Alternating Projections Method for Joint Precoding and Peak-to-Average-Power Ratio Reduction

Conference Paper

Author(s):
Taner, Sueda; Studer, Christoph

Alternating Projections Method for Joint Precoding and Peak-to-Average-Power Ratio Reduction

Sueda Taner and Christoph Studer

Department of Information Technology and Electrical Engineering, ETH Zurich, Switzerland
e-mail: taners@iis.ee.ethz.ch and studer@ethz.ch

Abstract—Orthogonal frequency-division multiplexing (OFDM) time-domain signals exhibit high peak-to-average (power) ratio (PAR), which requires linear radio-frequency chains to avoid an increase in error-vector magnitude (EVM) and out-of-band (OOB) emissions. In this paper, we propose a novel joint PAR reduction and precoding algorithm that relaxes these linearity requirements in massive multiuser (MU) multiple-input multiple-output (MIMO) wireless systems. Concretely, we develop a novel alternating projections method, which limits the PAR and transmit power increase while simultaneously suppressing MU interference. We provide a theoretical foundation of our algorithm and provide simulation results for a massive MU-MIMO-OFDM scenario. Our results demonstrate significant PAR reduction while limiting the transmit power, without causing EVM or OOB emissions.

I. INTRODUCTION

Massive multiuser (MU) multiple-input multiple-output (MIMO) is a popular technology to increase spectral efficiency [1]. Orthogonal frequency-division multiplexing (OFDM) [2] is widely used to deal with inter-symbol-interference. While the combination of massive MU-MIMO with OFDM has the potential to achieve high spectral efficiency, OFDM time-domain signals exhibit high dynamic range [3], which requires *linear* radio-frequency (RF) chains to avoid an increase in error-vector magnitude (EVM) and out-of-band (OOB) emissions. Unfortunately, such systems would require a large number of linear RF chains, which quickly results in excessively high RF circuit power consumption and system costs. To address this issue, a plethora of peak-to-average (power) ratio (PAR) reduction methods have been proposed in the OFDM literature; see, e.g., [4] and the references therein.

Among the recent literature, the methods in [5]–[10] modify the transmitted OFDM time-domain signals to reduce the PAR. Alternatively, as shown in [11], the excess degrees-of-freedom of massive MU-MIMO systems enables the design of precoding algorithms at the basestation (BS) that jointly reduce the PAR in the transmitted signals while canceling MU interference (MUI). To reduce the complexity of such emerging joint precoding and PAR reduction (JPP) methods, a number of algorithms have been proposed [12]–[20]. However, such

methods either increase MUI, EVM, or OOB emissions, or aim at minimizing the transmit signal's peaks (e.g., their ℓ^∞-norm), but not the actual PAR. Only recently, a novel $\ell^p - \ell^q$-norm minimization approach was proposed in [21], which is capable of finding minimum-PAR solutions unachievable by ℓ^∞-norm minimization methods without causing EVM, assuming perfect transmit-side channel-state information (CSI), or OOB emissions. Reference [21] also shows the existence of a fundamental trade-off between the PAR and the transmit power of the OFDM time-domain signals, so that reducing the PAR must come at the cost of a power increase (PINC), assuming that the EVM should remain zero. For a given transmit power constraint, this means that one has to back-off in transmit power, which lowers the signal-to-noise ratio (SNR) at the user-equipment (UE) side. Unfortunately, the $\ell^p - \ell^q$-norm minimization approach from [21] provides no control over the PINC of the transmitted time-domain signals.

A. Contributions

We propose a novel JPP method for massive MU-MIMO-OFDM systems based on an alternating projections method (APM) that enables precise control over the PAR and PINC. Our proposed algorithm is not only capable of finding solutions with lower PAR than the widely-used ℓ^∞-norm minimization approaches in [5], [12]–[19], but also enables lower PINC than the recent $\ell^p - \ell^q$-norm minimization approach from [21]. We demonstrate the efficacy of our approach for a massive MU-MIMO-OFDM scenario by showing that our APM achieves low PAR and low PINC, while perfectly removing MUI without causing EVM (assuming perfect CSI at the transmitter) or OOB emissions compared to least-squares precoding.

B. Notation

We represent column vectors and matrices by bold lowercase and uppercase letters, respectively. The kth entry of a vector $\mathbf{a}$ is a_k and the kth column of a matrix $\mathbf{A}$ is $[\mathbf{A}]_k = \mathbf{a}_k$. The matrix transpose and Hermitian are designated by the superscripts $(\cdot)^T$ and $(\cdot)^H$, respectively. We use $\mathbf{0}_{N \times M}$ for the $N \times M$ all-zeros matrix and $\mathbf{F}_N$ for the $N \times N$ unitary discrete Fourier transform (DFT) matrix. We define the ℓ^p-norm as $\|\mathbf{a}\|_p = (\sum_k |a_k|^p)^{1/p}$ and the Frobenius norm as $\|\mathbf{A}\|_F = \sqrt{\sum_k \|\mathbf{a}_k\|_2^2}$. We use calligraphic letters to denote sets and the superscript $(\cdot)^c$ for the set complement. Given a subset of indices $\mathcal{I}$, $\mathbf{a}_\mathcal{I}$ represents the vector whose entries

are given by $\{a_i\}_{i \in \mathcal{I}}$. The orthogonal projection of a vector $\mathbf{a}$ onto a set $\mathcal{A}$ is denoted by $\text{proj}_{\mathcal{A}}(\mathbf{a})$.

