PAPER

Joint Sequence Design for Robust Channel Estimation and PAPR Reduction for MIMO-OFDM Systems

Chin-Te CHIANG[†*a)], *Nonmember and* **Carrson C. FUNG**[†b)], *Member*

SUMMARY A joint superimposed sequence design, known as SuperImposed sequence for PAPR Reduction, or SIPR, using per-tone affine precoding technique is proposed to jointly estimate MIMO-OFDM channels and reduce the peak-to-average power ratio (PAPR) for MIMO-OFDM systems. The proposed technique optimizes the trade-off between BER, MSE of the channel estimate, and PAPR reduction performance. Moreover, it does not require side information to be transmitted for the removal of the sequence at the receiver, and the transmit redundancy can be as small as 1 symbol/subcarrier. The superimposed sequence is designed by solving a convex quadratically constrained quadratic programming problem and has a computational complexity comparable to previous technique using linear programming. It is shown that SIPR can be regarded as a generalization of the popular tone reservation (TR) technique, and thus, is able to outperform TR in terms PAPR reduction performance, with less transmit overhead. Simulation results and transmit redundancy analysis of SIPR and TR are shown to illustrate the efficacy of the proposed scheme.

key words: *MIMO-OFDM, PAPR, superimposed sequence, channel estimation, affine precoder*

1. Introduction

MIMO-OFDM has proven to be the key enabler of high-speed, bandwidth efficient broadband communication systems, such as IEEE 802.16e, 3GPP LTE and the LTE-Advanced. The use of MIMO-OFDM allows such systems to combat the adverse effects brought about by wideband channels so that intersymbol interference can be easily removed without incurring huge penalty in computational complexity. Unfortunately, due to the use of the IFFT at the transmitter, the transmitted signal follows a Gaussian-like, time-domain waveforms with relatively high peak-to-average power ratio (PAPR) compared to single-carrier systems. To avoid signal distortion, highly linear, and consequently inefficient power amplifiers, are required. The problem is worsened when the OFDM systems are combined with MIMO as more RF chains are required for transmission where different antennas may exhibit large and varying degrees of PAPR.

It is well known that nonlinearity in high power amplifiers can distort the transmit signal. Such distortion can lead to undesirable spectral regrowth, thus interfering with signals in the neighboring subcarriers. Large amounts of distortion can also cause in-band self-interference, which increases bit error rate. Hence, it is customary for power amplifiers to operate with a certain power backoff, which is defined as the ratio of maximum saturation output power to lower average output power [1]. However, such backoff schemes lowers the efficiency of the power amplifier and increases overall power consumption.

Such problems can be avoided by employing clipping [2] and clipping with filtering [3]. Unfortunately, clipping is a highly nonlinear technique which, similar to backoff schemes, can also cause in-band and out-of-band (OOB) distortion, resulting in increased bit error rate (BER). An iterative PAPR reduction technique using clipping without in-band and OOB distortion has been reported in [4]. The method requires the use of null subcarriers otherwise it will incur BER degradation or OOB distortion. Although increase in the number of iterations reduces the PAPR, it also increases the signal power in the unused subcarriers. Multiple signal representation techniques such as partial transmit sequence (PTS) [5]–[7], selected mapping (SLM) [8], [9], and interleaving [10]–[12] have also been proposed as distortionless techniques to reduce the PAPR. Unfortunately, either side information has to be transmitted, or an increase in computational complexity at the receiver is needed, to allow for the recovery of the data. Tone reservation (TR) techniques [13] are data-dependent methods which add signal to the information-bearing signal in order to lower the PAPR. The TR method does not distort the original signal as the added signal is injected into set of subcarriers that have been reserved for PAPR reduction. This, of course, lowers the spectral efficiency of the system. The TR method was extended by [14] such that the power injected into the reserved subcarriers for PAPR reduction is formulated as a power allocation problem which can be efficiently solved using linear programming. TR methods using virtual subcarriers [15] and single reserved subcarrier [16] were proposed as modified versions of the original TR method in [13] to alleviate the problem of data rate loss. Combination of TR and clipping methods have also been proposed in [17]. Recently, a combined channel estimation and PAPR reduction technique was proposed by [18] in which phase information of the pilot, together with TR, were used to mitigate PAPR of the information-bearing signal for MIMO-OFDM system with null subcarriers. An overview of previous proposed PAPR reduction techniques are given in [19], [20].

Herein, a superimposed sequence design, known as SuperImposed sequence for PAPR Reduction, or SIPR, using

Manuscript received January 8, 2013.
Manuscript revised May 27, 2013.

[†]The authors are with the Dept. of Electronics Engineering, National Chiao Tung University, Hsinchu, 300 Taiwan.
[*]Presently, with the Intellectual Property Office, Ministry of Economic Affairs, Taipei, Taiwan.
a) E-mail: jeremiah1214@gmail.com
b) E-mail: c.fung@ieee.org

DOI: 10.1587/transcom.E96.B.2693

per-tone affine precoding technique is proposed to reduce PAPR for MIMO-OFDM systems and estimate spatially correlated MIMO-OFDM channels even if spatial correlation uncertainty exists. The proposed technique can easily trade-off between mean-squared error (MSE) of the channel estimate and PAPR reduction performance. Moreover, it does not require side information to be transmitted for the removal of the sequence at the receiver, and the transmit redundancy can be as small as 1 symbol/subcarrier. The superimposed sequence is designed by solving a quadratically constrained quadratic programming (QCQP) problem whose computational complexity is comparable to previous technique using linear programming. It is shown that SIPR can be regarded as a generalization of the popular tone reservation (TR) technique, and thus, is able to outperform TR in terms PAPR reduction performance, with less transmit overhead. Simulation results and transmit redundancy analysis of SIPR and TR are shown to illustrate the efficacy of the proposed scheme. The rest of the paper is organized as follows. The system model and a detailed description of SIPR are given in Sect. 2, followed by the simulation results in Sect. 3. The paper is concluded in Sect. 4.

Notations: Upper (lower) bold face letters denote matrices (column vectors). ^{H} denotes Hermitian transpose, ^{T} denotes transposition, and * denotes conjugation. $E[\cdot]$ denotes expectation. $\mathcal{N}(\boldsymbol{A})$ denotes the nullspace of $\mathbf{A}$. The operation, $vec(\mathbf{A})$, forms a column vector by vertically stacking the column vectors of $\mathbf{A}$. tr($\mathbf{A}$) denotes the trace of the matrix $\mathbf{A}$. $diag(x)$ denotes a diagonal matrix with x on its main diagonal, $\mathbf{I}_N$ denotes an $N\times N$ identity matrix, $\mathbf{0}_{M\times N}$ denotes an $M \times N$ all zero matrix, and $\mathbf{1}_N$ denotes an $N \times 1$ vector containing all 1's. $\otimes$ denotes the Kronecker product. $\|\mathbf{A}\|_F$ denotes the Frobenius norm of the matrix $\mathbf{A}$. $\|\mathbf{a}\|_\infty$ denotes the ∞-norm of the vector $\mathbf{a}$. $\mathbf{A}(m_i, :)$ denotes a row vector extracted from the m_ith row of $\mathbf{A}$. $\mathbf{A}\left(m_i : m_s : m_f, :\right)$ denotes a matrix whose elements are taken from the m_ith row to the m_fth row of $\mathbf{A}$ with m_s increment. $\mathbf{a} \preceq \mathbf{b}$ denotes the element-wise less-than operator.

