

# A First-Order Primal-Dual Method for Saddle Point Optimization of PAPR Problem in MU-MIMO-OFDM Systems

*Davinder SINGH, R.K. SARIN*

Dept. of Electronics and Communication Engineering, Dr. B.R. Ambedkar National Institute of Technology,  
Jalandhar, Punjab, India-144011

dsdavinder1@gmail.com, sarinrk@nitj.ac.in

Submitted June 29, 2017 / Accepted January 5, 2018

**Abstract.** This paper investigates the use of a particular splitting-based optimization technique for constrained  $l_\infty$ -norm based peak-to-average power ratio (PAPR) reduction problem in multiuser orthogonal frequency-division multiplexing (OFDM) based multiple-input multi-output (MIMO) systems. PAPR reduction and multi-user interference (MUI) cancelation are considered in a saddle-point formulation on the downlink of a multi-user MIMO-OFDM system and an efficient primal-dual hybrid gradient (PDHG) inspired algorithm with easy-to-evaluate proximal operators is developed. The proposed algorithm converges significantly faster to satisfactory solutions with much improved asymptotical convergence rate than existing methods. Numerical results illustrate the superior performance of the proposed algorithm over existing methods in terms of PAPR reduction for different MIMO configurations.

## Keywords

MIMO-OFDM, peak-to-average power ratio (PAPR) reduction, saddle point problem, convex optimization, proximity operator

## 1. Introduction

Very-large multi-user multiple-input multiple-output (MU-MIMO) system or massive MIMO, with hundreds of base station antennas, is an emerging technology to meet the explosively increasing demand of throughput. Marzetta [1] firstly conceived the idea of employing large number of transmit antennas at the base-station (BS), in co-located or distributed manner, to simultaneously serve dozens of user-equipments (UEs) in the same time-frequency source. Massive MIMO arrays of realistic physical sizes provide substantial improvement in spectral and energy efficiency compared to the conventional small-scale MIMO systems. Moreover, the use of large-scale antenna arrays yields favorable propagation where the channel vectors between

the base station and the user terminals become mutually orthogonal. Hence, the simple linear signal processing methods become efficient and optimal for multi-user interference suppression [2]. Massive MIMO systems are promising candidates for 5G [3].

Practical wireless channels generally suffer from frequency-selective fading. The intersymbol interference resulting from the frequency-selective fading can be compensated by orthogonal frequency division multiplexing (OFDM). OFDM divides the entire bandwidth of a wide-band channel into a set of orthogonal narrowband sub-channels such that each individual sub-channel is exposed to flat fading rather than frequency selective fading.

MIMO and OFDM technology are widely used at the same time in the physical layer for wireless communication systems [4], [5]. However, the OFDM signal incurs high peak-to-average power ratio (PAPR) of the system due to the fact that phases of sub-carriers may combine in a constructive or destructive manner. Constructive superposition of subcarriers via inverse fast Fourier transform (IFFT) operation will result in a signal with high value of envelope peaks. To avoid high out-of-band radiation caused by high side lobes of modulated subcarriers, and inter-modulation distortion among subcarriers, we require power amplifiers and digital-to-analog converters (DAC) with large linear dynamic ranges, which is costly and power-inefficient [6]. Therefore, signals with low PAPR become more important to make massive MIMO systems affordable as well as power efficient.

Several schemes have been proposed to handle high PAPR in single-input single-output (SISO) OFDM systems, such as companding [7], [8], precoding [9], clipping [10], partial transmit sequence (PTS) [11], tone reservation [12], selected mapping (SLM) [13] and active constellation extension (ACE) [14]. Although it is easy to extend most of these schemes to single user MIMO systems [15–18], their applicability to the multi-user (MU) MIMO systems is not straightforward. This is due to the fact that, since the user terminals are spatially distributed, joint processing of the signals is only feasible at the transmitter end. So it is

more challenging to apply PAPR reduction schemes in multiuser scenarios. In [19], the null-spaces of massive MU-MIMO-OFDM channels are exploited based on a linear constrained  $l_\infty$  optimization problem to improve the PAPR. Convex optimization via fast iterative truncation algorithm (FITRA) jointly performs pre-coding, OFDM modulation, and PAPR minimization. It is shown theoretically that constant-envelope signals with PAPR close to unity can be obtained for infinitely large antenna limit. FITRA results in dramatic PAPR improvement (around 10 dB) for massive MIMO systems but shows the limited PAPR improvement when applied to the small scale MIMO configurations, i.e.,  $4 \times 2$  or  $8 \times 2$  MIMO. Similarly, in [20], the perturbation signals lying within the null spaces of the associated channel matrices are added to the pre-coded signals in order to reduce the PAPRs of the transmitted signals via convex optimization problem. Unlike [19], where the proximal operator of the  $l_\infty$ -norm is basically a clipping operator, an efficient algorithm is developed by resorting to the alternating direction method of multipliers (ADMM). A PAPR reduction scheme analogous to the tone reservation was proposed in [21], where a peak clipping scheme is used for some transmit antennas at base station, while other antennas are reserved to compensate for the distortions due to peak clipping. This antenna-reservation scheme has a low computational overhead but can achieve only a moderate PAPR minimization. Moreover, the antennas reserved to compensate for clipping distortion may also experience large PAPRs. Methods proposed in [22], [23] formulate the PAPR reduction problem as approximate message passing (AMP)-based Bayesian inference problem that use priors to encourage constant magnitude solutions. The Bayesian approach in [22] treats MUI cancelation as an underdetermined linear inverse problem and offers better PAPR reduction than the FITRA algorithm [19], with much lower computational complexity.

In this paper, we present a generalized usage of the primal-dual hybrid gradient (PDHG) algorithm [24] for joint consideration of PAPR reduction and multi-user precoding in multiuser MIMO-OFDM systems. PAPR reduction problem is formulated as a saddle-point problem and is solved by using a variant of the PDHG algorithm. The algorithm progresses by taking forward and backward proximal steps that alternately maximize and minimize a constrained form of the saddle function. The associated proximal operator for the convex non-differentiable  $l_\infty$ -norm is implemented in linear time with low computational cost.

The remainder of the paper is structured as follows: in Sec. 2, the description of the system model, basic assumptions and introduction to the PAPR reduction problem, are provided. Section 3 describes the formulation of PAPR reduction as a saddle-point problem along with the development of modified PDHG algorithm. In next section simulation results are presented and conclusions are drawn in Sec. 5.