II. Prerequisites

The joint precoding and PAR reduction problem in MU-MIMO-OFDM is a special case of a PAR reduction problem with linear constraints. Thus, we first review the PAR minimization problem under a linear constraint $\mathbf{y} = \mathbf{A}\mathbf{x}$ and study the limits of existing algorithms that find solution vectors $\mathbf{x}$ with low (or minimal) PAR in Sections II and III. We then apply the problem setup to the special, but more complicated case of an MU-MIMO-OFDM scenario in Sec. IV.

A. Peak-to-Average (Power) Ratio (PAR)

Given an underdetermined system of linear equations $\mathbf{y} = \mathbf{A}\mathbf{x}$, where $\mathbf{y} \in \mathbb{C}^M$ and $\mathbf{A} \in \mathbb{C}^{M \times N}$ is full-rank with $M < N$, we wish to compute solution vectors $\mathbf{x} \in \mathbb{C}^N$ with low dynamic range. In communication applications, the dynamic range is commonly measured by the peak-to-average-power ratio (PAR)

$$PAR(\mathbf{x}) \triangleq \frac{N\|\mathbf{x}\|_\infty^2}{\|\mathbf{x}\|_2^2}. \tag{1}$$

for non-zero vectors $\mathbf{x} \in \mathbb{C}^N$. The PAR satisfies $1 \le PAR(\mathbf{x}) \le N$, where the lower bound holds with equality if and only if $|x_i| = |x_j|$ for all $i, j \in \{1, \dots, N\}$. We will call such $\mathbf{x}$ a "minimum-PAR (min-PAR) vector" throughout this paper.

B. Computing Low-PAR Solutions via ℓ^∞-Norm Minimization

Since the PAR in (1) is nonconvex and nondifferentiable, directly minimizing $PAR(\mathbf{x})$ subject to the affine constraint $\mathbf{y} = \mathbf{A}\mathbf{x}$ is difficult. Thus, references [11], [14] proposed to minimize the ℓ^∞-norm of solution vectors:

$$(\text{P-}\infty) \quad \hat{\mathbf{x}}^\infty = \arg\min_{\tilde{\mathbf{x}} \in \mathbb{C}^N} \|\tilde{\mathbf{x}}\|_\infty \quad \text{subject to } \mathbf{y} = \mathbf{A}\tilde{\mathbf{x}}.$$

Although ℓ^∞-norm minimization provably reduces the PAR [14], such methods only reduce the signal's peaks and do, in general, not find min-PAR solutions. To compute min-PAR solutions, an $\ell^p - \ell^q$-norm problem formulation was proposed recently in [21]. In Sec. III, we propose a novel method that outperforms such approaches by directly dealing with the following trade-off between PAR and PINC.

C. Fundamental PAR vs. PINC Trade-off

Due to the power constraints of transmitters, we must also consider the power of the solution vectors along with their PAR. The least-squares (LS) vector $\hat{\mathbf{x}}^{\text{LS}} \triangleq \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}\mathbf{y}$, has the minimum power $\|\mathbf{x}\|_2^2$, among all the solutions $\mathbf{x}$ to $\mathbf{y} = \mathbf{A}\mathbf{x}$, by definition. Other solution vectors $\mathbf{x}$, e.g., vectors with low PAR which suit nonlinear RF circuitry better, typically have higher power than $\hat{\mathbf{x}}^{\text{LS}}$. In [21], this observation was made explicit by a fundamental trade-off between the PAR of a solution vector $\mathbf{x}$ satisfying $\mathbf{y} = \mathbf{A}\mathbf{x}$ and its power compared to the LS solution using the following definition [21]:

Definition 1. *Let $\mathbf{x}$ be any solution vector to $\mathbf{y} = \mathbf{A}\mathbf{x}$ and $\hat{\mathbf{x}}^{LS} = \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}\mathbf{y}$. Then, the power increase (PINC) of the vector $\mathbf{x}$ is defined as $PINC(\mathbf{x}) \triangleq \|\mathbf{x}\|_2^2/\|\hat{\mathbf{x}}^{LS}\|_2^2$.*

The following result, taken from [21, Lem. 1], reveals this fundamental trade-off between PAR and PINC.

Lemma 1. *Fix $\mathbf{A}$ and $\mathbf{y}$, and let $\mathbf{x}$ be any nonzero solution vector to $\mathbf{y} = \mathbf{A}\mathbf{x}$. Then, there exists a constant $c \ge 1$ (for fixed $\mathbf{y}$ and $\mathbf{A}$) which satisfies the following inequality:*

$$PAR(\mathbf{x})\,PINC(\mathbf{x}) \ge c. \tag{2}$$

Since the solution vector $\hat{\mathbf{x}}^\infty$ from (P-∞) achieves the lower bound in (2) with equality and defines the constant $c = \|\hat{\mathbf{x}}^\infty\|_\infty^2/\|\hat{\mathbf{x}}^{\text{LS}}\|_2^2$ [21], the vector $\hat{\mathbf{x}}^\infty$ is optimal in the PAR-PINC trade-off, but would not generally have minimal PAR. Furthermore, min-PAR solutions typically have higher PINC than other solutions. Hence, to normalize the transmit power according to a given power constraint, the transmitter must back-off more compared to the power of the LS solution, which will lower the SNR at the UE side.

