begin{cases} 0 & \text{for } x_{\mu,k} \le x_{target}\,, \\ \gamma(x_{\mu,k} - e^{j\angle(x_{\mu,k})} \cdot x_{target}) & \text{for } x_{\mu,k} > x_{target}\,. \end{cases} \tag{20}$$

Alternatively, one could instead write Step 5, Eq. (19) as

$$\tilde{\mathbf{s}}^{(i+1)} = \tilde{\mathbf{s}}^{(i)} - \gamma\tilde{\mathbf{r}}\,. \tag{21}$$

V. Results and Discussion

A massive MIMO-OFDM systems is expected to be equipped with a large number of antennas (up to one hundred or more), however, for our simulation results we have considered a 40×40 MIMO-OFDM system with $N = 128$ subcarriers and 16-QAM modulation. It is assumed that the channel is perfectly known at the transmitter and the receiver (perfect CSI). The channel matrix $\mathbf{H}$ is taken from [4], which can be described by a matrix polynomial $\mathbf{H}(z) = \sum_{k=0}^{l_H-1} \mathbf{H}_k z^{-k}$, where $\mathbf{H}_k$ is an $\mathbf{M}_r \times \mathbf{M}_t$ matrix having complex fading channel coefficients. The channel coefficients are i.i.d. complex Gaussian distributed with zero mean and variance $1/l_H$, with $l_H = 20$.

Due to transmitter-sided precoding, the average power is distributed over all spatial dimensions. Moreover, for PAR reduction a time-domain signal $\tilde{\mathbf{r}}$ is added to the original signal $\tilde{\mathbf{s}}$, thus, the average power increases per iteration. The PAR after applying the algorithm is defined in here as

$$\text{PAR} = \frac{\max_{\forall_\mu,\forall_k} |x_{\mu,k} + \gamma r_{\mu,k}|^2}{\sigma^2}\,, \tag{22}$$

[1]Illustration: Assume e and $M-1$ zeros to be approximated in least-squares sense. This means $\frac{d}{dx}[(x-e)^2 + (M-1)(x-0)^2] = 0 \Longrightarrow x = e/M$.

where k is the sample index, and $\sigma^2 = E_{\forall_\mu,\forall_k}\{|x_{\mu,k}|^2\}$ is chosen to be the average power (over all spatial dimensions) without any PAR measures, i.e., with an unused spatial dimension and without any increase in the average power after the algorithm. As a performance measure, we consider the complementary cumulative distribution function (ccdf), which is the probability that the current PAR exceeds a certain threshold τ_{th}, i.e., $Pr\{PAR > \tau_{th}\}$ [4].

Figure 4 shows the simulation results for a massive MIMO-OFDM system with a target PAR value of 6.5 dB, and a weighting factor $\gamma = 30$. A gain of approximately 5.0 dB is obtained at 10^{-5} with as few as 8 iterations. Moreover, Fig. 4 clearly shows a gain of 3.2 dB with the first iteration and as much as 4.2 dB with only 3 iterations. The higher number of iterations show that the algorithm converges to the target value. However, the computational complexity of the proposed algorithm increases with the number of iterations, as more IFFT/ FFTs are needed to alternate between time and DFT domains.

As mentioned earlier, for modeling the exceedence excursions, our algorithm searches for the peak values that exceeds a given target value on all spatial dimensions in time domain. These values are transformed into DFT domain, and are estimated by the last spatial dimension. The estimated model is re-transformed into time domain and are subtracted from the original signal for PAR reduction. However, in doing so, two facts needs to be taken into consideration:

1) Capacity loss: Loss in channel capacity due to the reserved eigenchannels.
2) Relative mean power increase ΔE: Increase in the aver-