2. Methodology

2.1 System Model

Consider an affine precoded MIMO-OFDM system with N_t transmit and N_r receive antennas as shown in Fig. 1. N_f-point DFT and IDFT are adopted. $\mathbf{u}(k) = \begin{bmatrix} u(kN_s) & u(kN_s+1) & \cdots & u(kN_s+N_s-1) \end{bmatrix}^T$ is the information-bearing signal vector, with k denoting the block index and N_s denoting the block size. Without loss of generality, $N_t = 2$ and 4 are assumed such that 2×2 and 4×4 orthogonal SFBC are used, respectively. The SFBC encodes two adjacent OFDM subcarriers at a time for a total of $K \geq N_t$ symbols. Note that the use of SFBC is not mandatory as it is only used to ease the decoding process. The output signal from the ith subcarrier at the ℓth antenna can then be grouped together to form $\mathbf{x}_{i,\ell} \in \mathbb{C}^K$, for $i =$

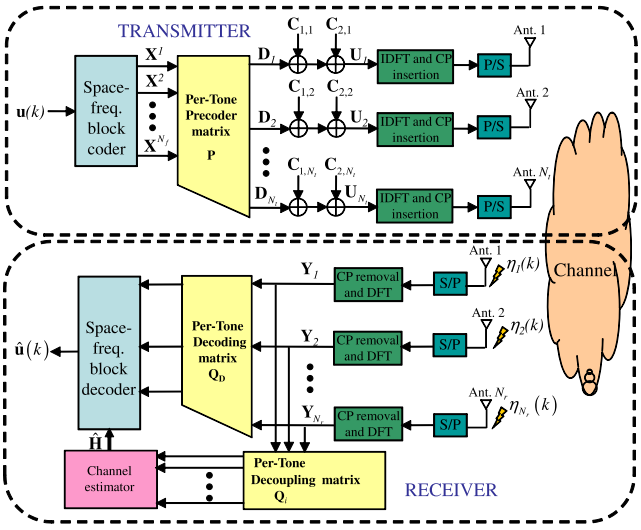


Fig. 1 Block diagram of MIMO-OFDM transceiver with per-tone affine precoder.

$1, 2, \ldots, N_f$. Defining $\mathbf{X}^i \triangleq \begin{bmatrix} \mathbf{x}_{i,1} & \mathbf{x}_{i,2} & \cdots & \mathbf{x}_{i,N_t} \end{bmatrix}^T \in \mathbb{C}^{N_t\times K}$. The information-bearing signal can then be written as an $N_fN_t \times N_fK$ complex-valued matrix

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}^1 & & & \\ & \mathbf{X}^2 & & \\ & & \ddots & \\ & & & \mathbf{X}^{N_f} \end{bmatrix}. \tag{1}$$

$\mathbf{X}$ is then post-multiplied by the precoder matrix $\mathbf{P} = \begin{bmatrix} \mathbf{P}_1^T & \mathbf{P}_2^T & \cdots & \mathbf{P}_{N_f}^T \end{bmatrix}^T \in \mathbb{C}^{N_fK\times(K+L)}$ to form

$$\mathbf{G} = \mathbf{XP} = \begin{bmatrix} \mathbf{X}^1\mathbf{P}_1 \\ \mathbf{X}^2\mathbf{P}_2 \\ \vdots \\ \mathbf{X}^{N_f}\mathbf{P}_{N_f} \end{bmatrix} \in \mathbb{C}^{N_fN_t\times(K+L)}, \tag{2}$$

where $\mathbf{P}_i = \begin{bmatrix} \mathbf{p}_{i,1} & \mathbf{p}_{i,2} & \cdots & \mathbf{p}_{i,K} \end{bmatrix}^T \in \mathbb{C}^{K\times(K+L)}$ is the per-tone affine precoder matrix which adds $L \geq N_t$ redundant symbols to the signal at the ith subcarrier. $\mathbf{p}_{i,j} \in \mathbb{C}^{K+L}$ denotes the precoder vector for $j = 1, 2, \ldots K$. The post-multiplication of the ith per-tone affine precoder to the data matrix $\mathbf{X}^i$ in (2) runs in contrast to conventional precoding techniques where the precoding matrix is pre-multiplied to the data matrix. It is clear that $\mathbf{x}_{i,\ell}^T\mathbf{P}_i$ belongs to the range space of $\mathbf{P}_i^T$, which as will be explained in the sequel, eases the removal of the superimposed sequences as the sequences and data belong to different orthogonal subspaces. As explained in [23], [24], removal of the sequences will be more difficult if the precoding matrix is pre-multiplied to the data matrix as the information-bearing signal belongs to the range space of the channel, which cannot guarantee to be orthogonal to the spaces occupied by the sequences. For ease of derivation, the precoded data shall be presented as

$$\mathbf{D}_\ell = \mathbf{G}\left(\ell : N_t : \ell + (N_f-1)N_t, :\right) \in \mathbb{C}^{N_f\times(K+L)}. \tag{3}$$

Defining the $N_fN_t \times (K+L)$ complex-valued aggregate precoded data matrix $\mathbf{D} = \begin{bmatrix} \mathbf{D}_1^T & \mathbf{D}_2^T & \cdots & \mathbf{D}_{N_t}^T \end{bmatrix}^T$, the $N_fN_t \times (K+L)$ complex-valued superimposed training (SIT) sequence matrix, $\mathbf{C}_1 = \begin{bmatrix} \mathbf{C}_{1,1}^T & \mathbf{C}_{1,2}^T & \cdots & \mathbf{C}_{1,N_t}^T \end{bmatrix}^T$, as proposed in [24], and the complex-valued SIPR matrix $\mathbf{C}_2 = \begin{bmatrix} \mathbf{C}_{2,1}^T & \mathbf{C}_{2,2}^T & \cdots & \mathbf{C}_{2,N_t}^T \end{bmatrix}^T \in \mathbb{C}^{N_fN_t \times (K+L)}$ of the same size, where $\mathbf{C}_{1,\ell} = \begin{bmatrix} \mathbf{c}_{1,1,\ell} & \cdots & \mathbf{c}_{1,N_f,\ell} \end{bmatrix}^T \in \mathbb{C}^{N_f \times (K+L)}$ and $\mathbf{C}_{2,\ell} = \begin{bmatrix} \mathbf{c}_{2,1,\ell} & \cdots & \mathbf{c}_{2,N_f,\ell} \end{bmatrix}^T \in \mathbb{C}^{N_f \times (K+L)}$ with $\mathbf{c}_{1,i,\ell}, \mathbf{c}_{2,i,\ell} \in \mathbb{C}^{K+L}$ denoting the SIT and SIPR vector for the ℓth antenna at the ith subcarrier, respectively. The superimposed sequence matrices will then be added to the precoded data before the IFFT resulting in

$$\mathbf{U} = \mathbf{C} + \mathbf{D} = \mathbf{C}_1 + \mathbf{C}_2 + \mathbf{D} \in \mathbb{C}^{N_tN_f \times (K+L)}, \tag{4}$$

as illustrated in Fig. 1.

Combining IFFT, cyclic prefix (CP) insertion, CP removal, and FFT with the frequency-selective MIMO channel results in an equivalent channel matrix

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} & \cdots & \mathbf{H}_{1N_t} \\ \mathbf{H}_{21} & \ddots & & \mathbf{H}_{2N_t} \\ \vdots & & \ddots & \vdots \\ \mathbf{H}_{N_r1} & \cdots & \cdots & \mathbf{H}_{N_rN_t} \end{bmatrix} \in \mathbb{C}^{N_fN_r \times N_fN_t}, \tag{5}$$

where $\mathbf{H}_{m\ell} \in \mathbb{C}^{N_f \times N_f}$ is a diagonal matrix containing the Fourier coefficients of the channel between the mth receive and ℓth transmit antenna. Thus, the receive data after CP removal can be written as

$$\mathbf{Y} = \mathbf{HU} + \mathbf{W} = \mathbf{H}(\mathbf{C}_1 + \mathbf{C}_2 + \mathbf{D}) + \mathbf{W}, \tag{6}$$

where $\mathbf{W} \in N_fN_r \times (K+L)$ denotes the channel noise matrix. It is assumed that the total signal power is normalized, i.e. $\sigma_{\mathbf{D}}^2 + \sigma_{\mathbf{C}_1}^2 + \sigma_{\mathbf{C}_2}^2 = 1$, where $\sigma_{\mathbf{D}}^2$, $\sigma_{\mathbf{C}_1}^2$ and $\sigma_{\mathbf{C}_2}^2$ denote the variance of $\mathbf{D}$, $\mathbf{C}_1$ and $\mathbf{C}_2$, respectively. It is further assumed that the information-bearing signal, and superimposed sequences all have zero-mean and are statistically independent from each other.

2.2 Proposed Sequence Design

It is clear from (6) that recovery of the data-bearing signal embedded inside $\mathbf{D}$ requires the removal of $\mathbf{C}_1$, $\mathbf{C}_2$ and $\mathbf{W}$, which can only be carried out after channel estimation has been performed using $\mathbf{C}_1$. However, $\mathbf{C}_2$ will have to be removed during the channel estimation and the detection phase as $\mathbf{C}_2$ is only exploited by the transmitter.