Although the $\ell^p - \ell^q$-norm minimization approach in [21] was shown to find min-PAR solutions, it provides no control over the PINC of the solutions. We next design a novel method that enables precise control over both the PAR and PINC.

III. PAR and PINC Reduction with an APM

We now introduce our approach to find solution vectors $\mathbf{x}$ with low PAR *and* low PINC for the general case of $\mathbf{y} = \mathbf{A}\mathbf{x}$. A concrete application of our approach to JPP in a massive MU-MIMO-OFDM scenario is shown in Sec. IV.

A. Computing Solutions with Bounded PAR and PINC

Our approach builds upon the following feasibility problem to find solution vectors $\mathbf{x}$ that (i) satisfy $\mathbf{y} = \mathbf{A}\mathbf{x}$, (ii) have bounded PAR, and (iii) have bounded PINC:

$$(\text{P-F}) \quad \begin{cases} \text{find} & \mathbf{x} \in \mathbb{C}^N \\ \text{subject to} & \mathbf{y} = \mathbf{A}\mathbf{x},\ PAR(\mathbf{x}) \le \rho,\ PINC(\mathbf{x}) \le \xi. \end{cases}$$

Here, the parameters ρ and ξ denote the desired PAR and PINC bounds satisfying $1 \le \rho \le N$ and $1 \le \xi$, respectively.

Given the sets $\mathcal{C}, \mathcal{D} \subseteq \mathbb{C}^N$, APMs typically solve constrained feasibility problems of the form [22]

$$\text{find } \mathbf{x} \in \mathbb{C}^N \text{ subject to } \mathbf{x} \in \mathcal{C} \cap \mathcal{D}, \tag{3}$$

by alternating between the projection onto sets $\mathcal{C}$ and $\mathcal{D}$ iteratively for iterations $k = 1, 2, \dots$ until convergence:

$$\mathbf{x}^{(k)} = \text{proj}_{\mathcal{C}}\left(\text{proj}_{\mathcal{D}}\left(\mathbf{x}^{(k-1)}\right)\right). \tag{4}$$

In order to solve (P-F), we define the following sets:

$$\mathcal{C} \triangleq \{\mathbf{x} \in \mathbb{C}^N \,|\, \mathbf{y} = \mathbf{A}\mathbf{x}\}, \tag{5}$$
$$\mathcal{D} \triangleq \{\mathbf{x} \in \mathbb{C}^N \,|\, PAR(\mathbf{x}) \le \rho, PINC(\mathbf{x}) \le \xi\}. \tag{6}$$

We detail the projection operators required in (4) below. To calculate solutions $\mathbf{x}$ with bounded PAR and PINC, we run our APM for a fixed number of iterations $K_{\max}$. We emphasize that $\mathbf{y} = \mathbf{A}\mathbf{x}^{(k)}$ is always satisfied since $\text{proj}_{\mathcal{C}}(\cdot)$ is carried out *after* evaluating $\text{proj}_{\mathcal{D}}(\cdot)$ in each iteration. For the same reason, $\mathbf{x}^{(k)}$ does not necessarily satisfy the PAR-PINC constraints defined by $\mathcal{D}$, and our APM procedure computes only approximate solutions to (P-F) as there are no guarantees on feasibility or convergence due to nonconvexity of the optimization problem.

B. The Orthogonal Projection Operators

1) Projection onto the Set $\mathcal{C}$: The orthogonal projection of a vector $\mathbf{z} \in \mathbb{C}^N$ onto the set $\mathcal{C}$ defined in (5) is given by [23]:

$$\text{proj}_{\mathcal{C}}(\mathbf{z}) = \mathbf{z} - \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}(\mathbf{A}\mathbf{z} - \mathbf{y}). \tag{7}$$

Note that $\text{proj}_{\mathcal{C}}(\mathbf{0}_{N\times 1}) = \hat{\mathbf{x}}^{\text{LS}}$. Hence, we initialize $\mathbf{x}^{(0)} = \mathbf{0}_N$ and, only in the first APM iteration, skip $\text{proj}_{\mathcal{D}}(\mathbf{x}^{(0)})$ to ensure that our method starts from the LS solution $\mathbf{x}^{(1)} = \hat{\mathbf{x}}^{\text{LS}}$.

2) Projection onto the Set $\mathcal{D}$: Using $\hat{\mathbf{x}}^{\text{LS}}$ from the first APM iteration, we compute $\mathbf{x} = \text{proj}_{\mathcal{D}}(\mathbf{z}) \in \mathbb{C}^N$ with the steps listed below. A rigorous derivation of the following projection algorithm can be found in App. A.