*Notations:* Throughout the paper, the lowercase and uppercase symbols such as “ $\mathbf{x}$ ” and “ $\mathbf{X}$ ” respectively are used to represent column vectors and matrices. The transpose is denoted as  $(\cdot)^T$  and conjugate transpose as  $(\cdot)^H$ . The inner-product of two vectors  $\mathbf{x}$  and  $\mathbf{y}$  will be denoted by  $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$ . Moreover  $\|\mathbf{x}\|_1$ ,  $\|\mathbf{x}\|_2$  and  $\|\mathbf{x}\|_\infty$  will be used to denote the  $l_1$ -norm,  $l_2$ -norm and  $l_\infty$ -norm of a vector  $\mathbf{x}$ .  $\Re\{\mathbf{x}\}$  and  $\Im\{\mathbf{x}\}$  respectively represent the real and imaginary part of a vector  $\mathbf{x}$ . The  $U \times U$  unitary discrete Fourier transform, the  $N \times N$  identity matrix and  $M \times N$  zero matrix are represented as  $\mathbf{F}_U$ ,  $\mathbf{I}_N$  and  $\mathbf{0}_{M \times N}$ , respectively. The symbol  $\otimes$  denotes the Kronecker product. The notation  $\text{abs}(\mathbf{x})$  denotes the component-wise application of the absolute value to a vector  $\mathbf{x}$ .

## 2. System Description

### 2.1 System Model

Consider a multiuser MIMO-OFDM downlink system depicted in Fig. 1 that consists of a base station (BS) having  $N$  transmit antennas. The BS simultaneously serves  $M$  single antenna equipped terminals, where  $M \ll N$ . For a total of  $U$  OFDM tones (subcarriers), the  $M \times 1$  signal vector  $\mathbf{s}_u$  contains information at  $u$ -th tone for each of the  $M$  terminals. The subcarriers are divided into two complementary sets  $\phi$  and  $\phi^C$ . The set  $\phi$  contains the data transmission tones and  $\phi^C$  indexes the inactive guard band tones, such that for each tone  $u \in \phi^C$ ,  $\mathbf{s}_u = \mathbf{0}_{M \times 1}$ .

To eliminate the multi-user interference (MUI) at the receivers, the information symbols on the  $u$ -th subcarrier are linearly precoded as

$$\mathbf{p}_u = \mathbf{W}_u \mathbf{s}_u \quad (1)$$

where  $\mathbf{p}_u \in \mathbb{C}^{N \times 1}$  is the pre-coded vector and  $\mathbf{W}_u \in \mathbb{C}^{N \times M}$  represents the precoding matrix for the  $u$ -th OFDM tone. To eliminate MUI, the widely used precoding schemes are maximum ratio transmission (MRT), zero forcing (ZF), and minimum mean-square error (MMSE) precoding [25]. For known channel matrix  $\mathbf{H} \in \mathbb{C}^{M \times N}$  at the transmitter, the standard ZF linear pre-coder can be written as

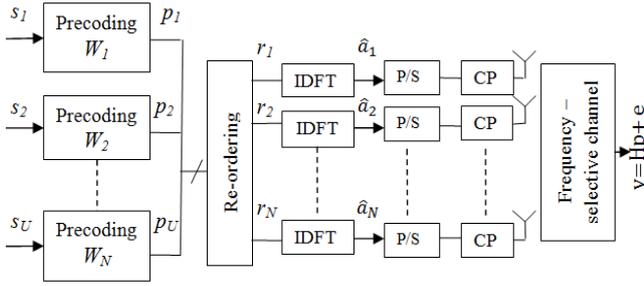
$$\mathbf{W}_u^{\text{ZF}} \propto \mathbf{H}_u^H (\mathbf{H}_u \mathbf{H}_u^H)^{-1}. \quad (2)$$

The normalized precoded vectors  $\mathbf{p}_u, \forall u$  are then reordered to  $N$  transmit antennas according to the following mapping:

$$[\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N] = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_U]^T \quad (3)$$

where  $U$ -dimensional vector  $\mathbf{r}_n$  contains the frequency-domain signal to be transmitted from the  $n$ th antenna.

Then an Inverse Discrete Fourier Transform (IDFT)  $F$  is applied on the precoded signal to obtain the time-domain signal, i.e.,



**Fig. 1.** System model of the MU-MIMO-OFDM downlink scenario, with  $U$  OFDM tones,  $N$  transmit antennas and  $M$  user-terminals.

$$\hat{\mathbf{a}}_n = \mathbf{F}_U^H \mathbf{r}_n. \quad (4)$$

After applying IDFT, a cyclic prefix is appended to the time domain samples of each antenna to avert intersymbol interference (ISI). Lastly, the time domain samples are transmitted through the wireless channel.

## 2.2 Peak-to-Average Power Ratio (PAPR) Reduction

The time domain samples  $\{\hat{\mathbf{a}}_n\}$  have large dynamic range due to the superposition of the individual subcarriers, which can be characterized by peak-to-average power ratio (PAPR) metric on the  $n$ -th antenna as [19]:

$$PAPR_n = \frac{2U \|\hat{\mathbf{a}}_n\|_\infty^2}{\|\hat{\mathbf{a}}_n\|_2^2} \quad (5)$$

where the  $l_\infty$ -norm is defined as  $\|\hat{\mathbf{a}}_n\|_\infty = \max\{\|\Re\{\hat{\mathbf{a}}_n\}\|_\infty, \|\Im\{\hat{\mathbf{a}}_n\}\|_\infty\}$ . The operator  $\|\cdot\|_\infty^2$  is used because RF-chains process and modulate the real and imaginary parts independently.

Alternative PAPR definitions also exist in the literature where  $l_\infty$ -norm instead of  $l_\infty$  and  $U$  instead of  $2U$  are used. The relation  $\frac{1}{2} \|\hat{\mathbf{a}}_n\|_\infty^2 \leq \|\hat{\mathbf{a}}_n\|_\infty^2 \leq \|\hat{\mathbf{a}}_n\|_\infty^2$  [26, Eq. 12] ensures that reducing the PAPR as defined in (5) also reduces an  $l_\infty$ -norm-based PAPR definition.