Defining

$$\mathbf{C}_1^i \triangleq \mathbf{C}_1\left(i : N_f : i + (N_t-1)N_f, :\right) \in \mathbb{C}^{N_t \times (K+L)} \tag{7}$$

$$\mathbf{C}_2^i \triangleq \mathbf{C}_2\left(i : N_f : i + (N_t-1)N_f, :\right) \in \mathbb{C}^{N_t \times (K+L)} \tag{8}$$

$$\mathbf{Y}^i \triangleq \mathbf{Y}\left(i : N_f : i + (N_r-1)N_f, :\right) \in \mathbb{C}^{N_r \times (K+L)} \tag{9}$$

$$\mathbf{W}^i \triangleq \mathbf{W}\left(i : N_f : i + (N_r-1)N_f, :\right) \in \mathbb{C}^{N_r \times (K+L)} \tag{10}$$

for ease of derivation, where ^{i} denotes the subcarrier index. Following the technique employed in [23], [24], the decoupling of the SIT sequence $\mathbf{C}_1^i$ can be achieved by post-multiplying $\mathbf{Y}^i$ by the per-tone decoder $\mathbf{Q}_i = \begin{bmatrix} \mathbf{q}_{i,1} & \mathbf{q}_{i,2} & \cdots & \mathbf{q}_{i,K+L} \end{bmatrix}^T \in \mathbb{C}^{(K+L) \times N_t}$, for $i = 1, 2 \ldots N_f$, where $\mathbf{Q}_i \in \mathcal{N}(\mathbf{P}_i)$, resulting in

$$\begin{aligned} \mathbf{Y}^i\mathbf{Q}_i &= \mathbf{H}^i\left(\mathbf{X}^i\mathbf{P}_i + \mathbf{C}_1^i + \mathbf{C}_2^i\right)\mathbf{Q}_i + \mathbf{W}^i\mathbf{Q}_i, \\ &= \mathbf{H}^i\left(\mathbf{C}_1^i + \mathbf{C}_2^i\right)\mathbf{Q}_i + \mathbf{W}^i\mathbf{Q}_i \end{aligned} \tag{11}$$

where $\mathbf{H}^i \in \mathbb{C}^{N_r \times N_t}$ is the matrix containing the Fourier coefficients at the ith subcarrier for all transmit and receive antennas. In other words, the SIT vector $\mathbf{c}_{1,i,\ell}$ should lie in the column space of $\mathbf{Q}_i$. Therefore, the condition that $\mathbf{C}_1^i\mathbf{P}_i^H = \mathbf{0}_{N_t \times K}$ guarantees the subspaces spanned by the vectors in $\mathbf{P}_i$ and $\mathbf{C}_1^i$ are complementary. In [24], the SVD of $\mathbf{C}_1^i$ is written as

$$\mathbf{C}_1^i = \mathbf{U}_{\mathbf{C}_1^i}\begin{bmatrix} \mathbf{\Sigma}_{\mathbf{C}_1^i} & \mathbf{0}_{N_t \times (K+L-N_t)} \end{bmatrix}\mathbf{V}_{\mathbf{C}_1^i}^H, \tag{12}$$

where $\mathbf{U}_{\mathbf{C}_1^i}$, $\mathbf{V}_{\mathbf{C}_1^i}$, and $\mathbf{\Sigma}_{\mathbf{C}_1^i}$ are the left and singular vector matrix of $\mathbf{C}_1^i$, and the invertible portion of the singular value matrix of $\mathbf{C}_1^i$, respectively. Notice that $\mathbf{C}_1^i\mathbf{P}_i^H = \mathbf{0}_{N_t \times K}$ if $\mathbf{V}_{\mathbf{C}_1^i} = \mathbf{U}_{\mathbf{Q}_i}$, where $\mathbf{U}_{\mathbf{Q}_i}$ is the eigenvector matrix of $\mathbf{Q}_i\mathbf{Q}_i^H$.

In the detection phase, the information-bearing signal can be recovered by post-multiplying $\mathbf{Y}$ by the decoding matrix $\mathbf{Q}_{D_i} = \mathbf{P}_i^H\left(\mathbf{P}_i\mathbf{P}_i^H\right)^{-1} \in \mathbb{C}^{(K+L) \times K}$, i.e.

$$\begin{aligned} \mathbf{Y}^i\mathbf{Q}_{D_i} &= \mathbf{H}^i\left(\mathbf{X}^i\mathbf{P}_i + \mathbf{C}_1^i + \mathbf{C}_2^i\right)\mathbf{Q}_{D_i} + \mathbf{W}\mathbf{Q}_{D_i}, \\ &= \mathbf{H}^i\left(\mathbf{X}^i + \mathbf{C}_2^i\mathbf{Q}_{D_i}\right) + \mathbf{W}\mathbf{Q}_{D_i}. \end{aligned} \tag{13}$$

Hence, a simple way to design $\mathbf{P}_i$ and $\mathbf{Q}_i$ is by extracting components off of an orthogonal matrix $\mathbf{O}_i \in \mathbb{C}^{(K+L) \times (K+L)}$, i.e.

$$\mathbf{P}_i = \sqrt{\frac{K+L}{K}}\mathbf{O}_i(1 : 1 : K, :) \in \mathbb{C}^{K \times (K+L)}, \tag{14}$$

$$\mathbf{Q}_i = \mathbf{O}_i^H((K+1) : 1 : (K+N_t), :) \in \mathbb{C}^{(K+L) \times N_t}. \tag{15}$$

Unfortunately, this is insufficient for the purpose herein as $\mathbf{C}_2^i$ remains. However, suppose

$$\mathbf{C}_2^i = \mathbf{B}_i\mathbf{M}_i, \tag{16}$$

where $\mathbf{B}_i = \begin{bmatrix} \mathbf{b}_{i,1} & \mathbf{b}_{i,2} & \cdots & \mathbf{b}_{i,N_t} \end{bmatrix}^T \in \mathbb{C}^{N_t \times (L-N_t)}$ has full rank, with $\mathbf{b}_{i,\ell} \in \mathbb{C}^{(L-N_t)}$, and

$$\mathbf{M}_i = \mathbf{O}_i((K+N_t+1) : 1 : (K+L), :) \tag{17}$$

denotes the $(L-N_t) \times (K+L)$ matrix extracted from the remaining rows of the orthogonal matrix $\mathbf{O}_i$ after $\mathbf{P}_i$ and $\mathbf{Q}_i$ have been extracted. That is, $\mathbf{C}_2^i$ is a linear combination of the remaining rows of the orthogonal matrix $\mathbf{O}_i$ after $\mathbf{P}_i$ and $\mathbf{Q}_i$ have been extracted. Therefore, the row vectors of $\mathbf{D}^i$, $\mathbf{C}_1^i$, and $\mathbf{C}_2^i$ are guaranteed to be orthogonal. Then, $\mathbf{C}_2^i$ will satisfy the constraints

$$\mathbf{C}_2^i\mathbf{Q}_i = \mathbf{0}_{N_t\times N_t}, \quad \text{and } \mathbf{C}_2^i\mathbf{Q}_{D_i} = \mathbf{0}_{N_t\times K}, \tag{18}$$

thus allowing the SIPR matrix $\mathbf{C}_2^i$ to be completely eliminated by $\mathbf{Q}_i$ and $\mathbf{Q}_{D_i}$ such that it will not interfere with either channel estimation nor symbol detection. It is clear that the existence of $\mathbf{C}_2^i$ does not impact the spectral efficiency of the system. Note that the number of superimposed subcarriers for $\mathbf{C}_2^i$ can be arbitrary chosen. For example, if there are N_f subcarriers, $\mathbf{C}_2$ is only added to the equally spaced $\frac{N_f}{\Delta_f}$ subcarriers, where Δ_f is a design parameter denoting the subcarrier spacing, with the constraint that N_f must be divisible by Δ_f. Note also that K and L can be different for each subcarrier as long as the sum of K and L for each subcarrier is identical. This is done so that it minimizes the impact on the spectral efficiency.