1: Set $P = \xi\|\hat{\mathbf{x}}^{\text{LS}}\|_2^2$.
2: If $PAR(\mathbf{z}) \leq \rho$, then set $\mathbf{x}' = \mathbf{z}$ and $P' = \|\mathbf{x}'\|_2^2$, and skip to Step 5. Otherwise, set $\alpha = \rho/N$ and initialize $L = 1$.
3: Let the set $\mathcal{I}$ index the L entries of $\mathbf{z}$ with the largest magnitude. If (i) this set is not uniquely determined, or (ii) $\max_{i\in\mathcal{I}^c} |z_i| \leq \sqrt{\frac{\alpha}{1-\alpha L}}\|\mathbf{z}_{\mathcal{I}^c}\|_2 < \min_{i\in\mathcal{I}} |z_i|$ is not satisfied, then increment L by 1 and repeat this step.
4: If $\mathbf{z}_{\mathcal{I}^c} = \mathbf{0}_{(N-L)\times 1}$, then compute $\mathbf{x}' \in \mathbb{C}^N$ using

$$x_i' = \begin{cases} \sqrt{\frac{(1-\alpha L)P'}{N-L}}, & i \in \mathcal{I}^c \\ \frac{\sqrt{\alpha P'}}{|z_i|}z_i, & i \in \mathcal{I}, \end{cases} \tag{8}$$

where $P' = \alpha\|\mathbf{z}_{\mathcal{I}}\|_1^2$. Otherwise, compute $\mathbf{x}' \in \mathbb{C}^N$ using

$$x_i' = \begin{cases} \frac{\sqrt{(1-\alpha L)P'}}{\|\mathbf{z}_{\mathcal{I}^c}\|_2}z_i, & i \in \mathcal{I}^c \\ \frac{\sqrt{\alpha P'}}{|z_i|}z_i, & i \in \mathcal{I}, \end{cases} \tag{9}$$

where $P' = \left(\sqrt{1-\alpha L}\|\mathbf{z}_{\mathcal{I}^c}\|_2 + \sqrt{\alpha}\|\mathbf{z}_{\mathcal{I}}\|_1\right)^2$.
5: $\mathbf{x} = \min\{1, \sqrt{P/P'}\}\mathbf{x}'$.

Since Step 3 requires sorting of the magnitudes of the entries of $\mathbf{z}$, the complexity of computing $\text{proj}_{\mathcal{D}}(\mathbf{z})$ is $O(N\log(N))$. We conclude by noting that a special case of this procedure was given in [24, Alg. 2], where an equality constraint is imposed on $\|\mathbf{x}\|_2^2$ instead of an upper bound—this reduces the PAR constraint to an ℓ^∞-norm constraint, which is different from $\text{proj}_{\mathcal{D}}(\cdot)$ and easier to solve. However, imposing an upper bound instead of an equality has the advantage of computing solutions with lower power, if they exist, in congruence with our aim of keeping the PINC small.

C. Example of PAR Reduction with our APM

In Fig. 1, we show an example of PAR reduction with our APM. We apply our algorithm to one fixed instance of a circularly-symmetric complex standard normal matrix $\mathbf{A} \in \mathbb{C}^{100\times 200}$ and vector $\mathbf{y} \in \mathbb{C}^{100}$. We consider the cases where the PAR and PINC bounds are: (i) $\rho_{\text{dB}} = 0.4$ and $\xi_{\text{dB}} = 1.6$, and (ii) $\rho_{\text{dB}} = 0.2$ and $\xi_{\text{dB}} = 2$ in decibel. As a baseline, we also show ℓ^∞-norm minimization via CRAMP [14][1] and $\ell^2-\ell^1$-norm minimization from [21]. All algorithms are initialized with $\mathbf{x}^{(1)} = \text{proj}_{\mathcal{C}}(\mathbf{0}_{100\times 1}) = \hat{\mathbf{x}}^{\text{LS}}$ and run until convergence.

---

[1]We select one ℓ^∞-norm minimization algorithm out of the many in the literature to demonstrate the limitation of such methods for reducing the PAR.
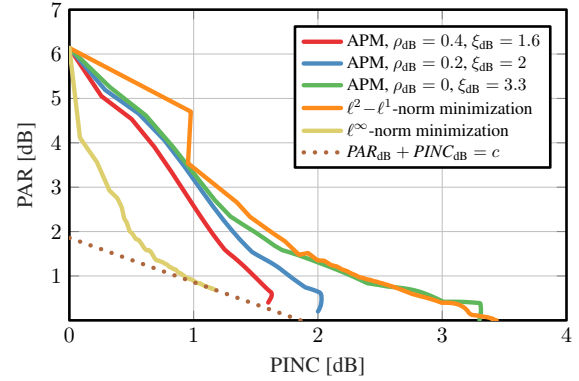

Fig. 1. Trade-off between PAR and PINC for the proposed APM, $\ell^p-\ell^q$-norm minimization [21], and ℓ^∞-norm minimization solved via CRAMP [14]. The APM achieves the same PAR as $\ell^2-\ell^1$-norm minimization at lower PINC.

Fig. 1 shows the PAR vs. PINC trade-off for all iterations together with the lower-bound given by Lem. 1. We observe that the ℓ^∞-norm minimization approach converges to a solution whose PAR is 0.7 dB, and has smaller PINC than the other methods for this PAR value. However, the min-PAR solutions, i.e., solutions whose PAR is 0 dB, are attained by our APM and $\ell^p-\ell^q$-norm minimization—but, as expected, at the cost of higher PINC. This example effectively demonstrates that the APM and $\ell^p-\ell^q$-norm minimization methods are more suitable to *minimize* the PAR than ℓ^∞-norm minimization, which instead yields solutions achieving the optimal PAR-PINC tradeoff. We also observe that the proposed APM outperforms $\ell^p-\ell^q$-norm minimization in the PAR-PINC trade-off: APM with $\rho_{\text{dB}} = 0.4, \xi_{\text{dB}} = 1.6$ and $\rho_{\text{dB}} = 0.2, \xi_{\text{dB}} = 2$ converge to 0.4 dB and 0.2 dB PAR with 1.4 dB and 1.2 dB lower PINC than $\ell^2-\ell^1$-norm minimization, respectively. This example showcases the APM's advantage over $\ell^p-\ell^q$-norm minimization in causing lower PINC while reducing the PAR of solutions.