The best and worst case PAPR is bounded as

$$1 \leq PAPR_n \leq 2U. \quad (6)$$

Since the transmit antennas are more in number than the user terminals i.e.,  $N \gg M$  so there will be infinitely many precoded signals,  $\mathbf{p} \triangleq [\mathbf{p}_1^T, \mathbf{p}_2^T, \dots, \mathbf{p}_U^T]^T$  that satisfy the precoding constraint  $\mathbf{s} = \mathbf{H}\mathbf{p}$ . In this paper, we search for the precoded signal  $\mathbf{p}$  whose related time domain signals  $\{\hat{\mathbf{a}}_n\}$  result in low dynamic range and which also satisfy the following conditions to cancel out the MUI.

$$\begin{cases} \mathbf{s}_u = \mathbf{H}_u \mathbf{p}_u, \mathbf{u} \in \Phi, \\ \mathbf{0}_{N \times 1} = \mathbf{p}_u, \mathbf{u} \in \Phi^c. \end{cases} \equiv \bar{\mathbf{s}} = \bar{\mathbf{H}}\mathbf{p} \quad (7)$$

where  $\bar{\mathbf{s}} \in C^{UM \times 1}$  is the aggregation of all user data symbols and inactive tones, and  $\bar{\mathbf{H}}$  is a block diagonal matrix having the main diagonal blocks as  $\mathbf{H}_u, \forall u \in \Phi$  and  $\mathbf{I}_N, \forall u \in \Phi^c$ .

As defined in (3), the entries of each precoded vector are assigned to  $N$  transmit antennas through a linear transformation  $\mathbf{T}$  as:

$$\mathbf{r} = \mathbf{T}\mathbf{p}. \quad (8)$$

Using (4) and (7) we obtain

$$\bar{\mathbf{s}} = \bar{\mathbf{H}}\mathbf{T}^T \bar{\mathbf{F}}\hat{\mathbf{a}} \quad (9)$$

where  $\bar{\mathbf{F}} \triangleq \mathbf{I}_N \otimes \mathbf{F}_U$  and  $\hat{\mathbf{a}} \triangleq [\hat{a}_1^T, \hat{a}_2^T, \dots, \hat{a}_N^T]^T$ .

For symbol vector  $\bar{\mathbf{s}}$  the aim is to find time domain samples  $\hat{\mathbf{a}}$  that satisfy (9) such that each antenna emits signals with low PAPR. Inspired by [19], the PAPR minimax reduction problem can be framed as a  $l_\infty$ -norm based constrained optimization which reduces the largest PAPR existing among all the transmit antennas as

$$\min \|\hat{\mathbf{a}}\|_\infty \text{ subject to } \bar{\mathbf{s}} = \bar{\mathbf{H}}\mathbf{T}^T \bar{\mathbf{F}}\hat{\mathbf{a}}. \quad (10)$$

This problem can eventually be transformed into equivalent real-valued  $l_\infty$ -norm based problem  $(\mathbf{P}_\infty^\delta)$  by relaxing the equality constraint  $\bar{\mathbf{s}} = \bar{\mathbf{H}}\mathbf{T}^T \bar{\mathbf{F}}\hat{\mathbf{a}}$  to  $\|\bar{\mathbf{s}} - \bar{\mathbf{H}}\mathbf{T}^T \bar{\mathbf{F}}\hat{\mathbf{a}}\|_2 \leq \delta$  [27], i.e.

$$(\mathbf{P}_\infty^\delta) \min_{\mathbf{x} \in C^N} \|\mathbf{x}\|_\infty \text{ subject to } \|\mathbf{s} - \mathbf{H}\mathbf{x}\|_2 \leq \delta. \quad (11)$$

where

$$\mathbf{s} \triangleq \begin{bmatrix} \Re\{\bar{\mathbf{s}}\} \\ \Im\{\bar{\mathbf{s}}\} \end{bmatrix}, \mathbf{x} \triangleq \begin{bmatrix} \Re\{\hat{\mathbf{a}}\} \\ \Im\{\hat{\mathbf{a}}\} \end{bmatrix}, \bar{\mathbf{H}} \triangleq \begin{bmatrix} \Re\{\bar{\mathbf{H}}\mathbf{T}^T \bar{\mathbf{F}}\} & -\Im\{\bar{\mathbf{H}}\mathbf{T}^T \bar{\mathbf{F}}\} \\ \Im\{\bar{\mathbf{H}}\mathbf{T}^T \bar{\mathbf{F}}\} & \Re\{\bar{\mathbf{H}}\mathbf{T}^T \bar{\mathbf{F}}\} \end{bmatrix}.$$

## 3. Problem Formulation and Proposed Algorithm

The problem  $(\mathbf{P}_\infty^\delta)$  in (11) can be rewritten in the constrained form

$$(\mathbf{P}_\infty^\delta) \begin{cases} \min_{\mathbf{x} \in C^N} \|\mathbf{x}\|_\infty \\ \text{subject to } \mathbf{v} = \mathbf{s} - \mathbf{H}\mathbf{x}, \|\mathbf{v}\|_2 \leq \delta. \end{cases} \quad (12)$$

After introducing *Lagrange multiplier*  $\lambda \in C^M$  to enforce the equality constraint  $\mathbf{v} = \mathbf{s} - \mathbf{H}\mathbf{x}$ , the saddle-point formulation takes the following form:

$$\max_{\lambda \in C^M} \min_{\mathbf{x} \in C^N, \mathbf{v} \in C^M} \|\mathbf{x}\|_\infty + \lambda^T (\mathbf{H}\mathbf{x} - \mathbf{v} - \mathbf{s}) + \chi_\delta(\mathbf{v}) \quad (13)$$

where the characteristic function  $\chi_\delta(\mathbf{v})$  used to remove the constraint  $\|\mathbf{v}\|_2 \leq \delta$  is defined as follows:

$$\chi_\delta(\mathbf{v}) = \begin{cases} 0, & \text{if } \|\mathbf{v}\|_2 \leq \delta \\ \infty, & \text{otherwise.} \end{cases} \quad (14)$$

### 3.1 PDHG-PROXINF Algorithm

The primal-dual hybrid gradient (PDHG) [24], [28], [29] is an efficient splitting method that works directly with the original saddle-point problem by alternating between primal and dual variables. In our case, the saddle point of the minimax Lagrangian formulation (13) corresponds to a minimizer of  $(P_\infty^\delta)$ . In order to find the saddle point of (13), we propose a variant of standard PDHG algorithm, referred to as the PDHG-PROXINF, in Algorithm 1.