To account for uncertainty in the spatial correlation matrix, $\mathbf{C}_1^i$ can be designed using the RoMMSE technique proposed in [24] in order to improve MSE performance over that of [23] for spatially correlated MIMO channels. That is, $\mathbf{C}_1^i$ (or equivalently $\widetilde{\mathbf{C}}_1^i$) can be found by solving

$$\min_{\left\|\widetilde{\mathbf{C}}_1^i\right\|_F^2\le P_1}\ \max_{\|\mathbf{E}\|_F\le\varepsilon} \operatorname{tr}\left(\left[(\widehat{\mathbf{R}}+\mathbf{E})^{-1}+\widetilde{\mathbf{C}}_1^{iH}(\sigma_{\mathbf{ww}}^2\mathbf{I}_{N_rN_t})^{-1}\widetilde{\mathbf{C}}_1^i\right]^{-1}\right), \tag{19}$$

where $\widehat{\mathbf{R}}$ is the estimate of the spatial correlation matrix $\mathbf{R}=\mathbf{\Sigma}_t\otimes\mathbf{\Sigma}_r$, with $\mathbf{\Sigma}_t\in\mathbb{C}^{N_r\times N_t}$ and $\mathbf{\Sigma}_r\in\mathbb{C}^{N_r\times N_t}$ denoting the spatial correlation matrix at the transmitter and receiver, respectively. $\mathbf{E}=\mathbf{R}-\widehat{\mathbf{R}}\in\mathbb{C}^{N_rN_t\times N_rN_t}$ models the mismatch between $\mathbf{R}$ and $\widehat{\mathbf{R}}$. $\widetilde{\mathbf{C}}_1^i\triangleq\left(\mathbf{C}_1^i\mathbf{Q}_i\right)^T\otimes\mathbf{I}_{N_r}$, $\sigma_{\mathbf{ww}}^2$ denotes the noise power, and $P_1=\alpha N_rN_t(K+L)\sigma_{\mathbf{cc}}^2$, with $0\le\alpha\le 1$. $\sigma_{\mathbf{cc}}^2=\sigma_{\mathbf{C}_1}^2+\sigma_{\mathbf{C}_2}^2$ is the power of $\mathbf{C}_1+\mathbf{C}_2$. The selection of α will be described in Sect. 2.3.

The design of $\mathbf{C}_2$ requires the minimization of [21]

$$PAPR_{MIMO}=\max_{\ell=1,2\ldots N_t}\frac{\max|s_\ell(t)|^2}{E(|s_\ell(t)|^2)}, \tag{20}$$

where $s_\ell(t)$ denotes the signal in time domain at the ℓth antenna. From (4), the output from the IFFT becomes

$$\mathbf{S}=\mathbf{F}^H(\mathbf{C}_1+\mathbf{C}_2+\mathbf{D}), \tag{21}$$

where

$$\mathbf{F}=\begin{bmatrix}\mathbf{F}_1 & & & \\ & \mathbf{F}_2 & & \\ & & \ddots & \\ & & & \mathbf{F}_{N_t}\end{bmatrix}\in\mathbb{C}^{N_fN_t\times N_fN_t}. \tag{22}$$

In other words, $\mathbf{F}$ is a block diagonal matrix composed of N_t number of N_f-point DFT matrices $\mathbf{F}_\ell$, for $\ell=1,\ldots,N_t$. Defining

$$\mathbf{B}=\begin{bmatrix}\mathbf{b}_{1,1}^T & & & & \\ & \mathbf{b}_{2,1}^T & & & \\ & & \ddots & & \\ & & & \mathbf{b}_{N_f,1}^T & \\ \mathbf{b}_{1,2}^T & & & & \\ & \mathbf{b}_{2,2}^T & & & \\ & & \ddots & & \\ & & & \mathbf{b}_{N_f,2}^T & \\ \vdots & & & \vdots & \\ \mathbf{b}_{1,N_t}^T & & & & \\ & \mathbf{b}_{2,N_t}^T & & & \\ & & \ddots & & \\ & & & \mathbf{b}_{N_f,N_t}^T & \end{bmatrix}\in\mathbb{C}^{N_fN_t\times N_f(L-N_t)} \tag{23}$$

and $\mathbf{M}=\begin{bmatrix}\mathbf{M}_1^T & \mathbf{M}_2^T & \cdots & \mathbf{M}_{N_f}^T\end{bmatrix}^T\in N_f(L-N_t)\times(K+L)$, then $\mathbf{C}_2$ can be expressed as

$$\mathbf{C}_2=\mathbf{BM}. \tag{24}$$

Since $\mathbf{M}$ is created by stacking $\mathbf{M}_i$, for $i=1,\ldots,N_f$ and $\mathbf{M}_i$ is obtained as proposed in (17), $\mathbf{C}_2$ is completely parameterized by $\mathbf{B}$. The transmitted block matrix $\mathbf{S}$ can be vectorized as

$$\mathbf{s}=vec\left(\mathbf{F}^H(\mathbf{C}_1+\mathbf{BM}+\mathbf{D})\right). \tag{25}$$

Using the property $vec(\mathbf{AEB})=\left(\mathbf{B}^T\otimes\mathbf{A}\right)vec(\mathbf{E})$, and letting $\mathbf{A}=\mathbf{F}^H$, $\mathbf{E}=\mathbf{C}_1+\mathbf{BM}+\mathbf{D}$, and $\mathbf{B}=\mathbf{I}_{K+L}$, then the PAPR reduction problem can be formulated as

$$\begin{aligned}&\min_{\|\mathbf{C}_2\|_F^2\le\widetilde{P}_2}\|\mathbf{s}\|_\infty\\ &=\min_{\|\mathbf{C}_2\|_F^2\le\widetilde{P}_2}\left\|\left(\mathbf{I}_{K+L}\otimes\mathbf{F}^H\right)vec\left(\mathbf{C}_1+\mathbf{C}_2+\mathbf{D}\right)\right\|_\infty\\ &=\min_{\|\mathbf{B}\|_F^2\le\widetilde{P}_2}\left\|\left(\mathbf{I}_{K+L}\otimes\mathbf{F}^H\right)vec\left(\mathbf{C}_1+\mathbf{BM}+\mathbf{D}\right)\right\|_\infty,\end{aligned} \tag{26}$$

where (26) is obtained by using (24), and

$$\begin{aligned}\|\mathbf{C}_2\|_F^2&=\sum_i\operatorname{tr}\left(\mathbf{C}_2^i\mathbf{C}_2^{iH}\right)=\sum_i\operatorname{tr}\left(\mathbf{B}_i\mathbf{M}_i\mathbf{M}_i^H\mathbf{B}_i^H\right),\\ &=\sum_i\operatorname{tr}\left(\mathbf{B}_i\mathbf{B}_i^H\right)\\ &=\operatorname{tr}\left(\mathbf{BB}^H\right)=\|\mathbf{B}\|_F^2.\end{aligned} \tag{27}$$

$\widetilde{P}_2\triangleq\frac{P_2}{\Delta_f}=\frac{(1-\alpha)N_fN_t(K+L)\sigma_{\mathbf{cc}}^2}{\Delta_f}$ is the power constraint on $\mathbf{C}_2$. Note that (26) can be easily recasted as a convex QCQP problem [25]

$$\begin{aligned}&\min_{\mathbf{B},t}\quad t\\ &s.t.-t\mathbf{1}_{N_s}\le\left(\mathbf{I}_{K+L}\otimes\mathbf{F}^H\right)vec\left(\mathbf{C}_1+\mathbf{BM}+\mathbf{D}\right)\le t\mathbf{1}_{N_s},\\ &\quad\|\mathbf{B}\|_F\le\widetilde{P}_2,\end{aligned} \tag{28}$$

where $N_s \triangleq N_f N_t(K+L)$ denotes the size of the vector $\mathbf{s}$. Assuming the primal-dual interior point method is used, and let $n = (L-N_t)N_tN_f$ to denote the number of variables, then the number of iterations required to solve the ε-solution of the above QCQP is $O(\ln n \ln(1/\varepsilon))$ [27]. Assuming each iteration requires $O\left(n^3\right)$ arithmetic operations, then a total of $O\left(n^3 \ln n \ln(1/\varepsilon)\right)$ arithmetic operations is required. Since the same interior point method can be used to solve for linear programming problems, hence, the number of iterations for the proposed algorithm is similar to that of [14].