IV. THE MASSIVE MU-MIMO-OFDM CASE

We now show an application example of the proposed APM for JPP in a massive MU-MIMO-OFDM scenario.

A. System Model

We consider the downlink of a massive MU-MIMO-OFDM system where a B-antenna BS transmits data to U single-antenna UEs with $U < B$. We assume that the total number of OFDM subcarriers is W, and $\mathcal{W}$ and $\mathcal{W}^c$ designate the sets of used and unused OFDM subcarriers, respectively, where $|\mathcal{W}| + |\mathcal{W}^c| = W$. For the used subcarriers, the signal vectors $\mathbf{s}_w \in \mathcal{S}^U$, $w \in \mathcal{W}$, contain the symbols to be transmitted from the constellation $\mathcal{S}$ for each UE. For the unused subcarriers, we set $\mathbf{s}_w = \mathbf{0}_{U\times 1}$, $w \in \mathcal{W}^c$. The signal vectors $\mathbf{s}_w$, $w \in \mathcal{W}$, are precoded into W frequency-domain vectors $\mathbf{x}_w \in \mathbb{C}^B$ with the goal of suppressing MUI. We define the matrix $\mathbf{X} \in \mathbb{C}^{B\times W}$ as $\mathbf{X} \triangleq [\mathbf{x}_1, \ldots, \mathbf{x}_W]$. The frequency-domain outputs of each BS antenna are given by $\mathbf{X}^T$, whose columns correspond to BS antennas and rows to OFDM subcarriers. The total transmit power $\|\mathbf{X}\|_F^2$ depends on the transmit signals $\mathbf{s}_w$, $\forall w$, and the channel state. Hence, we assume that the precoded vectors will be normalized prior to transmission via $\hat{\mathbf{x}}_w = \mathbf{x}_w/\|\mathbf{X}\|_F$ to

ensure unit transmit power. While this normalization step is necessary to satisfy regulatory power constraints, we omit this step while describing our precoder and recall it in Sec. V.

Let $\mathbf{t}_b$ denote the time-domain output of the bth BS antenna and define the matrix $\mathbf{T} \in \mathbb{C}^{W \times B}$ by $\mathbf{T} \triangleq [\mathbf{t}_1, \ldots, \mathbf{t}_B]$. In an OFDM system, the matrix $\mathbf{T}$ is given by the inverse DFT as $\mathbf{T} = \mathbf{F}^H \mathbf{X}^T$. To simplify notation, we define the linear mapping from $\{\mathbf{t}_b\}_{b=1}^{B}$ to $\{\mathbf{x}_w\}_{w=1}^{W}$ as follows:

$$\psi_w(\mathbf{t}_1, \ldots, \mathbf{t}_B) \triangleq [(\mathbf{F}\mathbf{T})^T]_w = \mathbf{x}_w, \ w \in \{1, \ldots, W\}. \quad (10)$$

In order to avoid ISI, a cyclic prefix is prepended to each time-domain vector $\mathbf{t}_b$, $b \in \{1, \ldots, B\}$, prior to transmission.

Let $\mathbf{y}_w \in \mathbb{C}^W$ denote the received vector at OFDM subcarrier w. We model $\mathbf{y}_w$ with the following input-output relation of the wireless channel in the frequency domain:

$$\mathbf{y}_w = \mathbf{H}_w \mathbf{x}_w + \mathbf{n}_w, \ w \in \{1, \ldots, W\}. \quad (11)$$

Here, $\mathbf{H}_w \in \mathbb{C}^{U \times B}$ denotes the MIMO channel matrix associated with the wth subcarrier and $\mathbf{n}_w \in \mathbb{C}^U$ models circularly-symmetric Gaussian noise. Finally, OFDM demodulation is performed by each of the U UEs in order to estimate the transmitted data symbols, i.e., data detection is carried out using $[\mathbf{y}_w]_u$ for UE u at each used subcarrier $w \in \mathcal{W}$.

B. Least-Squares Precoding

The BS must employ precoding in order to suppress MUI. Linear precoding simply computes $\mathbf{x}_w = \mathbf{G}_w \mathbf{s}_w$ on the used subcarriers $w \in \mathcal{W}$ with the precoding matrix $\mathbf{G}_w \in \mathbb{C}^{B \times U}$, and set $\mathbf{x}_w = \mathbf{0}_{B \times 1}$ on the unused subcarriers $w \in \mathcal{W}^c$. LS precoding is widely used and minimizes the transmit power while satisfying the following precoding constraints (under the assumption that the channel matrices $\mathbf{H}_w$, $\forall w$, are known perfectly at the BS-side[2]) (PC1) $\mathbf{s}_w = \mathbf{H}_w \mathbf{x}_w$, $w \in \mathcal{W}$, which ensure zero EVM, and (PC2) $\mathbf{x}_w = \mathbf{0}_{B \times 1}$, $w \in \mathcal{W}^c$, which ensure zero OOB emissions. For the LS precoder, the precoding constraints in (PC1) have the closed-form solution $\mathbf{x}_w^{\text{LS}} = \mathbf{G}_w^{\text{LS}} \mathbf{s}_w$ with $\mathbf{G}_w^{\text{LS}} = \mathbf{H}_w^H (\mathbf{H}_w \mathbf{H}_w^H)^{-1}$, $w \in \mathcal{W}$. Although LS precoding perfectly eliminates MUI and minimizes the transmit power, with $PINC = 1$ by definition, the PAR of the resulting time-domain signals is typically high [5], [11]–[19]. We next utilize the proposed APM to mitigate this drawback.