The parameters  $\tau_k$ , and  $\sigma_k$  are step sizes of the primal and dual steps, respectively. During any iteration  $k$ , PDHG-PROXINF updates both a primal (in step 1-2) and dual variable (in step 4-5) using a combination of forward and backward (or proximal) steps. In steps (1-2), PDHG-PROXINF updates the  $x \in C^N$  by first taking a gradient descent step followed by the proximal step involving  $\|x\|_\infty$ . Whereas the dual variable  $\lambda \in C^M$  is updated through a gradient ascent step 4 followed by the backward step 5. While the proposed algorithm is nearly as effective as constant non-adaptive parameter versions of PDHG, well selected adaptive step sizes  $\tau_k$ ,  $\sigma_k$  from recently developed method in [30] can enhance the performance. The FITRA [19] requires a manually chosen regularization parameter to achieve balance between PAPR reduction and multiuser interference, whereas PDHG-PROXINF is free of this issue.

---

#### Algorithm 1: PDHG-PROXINF Algorithm

---

**Input:**  $x_0 \in C^N, \tau_k, \sigma_k > 0$   
**while not converged do**  
 1.  $\hat{x}_{k+1} = x_k - \tau_k H^T \lambda_k$   
 2.  $x_{k+1} = \arg \min_{x \in C^N} \left\{ \|x\|_\infty + \frac{1}{2\tau_k} \|x - \hat{x}_{k+1}\|_2^2 \right\}$   
      $= \arg \min_{x \in C^N} \left\{ \|x\|_\infty + \langle \lambda_k, Hx \rangle + \frac{1}{2\tau_k} \|x - x_k\|_2^2 \right\}$   
 3.  $v_{k+1} = \delta(v_k - \tau_k \lambda_k) / \max\{\text{abs}(v_k - \tau_k \lambda_k), \delta\}$   
 4.  $\hat{\lambda}_{k+1} = \lambda_k + \sigma_k Hx_{k+1}$   
 5.  $\lambda_{k+1} = \hat{\lambda}_{k+1} - \sigma_k (v_{k+1} + s)$

---

### 3.2 Evaluation of the $l_\infty$ Proximal Operator

Algorithm 1 takes a gradient descent step initially followed by the computation of the proximal operator associated with  $l_\infty$ -norm in step 2. Proximal operator does not have a simple closed-form solution; however, it can be computed explicitly as explained in the following section.

The proximal operator  $\text{prox}_{\tau\| \cdot \|_\infty}(\mathbf{w}) = \text{prox}_\tau(x_k - \tau_k H^T \lambda_k)$  defined in step 2, finds a point close to the minimizer of  $\|x\|_\infty$  without straying too far from the starting point  $w = x_k - \tau_k H^T \lambda_k$  [31]. The step size  $\tau_k$  controls the extent

to which the proximal operator maps points towards the minimum of  $\|x\|_\infty$ . In plain terms,  $\text{prox}_{\tau\| \cdot \|_\infty}(\mathbf{w})$  is the function that maps the vector  $\mathbf{w} \in C^N$  to the unique solution of

$$\arg \min_{x \in C^N} \left\{ \|x\|_\infty + \frac{1}{2\tau_k} \|x - \mathbf{w}\|_2^2 \right\}. \quad (15)$$

The proximal operator  $\text{prox}_{\tau\| \cdot \|_\infty}(\mathbf{w})$  and the projector  $\text{Proj}_{\{\| \cdot \|_1 \leq \tau\}}(\mathbf{w})$  onto the closed ball of radius  $\tau$  of the dual  $l_1$ -norm, satisfy the relation

$$\text{prox}_{\tau\| \cdot \|_\infty}(\mathbf{w}) = \mathbf{w} - \text{Proj}_{\{\| \cdot \|_1 \leq \tau\}}(\mathbf{w}), \quad (16)$$

where

$$\text{Proj}_{\{\| \cdot \|_1 \leq \tau\}}(\mathbf{w}) = \left\{ \arg \min_x \|x - \mathbf{w}\|_2 \text{ subject to } \|x\|_1 \leq \tau \right\}. \quad (17)$$

Ignoring the trivial case  $\|\mathbf{w}\|_1 \leq \tau$ , there exists for  $\tau$  each a Lagrange multiplier  $\gamma(\tau)$  such that

$$\arg \min_x \frac{1}{2} \|x - \mathbf{w}\|_2^2 + \gamma(\tau) \|x\|_1, \quad (18)$$

shares the same solution as (17).

As shown by Chambolle et al. [32], the solution of separable convex optimization problem (18) has a closed form accomplished by applying component-wise soft thresholding operation  $S_\gamma(\mathbf{w})$  to the vector  $\mathbf{w}$  for a properly chosen value of  $\gamma$ .

$$\text{Proj}_{\{\| \cdot \|_1 \leq \tau\}}(\mathbf{w}) = S_\gamma(\mathbf{w}). \quad (19)$$

Thus, the proximal operator associated with  $l_\infty$ -norm in (16) becomes  $\text{prox}_{\tau\| \cdot \|_\infty}(\mathbf{w}) = \mathbf{w} - S_\gamma(\mathbf{w})$  that can be computed by sorting the elements of  $\mathbf{w} \in C^N$  in decreasing order [33] followed by component-wise truncation, which can be performed in  $\mathcal{O}(n \log n)$  expected time. An improvement of this algorithm was proposed in [34], that avoids having to sort the entire vector  $\mathbf{w}$ ; since only the largest elements of  $\mathbf{w}$  are involved in the determination of  $\gamma$ . This heap-sorting based approach reduces the complexity to  $\mathcal{O}(\log n)$ . We have exploited the linear-time  $\mathcal{O}(n)$  median-search like procedure [35] to compute the proximal operator. The overall algorithm to minimize (15) is detailed in Algorithm 2.