2.3 Power Allocation

Performance of joint channel estimation and PAPR reduction depends on how much power is allocated to $\mathbf{C}_1$ and $\mathbf{C}_2$. Unlike the power allocation problem for the channel estimation problem considered in [24], which only involves power allocation for $\mathbf{C}_1$, which can be done by optimizing the effective SNR at the receiver. The effective SNR in the scenario considered herein is

$$SNR_{\mathit{eff}} = \frac{E\left[\left\|\widehat{\mathbf{H}}^i\mathbf{X}^i\right\|^2\right]}{E\left[\left\|\widetilde{\mathbf{H}}^i\mathbf{W}_T\mathbf{Q}_{D_i} + \widetilde{\mathbf{H}}^i\mathbf{X}^i + \mathbf{W}\mathbf{Q}_{D_i}\right\|^2\right]}, \tag{29}$$

where $\mathbf{W}_T \in \mathbb{C}^{N_fN_f\times(K+L)}$ denotes the clipping noise, which results from in-band distortion caused by the use of clipping (and filtering) [28] to deal with symbols which still exceed the maximum power of the (linear) amplifier. As noted in [26], when the system is experiencing severe fading, the degradation of the SNR will be predominantly due to the channel noise, not the clipping noise. Under this premise, the power allocation between the information-bearing signal $\mathbf{X}$ and the sum of the superimposed sequences $\mathbf{C}_1 + \mathbf{C}_2$ in (6) can be carried out by utilizing the scheme proposed in [24], i.e. without considering the contribution of $\mathbf{W}_T$. The power of the superimposed sequences is selected to be

$$\sigma^2_{\mathbf{cc}} = \frac{\delta\beta\sigma^2_{\mathbf{ww}} - \sqrt{\delta\gamma\beta\sigma^2_{\mathbf{ww}}\left(\delta-\gamma+\beta\sigma^2_{\mathbf{ww}}\right)}}{\delta\left(\beta\sigma^2_{\mathbf{ww}}-\gamma\right)}, \tag{30}$$

where $\delta \triangleq N_rN_t$, $\gamma \triangleq N_r\sigma^2_{\mathbf{ww}}(K/(K+L))$, and $\beta \triangleq N_rN_t/(N_r+N_t)$.

Since there is no optimal way to allocate power between $\mathbf{C}_1$ and $\mathbf{C}_2$, an ad hoc method shall be adopted. This can be carried out by utilizing assumed a priori knowledge about the clipping ratio, which is defined as the ratio of the maximum signal amplitude that is allowed to pass unclipped to the root-mean-square power of the unclipped signal [26]. Therefore, increase in the clipping ratio (CR) results in smaller clipping noise. Figure 2 shows MSE vs. α for clipping ratio equals to 6 dB and 0 dB and for SNR equals to 10 and 25 dB for 2×2 MIMO-OFDM system. The total number of redundant symbols $L = 4$, where 2 symbols are used for channel estimation, and the rest for PAPR reduction, and $\Delta_f = 4$. The MSE reaches an optimal value at

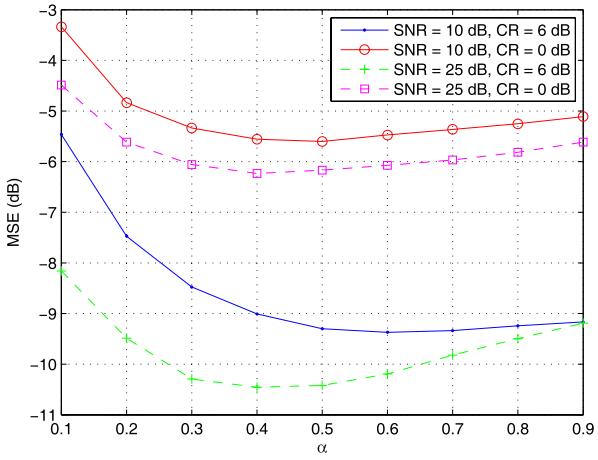


Fig. 2 MSE vs. α for 2×2 MIMO-OFDM system. $K = 10$, $L = 4$, $\Delta_f = 4$, $N_f = 512$.

$\alpha = 0.6$ for CR = 6 dB and $\alpha = 0.5$ for CR = 0 dB when SNR = 10 dB, and $\alpha = 0.4$ for CR = 6 dB and 0 dB when SNR = 25 dB. It is obvious as the level of clipping noise increases, more power is needed in $\mathbf{C}_2$ (smaller α) to further reduce the PAPR. Also, as SNR increases, the value of α that would be required to achieve the smallest MSE decreases as less power is needed in $\mathbf{C}_1$ for channel estimation. Since CR = 6 dB is typically encountered in actual system, $\alpha = 0.6$ is chosen in the next section for some of the simulations. Defining $w_1 \triangleq \alpha\sigma^2_{\mathbf{cc}}$ and $w_2 \triangleq (1-\alpha)\sigma^2_{\mathbf{cc}}$ as the normalized power allocated for $\mathbf{C}_1$ and $\mathbf{C}_2$, respectively (normalized by $N_fN_t(K+L)$). Consequently, with SNR = 10 dB, the normalized power allocated for channel estimation for 2×2 MIMO-OFDM system is $w_1 = (0.6)(\sigma^2_{\mathbf{cc}}) = 0.19$ and for PAPR reduction is $w_2 = (1-\alpha)(\sigma^2_{\mathbf{cc}}) = (1-0.6)(\sigma^2_{\mathbf{cc}}) = 0.13$. Note that for 4×4 MIMO system, w_1 and w_2 were computed to be 0.18 and 0.12, respectively. Since $\alpha = 0.6$ is chosen based on given set of design and simulation parameters, performance loss in MSE or PAPR reduction is expected when a different set is used, e.g. when the SNR is not equal to 10 dB. The choice of parameters will be further discussed in the sequel.

3. Simulation Results

4-QAM modulation is used to modulate the information-bearing signals in all simulations. For 2×2 system, Alamouti SFBC is used to provide transmit diversity. A half-rate complex orthogonal SFBC [29] is used for the 4×4 system.

3.1 2×2 MIMO-OFDM System

The simulation results in this section are generated using a MIMO-OFDM system with $N_t = N_r = 2$. Figure 3 shows the CCDF performance of the proposed SIPR scheme with different frequency spacings using the power allocation scheme described in Sect. 2.3. The figure also includes the CCDF performance for a system using only the robust superimposed training sequence channel estimator RoMMSE

proposed in [24] and one that does not use any superimposed sequence (No SIT). SNR = 10 dB, CR = 6 dB, $N_f = 512$, and $K = 10$ are used throughout all CCDF simulations in this subsection. The number of redundant vectors L equals to 4. For 2×2 SIPR, two of the redundant vectors are used for channel estimation with the rest allocated for PAPR reduction. Furthermore, $w_1 = 0.19$ and $w_2 = 0.13$. From the figure, it is clear that PAPR reduction performance improves as Δ_f decreases. This is reasonable since more symbols are allocated for PAPR reduction when smaller Δ_f is used. Unfortunately, even when $\Delta_f = 1$, the SIPR based system exhibits higher PAPR than the No SIT based system when $PAPR_0 \gtrsim 10.5$ dB. This is due mainly to the addition of $\mathbf{C}_1$, which pushed the CCDF curve for $\Delta_f = 16$ all to the way to $PAPR_0 = 17$ dB. This can be remedied by either increasing w_2 and/or L.

To increase PAPR reduction performance, w_1 is decreased to 0.1 and w_2 is increased to 0.2. $N_f = 512$. The CCDF results are shown in Fig. 4. Comparing Figs. 3 and 4, it is clear that SIPR with $\Delta_f = 2$ is now able to outperform RoMMSE and No SIT (for $PAPR_0 \lesssim 12$ dB). When $\Delta_f = 1$, the PAPR performance for SIPR outperforms all the others. For example, at CCDF = 10^{-3}, $PAPR_0 = 8$ dB for $\Delta_f = 1$, but for $\Delta_f = 2$, $PAPR_0 = 11$ dB. To see the improvement in PAPR reduction more clearly, Fig. 5 overlays the CCDF curves with identical Δ_f but different power allocation. As expected, increasing w_2 from 0.13 to 0.2 improves the PAPR reduction performance. Unfortunately, decreasing w_1 from 0.19 to 0.1 undoubtedly decreases the MSE performance of the channel estimate as seen in Fig. 6 when MSE performance between the two different power allocation is compared. Note that $N_f = 128$ in this figure. BER performance is shown in Fig. 7 when $L = 4$, $\Delta_f = 1$ and 4, and $N_f = 128$. Since part of the power for $\mathbf{C}_1$ has now been reallocated to $\mathbf{C}_2$ in the SIPR algorithm, higher BER is observed when compared to the RoMMSE method as evidenced in Fig. 7.