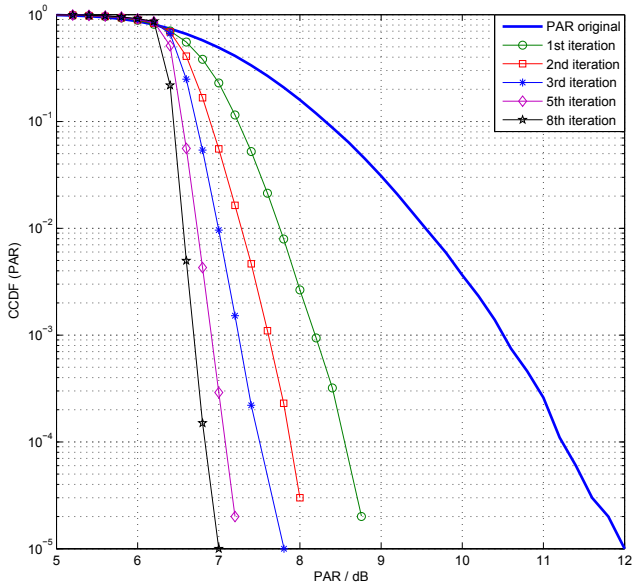

Fig. 4. CCDF(PAR) of the LS algorithm for a PAR target value of 6.5 dB without mean power constraints ($\gamma = 30$)

age power per iteration due to addition of the corrective signal, and performance of the algorithm with a mean power constraint.

A. Capacity analysis of the proposed algorithm

The proposed algorithm reserves the weakest eigenchannel (spatial dimension) to model the peak excursion due to the other spatial dimensions. Thus, there is a slight loss in the channel capacity. SVD decomposes the channel matrix $\mathbf{H}$ into independent parallel subchannels, thus, the capacity of a large-scale MIMO channel is equal to the sum of the capacities of the independent parallel subchannels,

$$C_{MIMO} = \sum_{i=1}^{\min(M_r, M_t)} \log_2(1 + \alpha_i \sigma_i^2) \,, \qquad (23)$$

where σ_i^2 is the ith singular value of the channel matrix $\mathbf{H}$, $\min(M_r, M_t)$ is the minimum of the number of receive or transmit antennas, and $\alpha_i = \Phi_i/\zeta^2$ denotes the transmitter-sided signal-to-noise ratio (SNR) of the ith parallel subchannel. For a MIMO-OFDM channel, (23) is formulated as

$$C_T = \sum_{n=1}^{N} \sum_{i=1}^{\min(M_r, M_t)} \log_2(1 + \alpha_i(n) \sigma_i^2(n)) \,, \qquad (24)$$

where $\sigma_i^2(n)$ and $\alpha_i(n)$ are the singular values and the SNRs at the nth subcarrier, respectively.
In case of an LS algorithm with no data transmission on the last spatial dimension, Eq. (24) becomes

$$C_{LS} = \sum_{n=1}^{N} \sum_{i=1}^{\min(M_r, M_t)-1} \log_2(1 + \alpha_i(n) \sigma_i^2(n)) \,. \qquad (25)$$

Figure 5 shows the capacity curves for a 40×40 MIMO-OFDM channel with and without reserved eigenchannels (using (24) and (25), respectively). It is obvious from Fig. 5 that at low SNR values the capacity curve for reserved eigenchannels overlap the actual capacity curve while there is a negligible capacity loss at higher SNR values.
Figure 5 also shows capacity curves for 2%, 5%, and 10% reserved tones for large-scale MIMO-OFDM obtained using the following equation

$$C_T = \sum_{\substack{N-\theta \\ \text{carriers}}} \sum_{i=1}^{\min(M_r, M_t)} \log_2(1 + \alpha_i(n) \sigma_i^2(n)) \,. \qquad (26)$$

where, θ is the number of reserved tones. It is clear from the figure that the capacity loss is very high for tone reservation. The capacity loss in our approach is far lower than for 2% reserved tones even at higher SNR values.

B. Mean power effects of the proposed algorithm

In order to reduce the clipping probability of the signal $\tilde{\mathbf{s}}$, a time-domain signal $\tilde{\mathbf{r}}$ is added to $\tilde{\mathbf{s}}$ for PAR reduction. However, in doing so, the mean transmit power is increased.