---

#### Algorithm 2: Procedure for evaluation of $\text{prox}_{\tau\| \cdot \|_\infty}(\mathbf{w})$

---

**Input:** A vector  $w \in C^N$  and a scalar  $\tau > 0$   
 $w_{\text{abs}} := \text{abs}(w)$   
**If**  $\|w_{\text{abs}}\|_1 \leq \tau$  **then** return  $w_{\text{abs}}$  [quick exit if  $w$  is feasible]  
**Initialize:**  $q := w_{\text{abs}}, c := 0$  [initialization]  
**While 1 do**  
    $k := \lceil \text{length}(q) / 2 \rceil$   
    $z_k := \text{upper - median}(q)$ ; [set the pivot]  
    $q_{\text{high}} := \text{find}(q < z_k)$ ; [partition around the pivot]  
    $q_{\text{low}} := \text{find}(q > z_k)$

---

---

```

 $c_{\text{low}} := \|q_{\text{low}}\|_1 + z_k;$ 
 $\psi(z_k) := c + c_{\text{low}} - k \cdot z_k; \quad [\text{soft-threshold upto } k]$ 
If  $\psi(z_k) < \tau$  then
     $q := q_{\text{low}};$ 
else
     $z_{k+1} := \max(q_{\text{high}})$ 
     $\psi(z_{k+1}) := c + c_{\text{low}} - k \cdot z_{k+1}; \quad [\text{soft-threshold upto } k+1]$ 
    If  $\psi(z_{k+1}) < \tau$  then break  $[\text{check termination condition}]$ 
     $q := q_{\text{high}};$ 
     $c := c + c_{\text{low}};$ 
 $\gamma := z_k - (\tau - \psi(z_k)) / k; \quad [\text{compute Lagrange multiplier}]$ 
for  $k=1, \dots, N$  do
     $t_k := \min\{(w_{\text{abs}})_k, \gamma\} \text{sign}(w_k); \quad [\text{truncation operator}]$ 
return t
    
```

---

## 4. Simulation Results

To illustrate the PAPR reduction performance of PDHG-PROXINF algorithm, simulations are carried out for a massive MIMO scenario having  $N = 100$  transmit antennas and 10 single-antenna user terminals. OFDM modulation with  $U = 128$  tones and spectral map  $\varphi$  with 114 data tones for transmission are considered [36]. For the purpose of simulation, coded transmission is employed, i.e., the information bits for each user are encoded by a convolutional encoder (rate-1/2, generator polynomials  $[133_0, 171_0]$  and constraint length 7) [37] which are then interleaved, mapped to 16-QAM constellation, precoded, and finally transmitted over the assumed frequency-selective channel. Convolutional code-1/2 rate is opted since it provides significant coding gain and low PAPR values as compared to higher coding rates. The channel is modelled as a tap-delay line with  $D = 4$  taps and impulse response matrices  $\hat{H}_d, d = 1, \dots, D$ , have independent and identically distributed entries drawn from circularly-symmetric normal distribution. The frequency-domain response  $H_u$  on the  $u$ -th tone can be expressed as [38]:

$$H_u = \sum_{d=1}^D \hat{H}_d \exp\left(\frac{-j2\pi du}{U}\right). \quad (20)$$

After demodulation at each user terminal, a soft-input Viterbi decoder is used to regenerate the transmitted bits.

In the simulations, the probability of PAPR exceeding a threshold  $\text{PAPR}_0$  is considered as the measurement index which is described as complementary cumulative distribution function (CCDF) [39].

$$\text{CCDF}(\text{PAPR}_0) = \Pr(\text{PAPR} > \text{PAPR}_0). \quad (21)$$

Also, in order to evaluate the out-of-band radiation of the solution, we define the out-of-band (power) ratio as:

$$\text{OBR} = \frac{|\varphi| \sum_{u \in \varphi^c} \|p_u\|_2^2}{|\varphi^c| \sum_{u \in \varphi} \|p_u\|_2^2}. \quad (22)$$

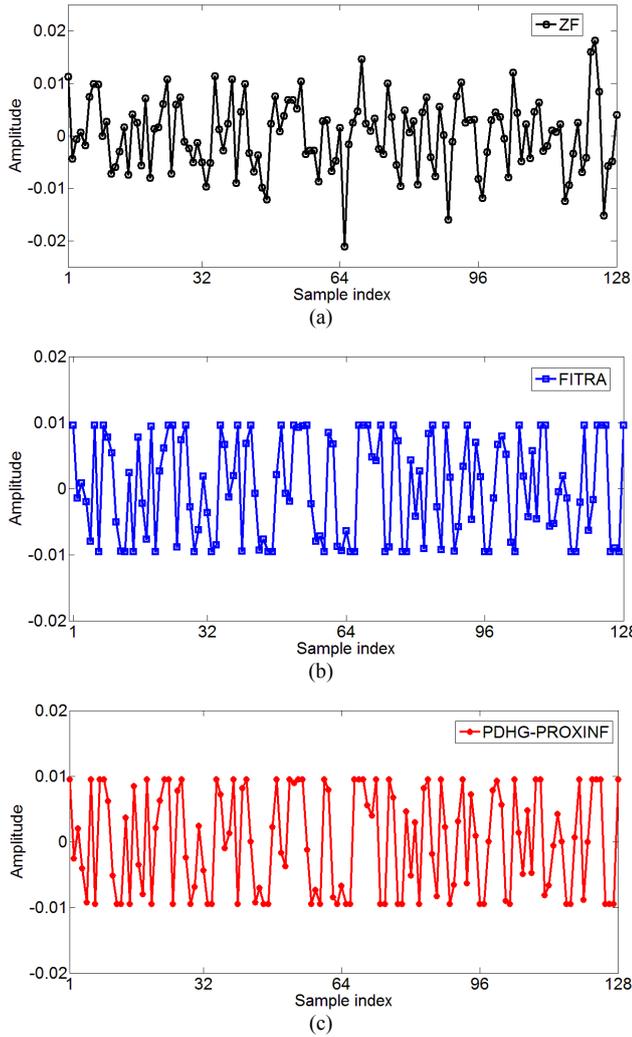
To evaluate the performance of the proposed primal-dual hybrid gradient proximal infinity (PDHG-PROXINF) algorithm, we compare it with the fast iterative truncation algorithm (FITRA) [19] and the conventional zero-forcing (ZF) precoding scheme for 1000 independent trials. It is worth mentioning here that the time taken for per packet transmission and overall simulation will not be the same for the considered iterative schemes. CPU computational time depends on the tolerance level and maximum number of iterations taken. The maximum number of iterations is, however, explicitly stated for each iterative scheme in the ensuing simulation results.

Firstly, we examine the signals estimated by respective schemes. Figure 2 depicts the real part of the first antenna's time-domain samples (i.e.  $\hat{a}_1$ ) obtained by all considered schemes. For PDHG-PROXINF and FITRA algorithms, their solutions have most of the entries located near to a ceiling, which results in small dynamic range in terms of PAPR. The solution of ZF scheme shows a large variation with few high peaks. Simulation results verify that PDHG-PROXINF obtains the lowest PAPR of 1.42 dB linked with the first transmit antenna, whereas ZF and the FITRA schemes render higher PAPRs of 9.62 dB and 1.75 dB, respectively.