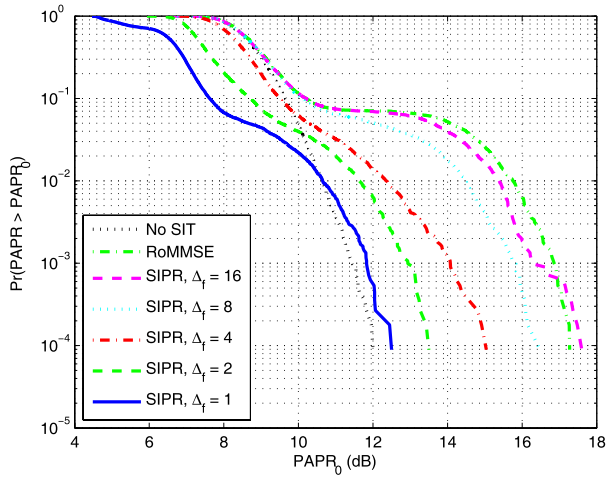


Fig. 3 CCDF comparison of proposed SIPR method with different frequency spacings to the RoMMSE method and without SIT for 2×2 MIMO-OFDM system. $L = 4$, CR = 6 dB, $\alpha = 0.6$ (equivalently, $w_1 = 0.19$ and $w_2 = 0.13$), $N_f = 512$.

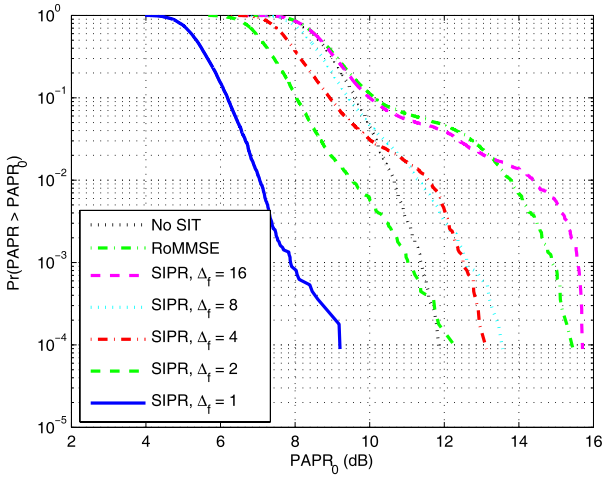
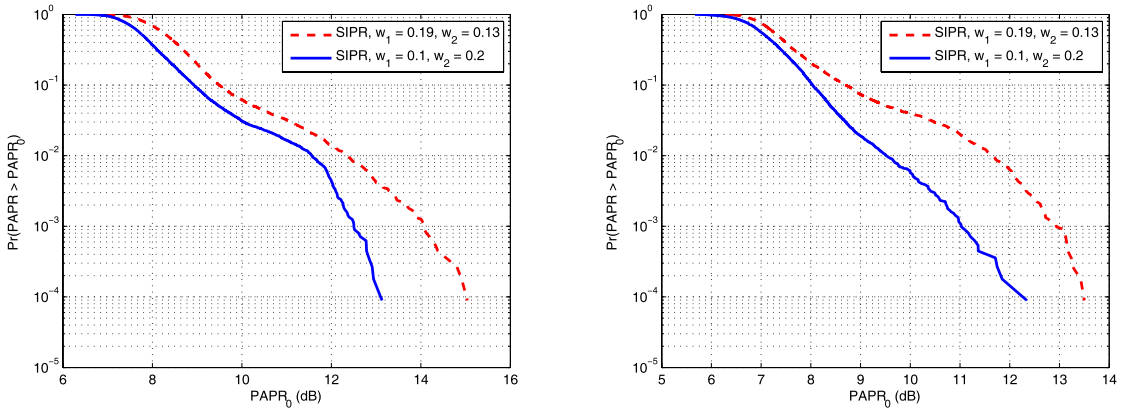


Fig. 4 CCDF comparison of proposed SIPR method with different frequency spacings to the RoMMSE method and without SIT for 2×2 MIMO-OFDM system. $L = 4$, CR = 6 dB, $w_1 = 0.1$ and $w_2 = 0.2$, $N_f = 512$.



(a) $\Delta_f = 4$. (b) $\Delta_f = 2$.

Fig. 5 CCDF comparison of SIPR with different power allocation for different Δ_f for 2×2 MIMO-OFDM system. $L = 4$, CR = 6 dB, $N_f = 512$.

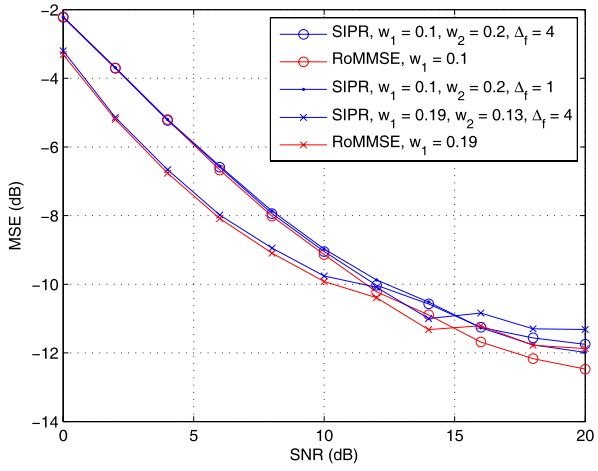


Fig. 6 MSE vs. SNR performance between RoMMSE and SIPR with different power allocation and different Δ_f for 2×2 MIMO-OFDM system. $L = 4$, CR = 6 dB, $N_f = 128$.

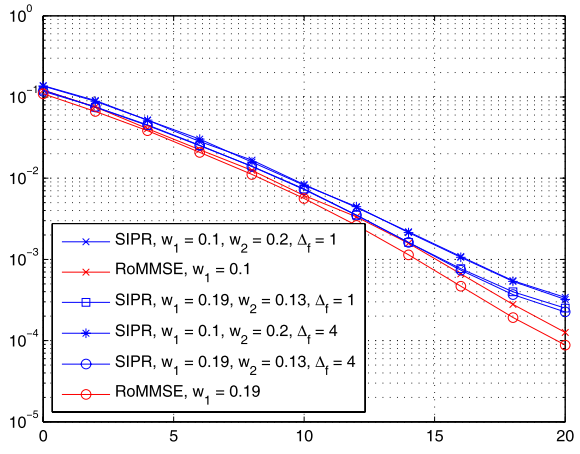
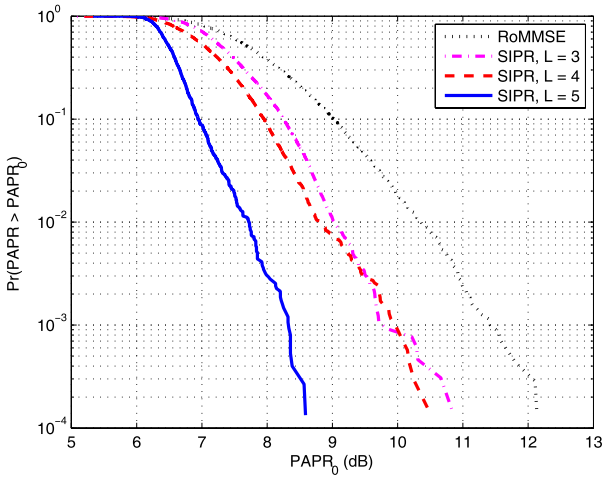
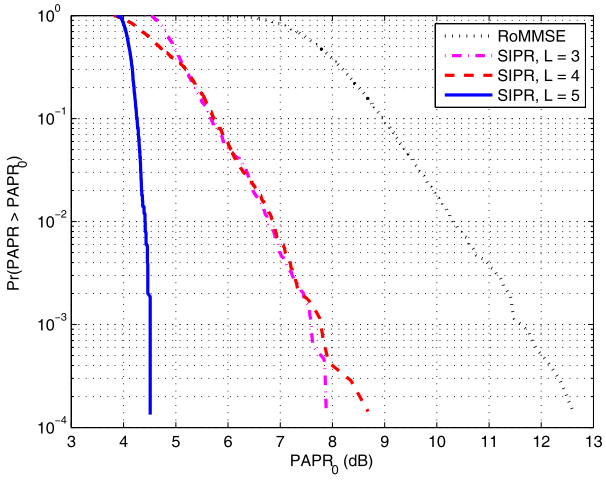


Fig. 7 BER vs. SNR performance between RoMMSE and SIPR with different power allocation for 2×2 MIMO-OFDM system. $L = 4$, CR = 6 dB, $N_f = 128$.



(a) $\Delta_f = 4$.



(b) $\Delta_f = 1$.

Fig. 8 CCDF comparison of proposed SIPR method with different number of redundant vectors L for 2 × 2 MIMO-OFDM system. CR = 6 dB, $w_1 = 0.1$, and $w_2 = 0.2$ for SIPR. $N_f = 128$.