C. Joint Precoding and PAR Reduction with our APM

Massive MU-MIMO systems have the unique property that the downlink channel matrices have a large nullspace, which can be exploited to simultaneously satisfy the precoding constraints (PC1) and (PC2), while shaping the transmitted time-domain signals to reduce the PAR [11]. Our goal is to find frequency-domain precoding vectors $\mathbf{x}_w$ such that the time-domain signals are bounded in both PAR and PINC, while the

---

[2]We assume perfect channel knowledge for simplicity. However, one can also use an estimate of the channel acquired by exploiting reciprocity. In the latter scenario, note that there would be nonzero EVM due to channel estimation errors, however, the EVM *caused by* precoding would still be zero.

vectors $\mathbf{x}_w$ satisfy (PC1) and (PC2). To achieve all of these goals, we propose the following feasibility problem:

$$(\text{JPP-F}) \begin{cases} \text{find} & \mathbf{t}_1, \ldots, \mathbf{t}_B \in \mathbb{C}^W \\ \text{subject to} & \mathbf{s}_w = \mathbf{H}_w \psi_w(\mathbf{t}_1, \ldots, \mathbf{t}_b), w \in \mathcal{W} \\ & \mathbf{0}_{U \times 1} = \psi_w(\mathbf{t}_1, \ldots, \mathbf{t}_b), \ w \in \mathcal{W}^c \\ & PAR(\mathbf{t}_b) \leq \rho, \ b \in \{1, \ldots, B\} \\ & PINC(\mathbf{T}) \leq \xi. \end{cases}$$

Here, we separately minimize the PAR at each transmit antenna and consider the total power of the transmit signals compared to the LS solution by defining $PINC(\mathbf{T}) \triangleq \|\mathbf{T}\|_F^2 / \|\mathbf{T}^{\text{LS}}\|_F^2$.

The remaining piece of the puzzle is to find a solution to (JPP-F) efficiently (and approximately) via our APM. Our approach is merely a more complicated version of the APM proposed in Sec. III and the details are as follows. Since there exists a one-to-one mapping between time and frequency via (10), we can apply the linear constraints in (JPP-F) separately to the columns of the frequency domain matrix $\mathbf{X}$. Here, we initialize with $\mathbf{x}_w^{(0)} = \mathbf{0}_{B \times 1}, \forall w$, which calculates the LS solution in the first iteration, i.e., $\mathbf{X}^{(1)} = \mathbf{X}^{\text{LS}}$. At every iteration, we use (7) to compute $\mathbf{x}_w^{(k)}$ for $w \in \mathcal{W}$ and set $\mathbf{x}_w^{(k)} = \mathbf{0}_{U \times 1}$ for $w \in \mathcal{W}^c$. As in Sec. III-B, we utilize $\|\mathbf{X}^{\text{LS}}\|_F^2 = \|\mathbf{T}^{\text{LS}}\|_F^2$ to rewrite the PINC constraint of (JPP-F) as a power constraint on $\|\mathbf{T}\|_F^2$. We project the time-domain vectors onto the PAR-bounded set individually for each antenna and then scale $\mathbf{T}^{(k)}$ to satisfy the PINC constraint. Analogous to our algorithm for the general case in Sec. III-A, we take the projection onto the linear (precoding) constraints as the output of each iteration, so that (PC1) and (PC2) are always satisfied. This implies that the APM does not increase the EVM or OOB emissions compared to LS precoding, while decreasing the PAR. Note that our APM calculates *approximate* solutions to (JPP-F) with no guarantees on feasibility or convergence due to its nonconvex nature.

V. SIMULATION RESULTS

We now demonstrate the efficacy of the proposed APM for JPP in a massive MU-MIMO-OFDM system.

A. Simulation Setup

As in [21], we consider a massive MU-MIMO-OFDM system with $B = 128$ BS antennas, $U = 16$ single-antenna UEs, and 16-QAM transmission. We consider $W = 2048$ subcarriers and a bandwidth of $20\,\text{MHz}$; the used and unused subcarriers are defined in [25]. We assume a simple channel model with 4 taps, where the entries of the non-zero time-domain matrices are assumed i.i.d. circularly complex Gaussian with unit variance. The frequency-domain channel matrices are obtained using the Fourier transform [26]. As mentioned in Sec. IV-A, the precoded vectors will be normalized to unit power prior to transmission. This back-off in transmit power is equivalent to an SNR_{dB} loss of exactly $PINC_{\text{dB}}$ compared to the LS precoder.

We solve (JPP-F) using our APM as explained in Sec. IV-C and consider two parameter settings (i) $\rho_{\text{dB}} = 3$, $\xi_{\text{dB}} = 0.3$ and (ii) $\rho_{\text{dB}} = 4$, $\xi_{\text{dB}} = 0.1$. As baselines, we also compare with ℓ^{∞}-norm minimization solved using CRAMP [14] and ℓ^4–ℓ^2-norm

minimization[3] from [21]. All algorithms run for $K_{\max} = 20$ iterations. We reiterate that none of these algorithms increase the EVM or OOB emissions compared to LS precoding; this implies that in the resulting error-rate performance, the SNR loss compared to the LS precoder is determined solely by the PINC. Concretely, these JPP algorithms require exactly $PINC_{dB}$ higher SNR than LS precoder to achieve the same BER.