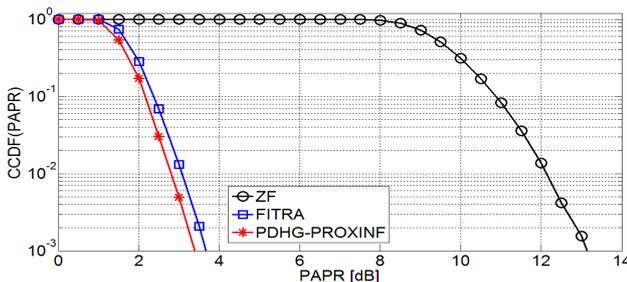
In order to assess the PAPR reduction performance, we compare the CCDF of the PAPR values for respective schemes obtained in 1000 simulation trials in Fig. 3. The PAPR associated with all  $N$  transmit antennas is considered while calculating the empirical CCDF. For system configuration with  $(N, M) = (32, 4)$ , our proposed algorithm reduces the PAPR by 9.7 dB (corresponding to a complementary CDF of  $10^{-3}$ ) compared to ZF scheme and by 0.3 dB compared to the FITRA algorithm with 2000 iterations. Meanwhile, PDHG-PROXINF exhibits a much lower OBR (averaged over 1000 independent runs) than FITRA algorithm. Their OBRs are given by  $-62.49$  dB and  $-56.51$  dB, respectively.

The symbol error rate (SER) performance of each considered scheme is represented in Fig. 4. The average signal-to-noise ratio (SNR) across user terminals is defined as  $\|\mathbf{x}\|_2^2 / N\sigma_n^2$ , where  $N\sigma_n^2$  denotes the noise variance at the receivers. At SER of  $10^{-3}$ , we observe that PDHG-PROXINF suffers a SNR loss of about 1.66 dB and 0.38 dB compared to ZF and FITRA schemes, respectively. This performance loss is mainly caused due to the fact that norm of the obtained solution has more significant increase. It is also worth noting that the ZF precoding attains the least-norm solution.

We now examine the convergence behavior of FITRA and PDHG-PROXINF algorithms for system configuration  $(N, M) = (100, 10)$ . Figures 5(a) and 5(b) show

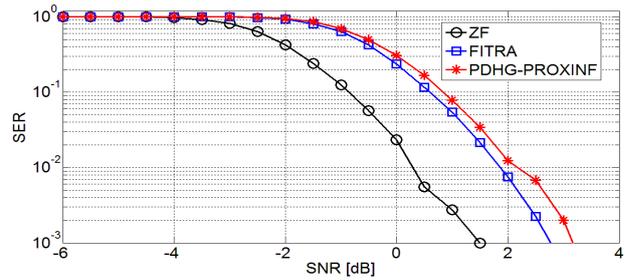


**Fig. 2.** Time representation for different schemes. (a), (b) and (c) are time-domain signals for ZF, FITRA and PDHG-PROXINF respectively (PAPR: ZF = 9.62 dB, FITRA = 1.75 dB and PDHG-PROXINF = 1.42 dB).

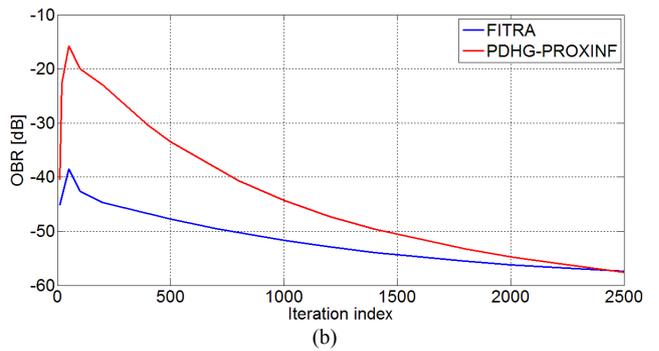
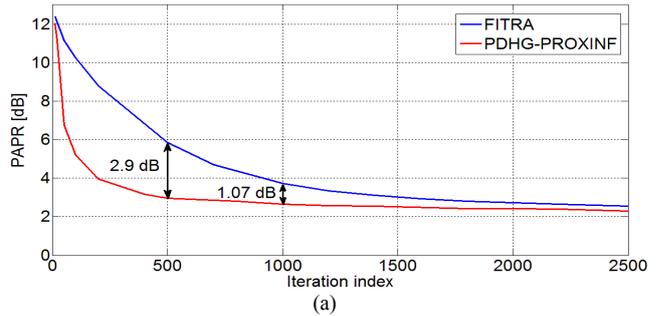


**Fig. 3.** PAPR performance for different schemes for  $(N, M) = (32, 4)$ . (PAPR: ZF = 13.1 dB, FITRA = 3.7 dB and PDHG-PROXINF = 3.4 dB).

the PAPR and OBR performance of these iterative-based algorithms as a function of the number of iterations, respectively. We observe in Fig. 5(a) that PDHG-PROXINF clearly yields much faster initial as well as asymptotic convergence than the FITRA algorithm. PDHG-PROXINF algorithm obtains a PAPR of 2.96 dB within only 500 iterations, while the FITRA algorithm requires about 1550 iterations to achieve the same PAPR reduction performance



**Fig. 4.** Symbol error rate (SER) performance for various schemes.



**Fig. 5.** Comparison of the convergence rates of different metrics for FITRA and PDHG-PROXINF algorithms for  $(N, M) = (100, 10)$ : (a) PAPR, (b) OBR.

Also, our proposed method reduces the PAPR down to 2.65 dB within 1000 iterations, while the FITRA algorithm needs as many as 2200 iterations to obtain a similar result. From Fig. 5(b), we notice that PDHG-PROXINF has OBR of -40.66 dB during the first iteration. The highest average OBR of -15.29 dB is obtained at the 40th iteration which indicates that initially, more amount of power is transmitted outside the active tones  $\phi$ . As the algorithm progresses, the high value of OBR observed in the first few iterations falls rapidly. Although FITRA has fast initial convergence rate, both algorithms obtain almost same OBR at the 2500th iteration. Average OBR of FITRA and PDHG-PROXINF at the 2500th iteration are -57.41 dB and -57.56 dB, respectively.