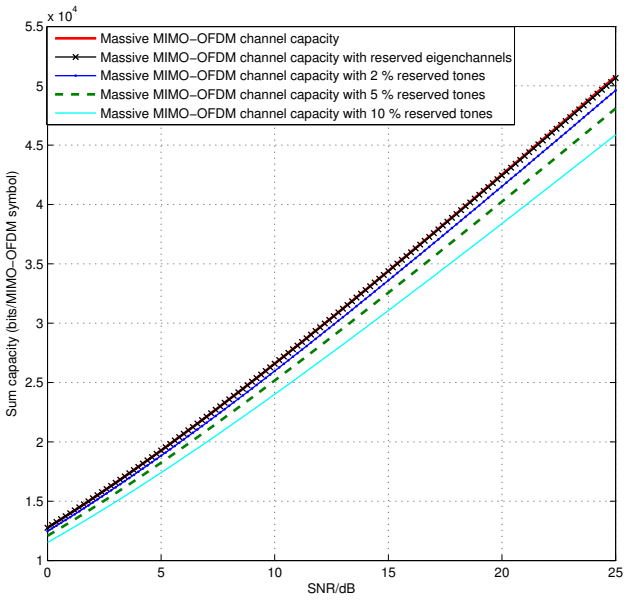

Fig. 5. Capacity curve of a 40×40 MIMO-OFDM channel with and without reserved eigenchannels and reserved tones (2%, 5%, and 10%), averaged over 100000 channel models

Thus, the relative mean transmit power increase ΔE is defined as [7]

$$\Delta \mathrm{E} = 10 \log_{10} \frac{\mathrm{E}\{||\tilde{\mathbf{s}}^i + \tilde{\mathbf{r}}^i||_2^2\}}{\mathrm{E}\{||\tilde{\mathbf{s}}||_2^2\}} \,, \qquad (27)$$

where $\mathrm{E}\{||\tilde{\mathbf{s}}||_2^2\} = \sigma^2$ and $\mathrm{E}\{||\tilde{\mathbf{s}}^i + \tilde{\mathbf{r}}^i||_2^2\}$ are the average powers at the ith iteration.
Figure 6 shows the mean power increase ΔE per iteration in dB for different PAR target values. Obviously, the increase in average power is high for low PAR target values, as more peaks are approximated. However, the increase in the average power is still marginal, in a range of one-two tenth of a decibel for as many as 10 iterations.
The simulation results under different mean power constraints are shown in Fig. 7. A gain of approximately 5.2 dB and 4.8 dB is obtained at 10^{-5} for a slight increase in the mean power, i.e., ΔE = 0.5 dB and 0.4 dB, respectively.

VI. Conclusions

We presented a novel PAR reduction algorithm for OFDM based massive MIMO systems in point-to-point scenarios. The proposed least-squares iterative approach exploits the massive degree of freedom (DoF) of a massive MIMO by reserving the weakest eigenchannels. These eigenchannels are used to estimate and model the peak excursion values exceeding a chosen PAR target value. This model function is then subtracted from the original signal in time domain for PAR reduction. We proposed two different approaches to estimate and model the peak values, which are actually equivalent.
Simulation results shows that a gain of approximately 5.0 dB

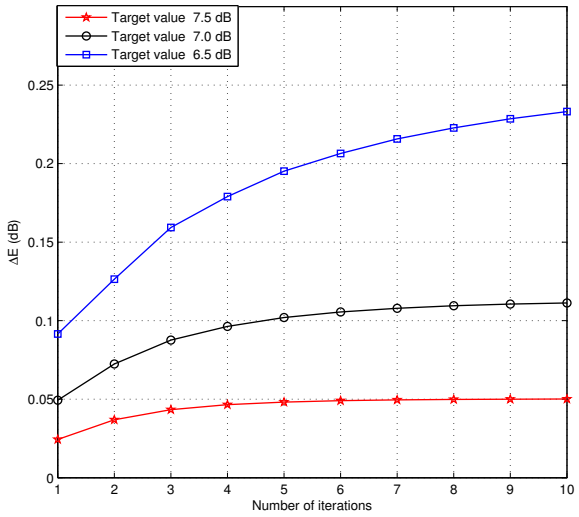

Fig. 6. Mean power increase ΔE, in dB, per iteration for different target values, weighting factor $\gamma = 30$