B. Results and Discussion

Fig. 2 shows the trade-off between PINC and PAR for the JPP methods and the massive MU-MIMO-OFDM scenario described in Sec. V-A. We use Monte-Carlo sampling to compute the complementary cumulative distribution function (CCDF) for the PAR at each antenna's time domain output and the total PINC at each iteration for the different JPP methods. The CCDF of a random variable Z is defined as $\text{CCDF}_Z(z) = \mathbb{P}(Z > z)$. For example, the value z for which $\text{CCDF}_Z(z) = 1\%$ is the 99th percentile of Z. In order to demonstrate the per-iteration *pessimistic* behavior of the JPP algorithms, we pick the 99th percentile as the operating point for both PAR_{dB} and $PINC_{dB}$, and show the PAR-PINC trade-off for each algorithm iteration in Fig. 2, where the iterations are indicated by markers. Here, we no longer have the notion of an optimal trade-off since the vectors whose PAR we aim to reduce are not directly the solutions of one linear system (as done in the example of Sec. III-C). We therefore plot the 99th percentile of PAR and PINC over 100 randomized trials. Analogously to Sec. III-C, we observe that our APM and the $\ell^4 - \ell^2$-norm method are able to compute lower-PAR solutions than ℓ^∞-norm minimization. Since we perform only 20 iterations, we do not expect the APM variants to satisfy the specified PAR and PINC bounds ρ_{dB} and ξ_{dB}, respectively. Nonetheless, the proposed APM has the advantage of achieving 4.7 dB PAR at 1 dB and 1.1 dB lower PINC than ℓ^∞-norm minimization, with $\rho_{dB} = 3, \xi_{dB} = 0.3$ and $\rho = 4, \xi_{dB} = 0.1$, respectively; this demonstrates our APM's capability to find solutions with lower PINC than $\ell^p - \ell^q$-norm minimization at the same PAR.

Fig. 3 shows the CCDF of PAR and PINC for the 5th iteration of the considered JPP algorithms together with the CCDF resulting from LS precoding. In only five iterations, all of the iterative JPP methods decrease the PAR at a CCDF target value of 1% by at least 5 dB compared to LS precoding, while the APM variants result in at least 0.4 dB lower PINC than ℓ^∞-norm and $\ell^4 - \ell^2$-norm minimizing approaches. Thus, our APM is a promising low-complexity PAR reduction method for massive MU-MIMO-OFDM systems with power constraints.

VI. CONCLUSIONS

We have proposed a novel formulation for finding min-PAR solutions to an underdetermined linear system using an alternating projections method (APM). Our method simultaneously bounds the PAR of the solution vectors and their power increase (PINC), which leads to solutions that are, in general, unattainable by ℓ^∞-norm minimization and have lower PINC

---

[3]We use the $\ell^4 - \ell^2$-norm instead of the $\ell^2 - \ell^1$-norm, as it was shown in [21] to outperform the $\ell^2 - \ell^1$-norm-based approach in this scenario.
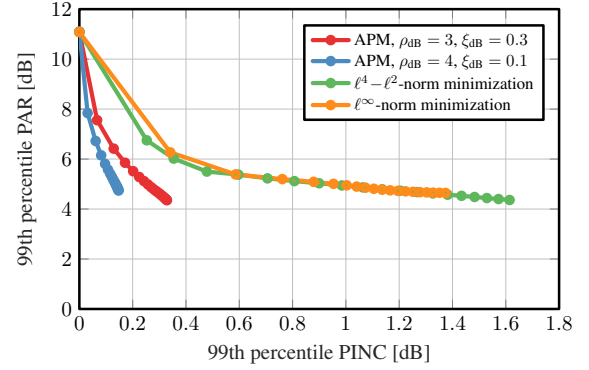

Fig. 2. PAR vs. PINC trade-off for PPB-based JPP methods in a massive MU-MIMO-OFDM system. The markers correspond to iterations and all algorithms start with the LS solution at 11.1 dB PAR and 0 dB PINC.
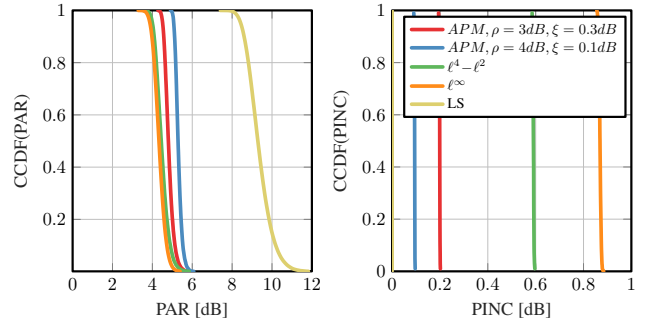

Fig. 3. CCDF of PAR and PINC at the 5th iteration of the considered JPP algorithms and of a baseline LS precoder.

compared to the recent $\ell^p - \ell^q$-norm formulation in [21]. We have applied our approach to a massive MU-MIMO-OFDM scenario, which has shown that the proposed APM is able to decrease the 99th percentile PAR by 5 dB in only 5 algorithm iterations while keeping the PINC below 0.2 dB.