This loss, however, is less than 1 dB at BER = 10^{-2}. Also, note that the MSE and BER do not change significantly between $\Delta_f = 1$ and 4 as changing Δ_f mainly affects PAPR reduction performance.

PAPR reduction performance of SIPR with different number of redundant vectors and RoMMSE is shown in Fig. 8, with $N_f = 128$. For all three SIPR curves, two redundant symbols are allocated to channel estimation, with the rest allocated for PAPR reduction. For example, when $L = 3$, one redundant symbol is allocated for PAPR reduction. When $L = 5$, three redundant symbols are allocated for PAPR reduction. The normalized power $w_1 = 0.1$ and $w_2 = 0.2$. When $L = 3$ and CCDF = 10^{-3}, the PAPR of SIPR outperforms that of the RoMMSE by 1.5 dB. When Δ_f decreases to 1, this gap increases to almost 4 dB. Furthermore, as the number of redundant vectors increases, the PAPR performance improves. At CCDF = 10^{-3} and $\Delta_f = 4$, the PAPR gap between $L = 5$ and 3 equals to 1.5 dB. When Δ_f decreases to 1, this gap increases to 3.5 dB. Such performance gain is expected because as L increases, the number of rows in $\mathbf{M}_i$ also increases, thereby expanding the dimension of the "PAPR reduction subspace", allowing more redundant symbols to be utilized for PAPR reduction.

PAPR reduction performance between tone reservation and SIPR is compared in Fig. 9 with $N_f = 128$. To make the comparison fair, $L = 2$ redundant symbols are added in the simulation for TR as this is the same number of redundant symbols used in the simulations for SIPR. 32 subcarriers, or 25% of the total number of subcarriers, were reserved for PAPR reduction in the TR simulation. This is comparable to the simulation conditions in [13]. $\Delta_f = 4$, 2, and 1 were used for SIPR. When CCDF equals 10^{-3}, SIPR with $\Delta_f = 1$ is able to outperform TR in terms of PAPR reduction by 0.7 dB. Moreover, the spectral efficiency of a SIPR system will be comparable to that of a "No SIT" system as the transmit redundancy can be as small as 1 symbol/subcarrier if $\mathbf{C}_1$

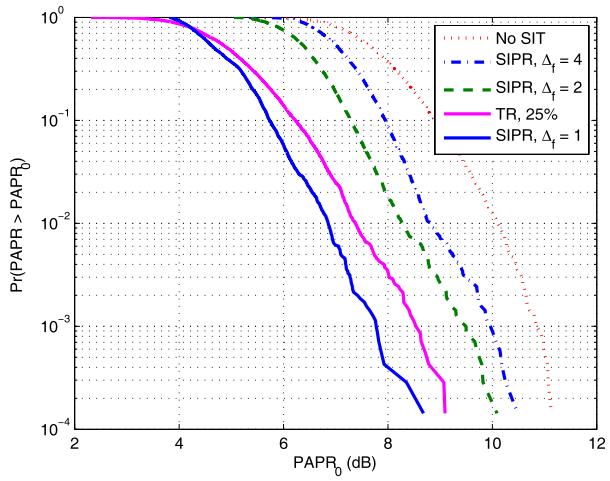


Fig. 9 CCDF comparison of SIPR and TR for 2×2 MIMO-OFDM system. $L = 4$, CR = 6 dB, $w_1 = 0.1$, $w_2 = 0.2$ for SIPR. $N_f = 128$.

is not needed for channel estimation. This is not the case for TR as dedicated subcarriers have to be utilized, further impacting the spectral efficiency. To gain further insight into the impact of the transmit redundancy, define the transmit efficiency as

$$\rho \triangleq \frac{\kappa}{\tau}, \tag{31}$$

where κ denotes the number of the transmitted symbols used for channel estimation and data detection, and τ denotes the total transmitted symbols. In Fig. 9, ρ is equal to 75% for the TR method because the number of reserved subcarriers used is 25% of the total number of subcarriers. However,

$$\begin{aligned}\rho &= \frac{\frac{N_f}{\Delta_f}K_1 + \left(N_f - \frac{N_f}{\Delta_f}\right)K_2}{K_2 N_f} \\ &= \frac{(32)(12) + (96)(14)}{(14)(128)} = 96.4\% \end{aligned} \tag{32}$$

for the proposed SIPR scheme with $\Delta_f = 4$. $K_1 = 12$ is the number of symbols per subcarrier which are transmitted when $\mathbf{C}_2$ is employed on that particular subcarrier, and $K_2 = 14$ is the number of symbols per subcarrier which are transmitted when $\mathbf{C}_2$ is not used. Table 1 summarizes the transmit efficiency of methods used in Fig. 9, and it is clear that the proposed scheme outperforms the tone reservation method not only in PAPR reduction performance, but also in terms of transmit efficiency, because less overhead is needed. This performance gain can be attributed to the fact that SIPR is able to spread the PAPR reduction symbols across the entire transmission spectrum without affecting the transmit efficiency. This allows the proposed method more capability to reduce the PAPR compared to the TR technique. Hence, SIPR can be viewed as a generalization of TR. It is important to note that decreasing Δ_f from 4 to 1 does not impact the MSE performance of SIPR as long as the normalized power w_1 allocated to $\mathbf{C}_1$ remains the same. This is evident in Fig. 6.

Table 1 Transmit efficiency of TR and SIPR method.

	TR, 25%	SIPR, $\Delta_f = 4$	SIPR, $\Delta_f = 2$	SIPR, $\Delta_f = 1$
ρ	75%	96.4%	92.9%	85.7%

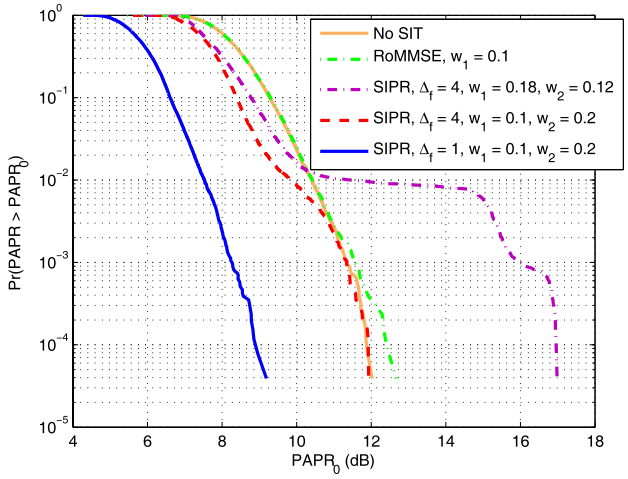


Fig. 10 CCDF comparison of SIPR with different power allocation for 4×4 MIMO-OFDM system. $L = 8$, CR = 6 dB, $N_f = 128$.

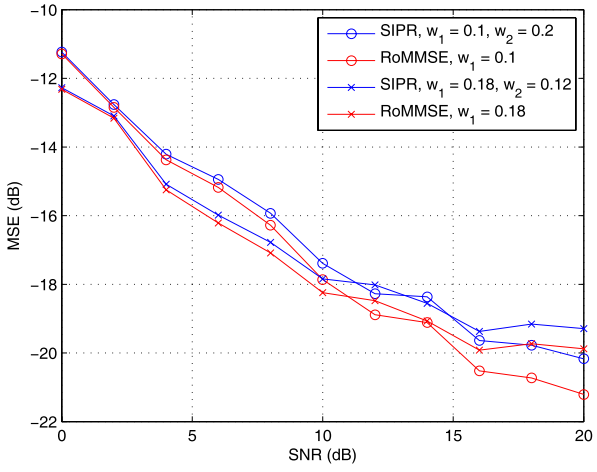


Fig. 11 MSE vs. SNR performance between RoMMSE and SIPR with different power allocation for 4×4 MIMO-OFDM system. $L = 8$, CR = 6 dB, $\Delta_f = 4$, $N_f = 128$.

3.2 4×4 MIMO-OFDM System

Figure 10 shows the CCDF result for 4×4 system with different power distribution for SIPR for $\Delta_f = 1$ and 4. SNR = 10 dB, CR = 6 dB, and $N_f = 128$. The number of redundant vectors L equals to 8. For 4×4 SIPR, four of the redundant vectors are used for channel estimation with the rest allocated for PAPR reduction. Similar to the behavior seen in the simulation results of the 2×2 system, as w_2 increases, the PAPR reduction performance of the SIPR based system improves. Such improvement is also observed when Δ_f decreases from 4 to 1 as more symbols are used for PAPR reduction.