The impact of the transmit antenna configuration and the channel model to the PAPR reduction is investigated in Fig. 6 for fixed number of user terminals  $M = 10$ , where results are averaged over 1000 independent trials. It is evident that increasing the number of transmit antennas from 20 to 100 and non-zero channel taps from 4 to 8, yields improved PAPR performance for both FITRA and

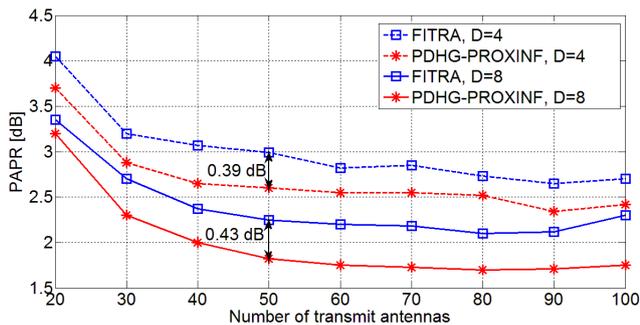


Fig. 6. PAPR performance of FITRA and PDHG- PROXINF algorithms vs. number of transmit antennas.

PDHG-PROXINF algorithms. This result is expected because increasing  $N$  and number of channel taps  $D$  will result in the increase of the degrees-of-freedom (DoF) at the base station. However, the PDHG-PROXINF exploits the inherent DoFs more efficiently as the number of transmit antennas and channel taps increases.

## 5. Conclusion

We have introduced a variant of primal-dual hybrid gradient (PDHG) algorithm for joint consideration of PAPR reduction and MUI cancelation in multiuser MIMO-OFDM systems. The associated proximal map is computed using a linear-time projection algorithm. The proposed method exhibits an optimal rate of convergence for PAPR and OBR metrics in terms of its dependence on the number of iterations. Simulation results show that the proposed algorithm efficiently exploits the large degrees-of-freedom inherent in large-scale antenna systems in order to obtain lower PAPRs than FITRA algorithm [19], meanwhile producing lower out-of-band radiation.

## References

- [1] MARZETTA, T. L. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, 2010, vol. 9, no. 11, p. 3590–3600. DOI: 10.1109/TWC.2010.092810.091092
- [2] BJORNSSON, E., LARSSON, E. G., MARZETTA, T. L. Massive MIMO: Ten myths and one critical question. *IEEE Communications Magazine*, 2016, vol. 54, no.2, p. 114–123. DOI: 10.1109/MCOM.2016.7402270
- [3] SHUMINOSKI, T., JANEVSKI, T. 5G terminals with multi-streaming features for real-time mobile broadband applications. *Radioengineering*, 2017, vol. 26, no. 2, p. 470–478. DOI: 10.13164/re.2017.0470
- [4] RAIDA, Z., KOLKA, Z., MARSALEK, R., PETRZELA, J., et al. Communication subsystems for emerging wireless technologies. *Radioengineering*, 2012, vol. 21, no. 4, p. 1036–1049. ISSN: 1210-2512
- [5] ANGELIS, C. T., CHRONOPOULOS, S. K. System performance of an LTE MIMO downlink in various fading environments. In *International Conference on Ambient Media and Systems (AMBI-SYS)*. Porto (Portugal), 2011, p. 36–43. DOI: 10.1007/978-3-642-23902-1\_5
- [6] WUNDER, G., FISCHER, R.F.H., BOCHE, H., et al. The PAPR problem in OFDM transmission: New directions for a long-lasting problem. *IEEE Signal Processing Magazine*, 2013, vol. 30, no. 6, p. 130–144. DOI: 10.1109/MSP.2012.2218138
- [7] CHRONOPOULOS, S., TATSIS, G., RAPTIS, V., KOSTARAKIS, P. Enhanced PAPR in OFDM without deteriorating BER performance. *International Journal of Communications, Network and System Sciences*, 2011, vol. 4, no. 3, p. 164–169. DOI: 10.4236/ijcns.2011.43020
- [8] CHRONOPOULOS, S., CHRISTOFILAKIS, V., TATSIS, G., KOSTARAKIS, P. Reducing peak-to-average power ratio of a turbo coded OFDM. *Wireless Engineering and Technology*, 2012, vol. 3, no. 4, p. 195–202. DOI: 10.4236/wet.2012.34028
- [9] HAO, M. J., LAI, C. H. Precoding for PAPR reduction of OFDM signals with minimum error probability. *IEEE Transactions on Broadcasting*, 2010, vol. 56, no. 1, p. 120–128. DOI: 10.1109/TBC.2009.2034512
- [10] LI, X., CIMINI, L. J. Effects of clipping and filtering on the performance of OFDM. *IEEE Communication Letters*, 1998, vol. 2, no. 5, p. 131–133. DOI: 10.1109/4234.673657
- [11] ALAVIA, A., TELLAMBURA, C., FAIR, I. PAPR reduction of OFDM signals using partial transmit sequence: An optimal approach using sphere decoding. *IEEE Communication Letters*, 2005, vol. 9, no. 11, p. 982–984. DOI: 10.1109/LCOMM.2005.11014
- [12] CHEN, J.-C., CHIU, M.-H., YANG, Y.-S., et al. A suboptimal tone reservation algorithm based on cross-entropy method for PAPR reduction in OFDM systems. *IEEE Transactions on Broadcasting*, 2011, vol. 57, no. 3, p. 752–756. DOI: 10.1109/TBC.2011.2127590
- [13] BAUML, R. W., FISCHER, R. F. H., HUBER, J. B. Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping. *IEEE Electronics Letters*, 1996, vol. 32, no. 22, p. 2056–2057. DOI: 10.1049/el:19961384
- [14] KRONGOLD, B. S., JONES, D. L. PAR reduction in OFDM via active constellation extension. *IEEE Transactions on Broadcasting*, 2003, vol. 49, no. 3, p. 258–268. DOI: 10.1109/TBC.2003.817088
- [15] ZHU, X. A low-BER clipping scheme for PAPR reduction in STBC MIMO-OFDM system. *Wireless Personal Communications*, 2012, vol. 65, no. 2, p. 335–346. DOI: 10.1007/s11277-011-0259-x
- [16] KU, S.-J. Low-complexity PTS-based schemes for PAPR reduction in SFBC MIMO-OFDM systems. *IEEE Transactions on Broadcasting*, 2014, vol. 60, no. 4, p. 650–658. DOI: 10.1109/TBC.2014.2364966
- [17] MANASSEH, E., OHNO, S., NAKAMOTO, M. Combined channel estimation and PAPR reduction technique for MIMO-OFDM systems with null subcarriers. *EURASIP Journal on Wireless Communications and Networking*, 2012, vol. 2012, no. 1, p. 1–15. DOI: 10.1186/10.1186/1687-1499-2012-201
- [18] JIANG, T., NI, C., GUAN, L. A novel phase offset SLM scheme for PAPR reduction in Alamouti MIMO -OFDM systems without side information. *IEEE Signal Processing Letters*, 2013, vol. 20, no. 4, p. 383–386. DOI: 10.1109/LSP.2013.2245119
- [19] STUDER, C., LARSSON, E. G. PAR-aware large-scale multi-user MIMO-OFDM downlink. *IEEE Journal on Selected Areas in Communications*, 2013, vol. 31, no. 2, p. 303–313. DOI: 10.1109/JSAC.2013.130217
- [20] BAO, H., FANG, J., CHEN, Z., et al. *Perturbation-assisted PAPR Reduction for Large-scale MIMO-OFDM Systems via ADMM*. 2016, 9 p. [Online] Available at: <https://arxiv.org/abs/1607.02681>