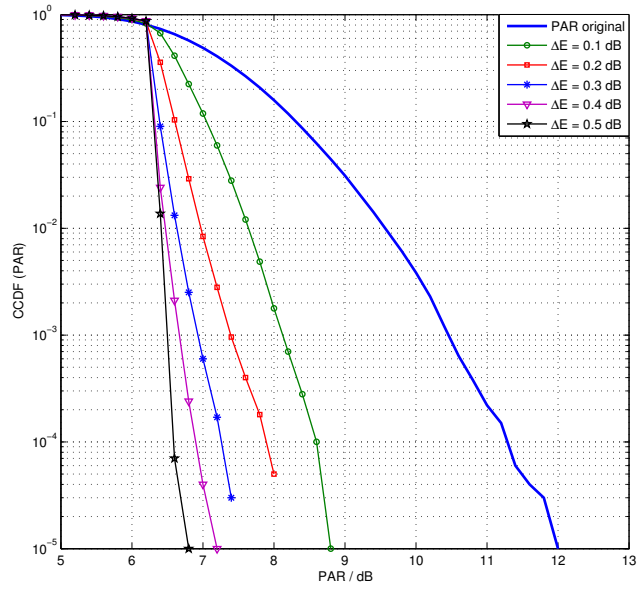

Fig. 7. CCDF(PAR) of the LS algorithm with a weighting factor γ = 30, for a PAR target value of 6.5 dB with different mean power constraints ΔE

is obtained with as few as 8 iterations at a CCDF of 10^{-5}. This gain is obtained at the expense of a minor increase in average power and almost no loss in the channel capacity.

REFERENCES

[1] S.H. Han, and J.H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Communications*, pp. 56–64, April 2005.

[2] T.L. Marzetta, "Noncooperative cellular wireless with unlimited number of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, No. 11, pp. 3590–3600, November 2010.

[3] E.G. Larson, F. Tufvesson, O. Edfors, and T.L. Marzetta, "Massive MIMO for next generation wireless systems," *arXiv:1304.6690v2*, May 2013.

[4] C. Siegl and R.F.H. Fischer, "Peak-to-average ratio reduction in multi-user OFDM," *International Symposium on Information Theory*, Nice, France, June 24-29, 2007.

[5] R.F.H. Fischer and, M. Hoch "Peak-to-average ratio reduction in MIMO-OFDM," *IEEE International Conference on Communication*, Glasgow, Scotland, June 24-28, 2007.

[6] C. Siegl and R.F.H Fischer, "Partial transmit sequences for peak-to-average power ratio reduction in multiantenna OFDM," *EURASIP Journal on Wireless Communications and Networking*, vol. 2008.

[7] J. Tellado, *Peak-to-average power reduction for multicarrier modulation*, Ph.D. thesis, Stanford University, 1999.

[8] W. Henkel, A. Wakeel, and M. Taseska, "Peak-to-average ratio reduction with tone reservation in multi-user and MIMO-OFDM," *1st IEEE International Communication Conference China (ICCC)*, Beijing, 2012.

[9] S.K. Muhammad and E.G. Larsson, "Per antenna constant envelop precoding for large multi-user MIMO systems," *arXiv:1111.3752v1*, January 2012.

[10] C. Studes and E.G. Larsson, "PAR-aware larege-scale multi-user MIMO-OFDM downlink," *in Proc. of the 9th International Symposium on Wireless Communication Systems (ISWCS)*, Paris, France, August 2012.

[11] A. Goldsmith, *Wireless Communications*, Cambridge University Press 2005.

[12] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecomm. ETT*, vol. 10, no. 6, pp.585-596, November 1998.