Figure 11 shows the MSE vs. SNR performance be-

tween SIPR and RoMMSE for two different values of w_1 and w_2. $L = 8$, $\Delta_f = 4$, CR = 6 dB, and $N_f = 128$. Similar to the performance seen in Fig. 6, as w_1 increases (and w_2 decreases for SIPR), the MSE performance of SIPR and RoMMSE improves. This is because more power is allocated for robust channel estimation. Also, since all the power allocated to the superimposed training sequence in the RoMMSE algorithm is dedicated to robust channel estimation, the RoMMSE method always outperforms the proposed SIPR method in terms of MSE.

4. Conclusion

A joint sequence design called SIPR is proposed for robust channel estimation and PAPR reduction for MIMO-OFDM systems. The proposed SIPR method is able to accurately estimate spatially correlated MIMO-OFDM channels even if the spatial correlation estimate is inaccurate, and also significantly lower the PAPR without significantly impacting the spectral efficiency. When channel estimation using superimposed training sequence is not required, the transmit redundancy has been shown to be as small as 1 symbol/subcarrier. The computational complexity is shown to be comparable to a previously proposed method that uses linear programming. Simulation results show that the proposed scheme is able to outperform the tone reservation method in terms of PAPR reduction and transmit efficiency.

Acknowledgements

The work described in this paper has been supported by the National Science Council Grant 101-2219-E-009-019.

References

[1] F. Danilo-Lemoine, et al., "Power backoff reduction techniques for generalized multicarrier waveforms," EURASIP Journal on Wireless Communications and Networking, vol.2008, pp.1–13, Jan. 2008.

[2] A. Bahai, M. Singh, A. Goldsmith, and B. Saltzberg, "A new approach for evaluating clipping distortion in multicarrier systems," IEEE J. Sel. Areas Commun., vol.20, no.5, pp.1037–1046, June 2002.

[3] X. Li and L.J. Cimini, Jr., "Effects of clipping and filtering on the performance of OFDM," IEEE Commun. Lett., vol.2, no.5, pp.131–133, May 1998.

[4] D. Guel and J. Palicot, "OFDM PAPR reduction based on nonlinear functions without BER degradation and out-of-band emission," Proc. Intl. Conf. on Signal Processing Systems, pp.167–171, May 2009.

[5] S.H. Müller and J.B. Huber, "OFDM with reduced peak-to-average power ratio by optimum combination of partial transmit sequences," Electron. Lett., vol.33, no.5, pp.368–369, Feb. 1997.

[6] A.D.S. Jayalath and C. Tellambura, "Adaptive PTS approach for reduction of peak-to-average power ratio of OFDM signal," Electron. Lett., vol.36, no.14, pp.1225–1228, July 2000.

[7] S.H. Han and J.H. Lee, "PAPR reduction of OFDM signals using a reduced complexity PTS technique," IEEE Signal Process. Lett., vol.11, no.11, pp.887–890, Nov. 2004.

[8] R.W. Baüml, R.F.H. Fisher, and J.B. Huber, "Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping," Electron. Lett., vol.32, no.22, pp.2056–2057, Oct. 1996.

[9] H. Breiling, S.H. Müller-Weinfurtner, and J.B. Huber, "SLM peak-power reduction without explicit side information," IEEE Commun. Lett., vol.5, no.6, pp.239–241, June 2001.

[10] P. Van Eetvelt, G. Wade, and M. Tomlinson, "Peak to average power reduction for OFDM schemes by selective scrambling," Electron. Lett., vol.32, no.21, pp.1963–1964, Oct. 1996.

[11] G.R. Hill, M. Faulkner, and J. Singh, "Reducing the peak-to-average power ratio in OFDM by cyclically shifting partial transmit sequences," Electron. Lett., vol.36, no.6, pp.560–561, March 2000.

[12] A.D.S. Jayalath and C. Tellambura, "Reducing the peak-to-average power ratio of orthogonal frequency division multiplexing signal through bit or symbol interleaving," Electron. Lett., vol.36, no.13, pp.1161–1163, June 2000.

[13] J. Tellado and J.M. Cioffi, "Peak power reduction for multicarrier transmission, Proc. IEEE GlobeCom Communications Theory Mini-Conf, pp.219–224, 1998.

[14] J. Tellado and J.M. Cioffi, "Efficient algorithms for reducing PAR in multicarrier systems," Proc. IEEE Intl. Sym. on Information Theory, Aug. 1998.

[15] S. Hu, et al., "Analysis of tone reservation method for WiMAX system," Proc. Intl. Sym. on Communications and Information Technologies, pp.498–502, 2006.

[16] Y. Son, C.H. Nam, and H.S. Lee, "An approach for PAPR reduction baesd on tone reservation method," Proc. 6th IEEE Consumer Communications and Networking Conference, pp.1–2, 2009.

[17] V. Cuteanu, A. Isar, and C. Nafornita, "PAPR reduction of OFDM signals using sequential tone reservation-clipping hybrid scheme," Proc. Signal Prcoessing and Applied Mathematics for Electronics and Communications Workshop, pp.49–52, Cluj-Napoca, Romania, Aug. 2011.

[18] E.C. Manasseh, S. Ohno, and M. Nakamoto, "Combined channel estimation and PAPR reduction technique for MIMO-OFDM systems with null subcarriers," EURASIP Journal on Wireless Communications and Networking, June 2012.

[19] 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, April 2005.

[20] 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, June 2008.

[21] R.F.H. Fischer and M. Hoch, "Peak-to-average power ratio reduction in MIMO OFDM," IEEE Intl. Conf. on Communications, June 2007.

[22] D.H. Pham and J.H. Manton, "Orthogonal superimposed training on linear precoding: A new affine precoder design," Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications, pp.445–449, June 2005.

[23] V. Nguyen, et al., "Optimal superimposed training design for spatially correlated fading MIMO channels," IEEE Trans. Wireless Commun., vol.7, no.8, pp.3206–3217, Aug. 2008.

[24] C.-T. Chiang and C.C. Fung, "Robust training sequence design for spatially correlated MIMO channel estimation," IEEE Trans. Vehicular Technol., vol.60, no.7, pp.2882–2894, Sept. 2011.

[25] S. Boyd and L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.

[26] K.R. Panta and J. Armstrong, "Effects of clipping on the error performance of OFDM in frequency selective fading channels," IEEE Trans. Wireless Commun., vol.3, no.2, pp.668–671, March 2004.

[27] A. Nemirovskii and K. Scheinberg, "Extension of Karmarkar's algorithm onto convex quadratically constrained quadratic problems," Mathematical Programming, vol.72, no.3, pp.273–289, 1996.

[28] J. Armstrong, "New OFDM peak-to-average power reduction scheme," Proc. IEEE Vehicular Technology Conf., Rhodes, Greece, May 2001.

[29] V. Tarokh, H. Jafarkhani, and A.R. Calderbank, "Space-time block codes from orthogonal designs," IEEE Trans. Inf. Theory, vol.45, no.5, pp.1456–1467, July 1999.

Chin-Te Chiang received his the B.S. and M.S. degrees in electronics engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2007 and 2010, respectively. His research interests include signal processing aspects of wireless communication systems and biomedical engineering. He is currently with the Intellectual Property Office, Ministry of Economic Affairs, in Taipei, Taiwan.

Carrson C. Fung received his B.S. degree from Carnegie Mellon University, Pittsburgh, PA, USA, in 1994, M.S. degree from Columbia University, New York City, NY, USA, in 1996 and Ph.D. degree at The Hong Kong University of Science and Technology in 2005, all in electrical engineering. He was the recipient of the prestigious Sir Edward Youde Ph.D. Fellowship in 2001–2002. He has also been a Member of Technical Staff at AT&T and Lucent Technologies Bell Laboratories, Holmdel, NJ, USA, from 1994–1999, where he worked on video and audio coding. He was also a Researcher at the Hong Kong Applied Science and Technology Research Institute (ASTRI) in 2005, where he worked on MIMO-OFDM systems and a Senior DSP Engineer at Sennheiser Research Lab in Palo Alto, CA, USA, in 2006, where he worked on microphone and microphone array technologies. He has, since 2006, been an Assistant Professor with the National Chiao Tung University in Hsinchu, Taiwan. His research interests include statistical signal processing, communications, and optimization.