- [21] PRABHU, H., EDFORS, O., RODRIGUES, J., et al. A low complex peak-to-average power reduction scheme for OFDM based massive MIMO systems. In *6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*. Athens, Greece, 2014, p. 114–117. DOI: 10.1109/ISCCSP.2014.6877829
- [22] CHEN, J., WANG, C., WONG, K., et al. Low-complexity precoding design for massive multiuser MIMO systems using approximate message passing. *IEEE Transactions on Vehicular Technology*, 2016, vol. 65, no. 7, p. 5707–5714. DOI: 10.1109/TVT.2015.2457450
- [23] BAO, H., FANG, J., LI, H., CHEN, Z., et al. An efficient Bayesian PAPR reduction method for OFDM-based massive MIMO systems. *IEEE Transactions on Wireless Communications*, 2016, vol. 15, no. 6, p. 4183–4195. DOI: 10.1109/TWC.2016.2536662
- [24] ZHU, M., CHAN, T. An efficient primal-dual hybrid gradient algorithm for total variation image restoration. *UCLA CAM Technical Report*, 2008, p. 08–34.
- [25] JOHAM, M., UTSCHICK, W., NOSSEK, J. Linear transmit processing in MIMO communications systems. *IEEE Transactions on Signal Processing*, 2005, vol. 53, no. 8, p. 2700–2712. DOI: 10.1109/TSP.2005.850331
- [26] SEETHALER, D., BOLCSKEI, H. Performance and complexity analysis of infinite-norm sphere coding. *IEEE Transactions on Information Theory*, 2010, vol. 56, no. 3, p. 1085–1105. DOI: 10.1109/TIT.2009.2039034
- [27] STUDER, C., GOLDSTEIN, T., YIN, W., et al. *Democratic Representations*. 2015, 43 p. [Online] Available at: <https://arxiv.org/abs/1401.3420>
- [28] HE, B., YUAN, X. Convergence analysis of primal-dual algorithms for a saddle-point problem: From contraction perspective. *SIAM Journal on Imaging Sciences*, 2012, vol. 5, no. 1, p. 119–149. DOI: 10.1137/100814494
- [29] CONDAT, L. A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. *Journal of Optimization Theory and Applications*, 2013, vol. 158, no. 2, p. 460–479. DOI: 10.1007/s10957-012-0245-9
- [30] GOLDSTEIN, T., ESSER, E., BARANIUK, R. *Adaptive Primal-Dual Hybrid Gradient Methods for Saddle-point Problems*. 2015, 24 p. [Online] Available at: <https://arxiv.org/abs/1305.0546>
- [31] PARIKH, N., BOYD, S. Proximal algorithms. *Foundations and Trends in Optimization*, 2014, vol. 1, no. 3, p. 127–239. DOI: 10.1561/2400000003
- [32] CHAMBOLLE, A., VORE, R. D., LEE, N.-Y., et al. Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Transactions on Image Processing*, 1998, vol. 7, no. 3, p. 319–335. DOI: 10.1109/83.661182
- [33] DUCHI, J., SHALEV-SHWARTZ, S., SINGER, Y., et al. Efficient projections onto the  $\ell_1$ -ball for learning in high dimensions. In *Proceedings of the 25th International Conference on Machine Learning*. Helsinki (Finland), 2008, p. 272–279. DOI: 10.1145/1390156.1390191
- [34] BERG, E. V. D., FRIEDLANDER, M. P. Probing the Pareto frontier for basis pursuit solutions. *SIAM Journal on Scientific Computing*, 2008, vol. 31, no. 2, p. 890–912. DOI: 10.1137/080714488
- [35] BERG, E. V. D., SCHMIDT, M., FRIEDLANDER, M. P., et al. *Group Sparsity via Linear Time Projection*. Dept. of Computer Science, University of British Columbia, 2008, Tech. Rep. TR-2008-09.
- [36] IEEE 802.11 WORKING GROUP. Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications, amendment 5: enhancements for higher throughput. *IEEE 802.11n*, 2009.
- [37] SCHLEGEL, C., PEREZ, L. Trellis representations. In *Trellis and Turbo Coding*. John Wiley, 2003, p. 103–157. ISBN: 0471227552
- [38] PARSONS, J. D. Wideband channel characterisation. In *The Mobile Radio Propagation Channel*. 2<sup>nd</sup> ed., Wiley, 2000, p. 164–189. ISBN: 978-0-471-98857-1.
- [39] OCHIAI, H., IMAI, H. On the distribution of the peak-to-average power ratio in OFDM signals. *IEEE Transactions on Communications*, 2001, vol. 49, no. 2, p. 282–289. DOI: 10.1109/26.905885

## About the Authors ...

**Davinder SINGH** was born in Punjab, India, in 1989. He received his B.Tech. degree in Electronics and Communication Engineering from I.K. Gujral Punjab Technical University, Punjab, India, in 2011, M.E. (ECE) from Panjab University, Chandigarh, India, in 2013, and is currently working toward the Ph.D. degree at Dr. B.R. Ambedkar National Institute of Technology, Jalandhar, India. His research interests include signal processing for wireless communication and MIMO wireless systems.

**R.K. SARIN** was born in New Delhi, India, in 1955. He did his BSc (Engg) ECE, from NIT-Kurukshetra in 1978, ME (ECE) from IIT-Roorkee in 1980 and Ph.D. from A F Ioffe Physico Technical Institute-St Petersburg (Russia) in 1987. His research interests are in semiconductors, optoelectronics, microelectronics/VLSI, microwaves and RF; and he has published papers in these areas.