APPENDIX A
DERIVATION OF THE PROJECTION IN SEC. III-B2

As a result of the PAR definition in (1), we can express the PAR constraint of the set $\mathcal{D}$ in (6) as $|x_i|^2 \leq \alpha \|\mathbf{x}\|_2^2, \forall i$, where $\alpha \triangleq \rho/N$ with $1/N \leq \alpha \leq 1$. Furthermore, since $PINC(\mathbf{x}) \triangleq \|\mathbf{x}\|_2^2/\|\mathbf{x}^{LS}\|_2^2$, we can express the PINC constraint of the set $\mathcal{D}$ as $\|\mathbf{x}\|_2^2 \leq P$ with $P \triangleq \xi\|\mathbf{x}^{LS}\|_2^2$. Hence, the optimization problem we wish to solve is as follows:

$$\text{proj}_{\mathcal{D}}(\mathbf{z}) = \begin{cases} \underset{\tilde{\mathbf{x}} \in \mathbb{C}^N}{\arg\min} & \|\mathbf{z} - \tilde{\mathbf{x}}\|_2^2 \\ \text{subject to} & |\tilde{x}_i|^2 \leq \alpha \|\tilde{\mathbf{x}}\|_2^2, \forall i, \|\tilde{\mathbf{x}}\|_2^2 \leq P. \end{cases} \quad (12)$$

Here, we define the sets $\mathcal{D}_1 \triangleq \{\mathbf{x} \in \mathbb{C}^N \mid \|\mathbf{x}\|_2^2 \leq P\}$ and $\mathcal{D}_2 \triangleq \{\mathbf{x} \in \mathbb{C}^N \mid |x_i|^2 \leq \alpha\|\mathbf{x}\|_2^2, \forall i\}$, so that $\mathcal{D} = \mathcal{D}_1 \cap \mathcal{D}_2$. We solve (12) by forming the following Lagrangian function with the dual variables $\mathbf{u} \in \mathbb{R}^N$ and $v \in \mathbb{R}$:

$$\mathfrak{L}(\tilde{\mathbf{x}}, \mathbf{u}, v) = \|\mathbf{z} - \tilde{\mathbf{x}}\|_2^2 + \sum_{i=1}^N u_i(|\tilde{x}_i|^2 - \alpha\|\tilde{\mathbf{x}}\|_2^2) + v(\|\tilde{\mathbf{x}}\|_2^2 - P). \quad (13)$$

Following the definitions of [27] for complex-valued gradients, we arrive at the following partial derivatives:

$$\nabla_{\tilde{x}_i}\mathfrak{L}(\tilde{\mathbf{x}},\mathbf{u},v) = \left(u_i + v + 1 - \alpha\sum_{j=1}^N u_j\right)\tilde{x}_i - z_i. \quad (14)$$

Let $t \triangleq 1 - \alpha\sum_{j=1}^N u_j$ and $\mathbf{x} = \text{proj}_{\mathcal{D}}(\mathbf{z})$ be a minimizer of (12). Then, the vector $\mathbf{x}$ must satisfy the KKT conditions:

(C1) $\nabla_{x_i}\mathfrak{L}(\mathbf{x},\mathbf{u},v) = 0 \implies (u_i + v + t)x_i = z_i, \forall i$.
(C2) $u_i(|x_i|^2 - \alpha\|\mathbf{x}\|_2^2) = 0, \forall i$, and $v(\|\mathbf{x}\|_2^2 - P) = 0$.
(C3) $|x_i|^2 \le \alpha\|\mathbf{x}\|_2^2, \forall i$, and $\|\mathbf{x}\|_2^2 \le P$.
(C4) $u_i \ge 0, \forall i$, and $v \ge 0$.

We note that neither the PAR nor the power constraint depends on the phases of the entries in $\mathbf{x}$. Hence, the entries of a minimizer $\mathbf{x}$ must match the phases of the entries of the input $\mathbf{z}$, which implies that $u_i + v + t \ge 0$ must hold.

Let $P' \triangleq \|\mathbf{x}\|_2^2$. We define the set $\mathcal{I} \triangleq \{i : u_i > 0\}$ with $|\mathcal{I}| = L$ and, thus, by definition, $\mathcal{I}^c \triangleq \{i : u_i = 0\}$ with $|\mathcal{I}^c| = N - L$; we will first derive the solution assuming that $\mathcal{I}$ is known, and then show how to determine $\mathcal{I}$. In order to satisfy (C2), we need that $|x_i|^2 = \alpha P', \forall i \in \mathcal{I}$. Then, $\mathcal{I}$ denotes the indices of the largest L entries of $\mathbf{x}$ (and, hence, of $\mathbf{z}$) in magnitude.

If $\mathcal{I}$ is empty, i.e., $L = 0$, then $u_i = 0, \forall i$ and $t = 1$, hence, we have $vx_i = z_i, \forall i$, by (C1). In order to satisfy (C3), this means that $|z_i|^2 \le \alpha\|\mathbf{z}\|_2^2, \forall i$. In other words, $u_i = 0, \forall i$ if and only if the vector $\mathbf{z}$ already satisfies $PAR(\mathbf{z}) \le \rho$. Then, $\mathbf{x} = \text{proj}_{\mathcal{D}_1}(\mathbf{z}) = \min\left\{1, \sqrt{P}/\|\mathbf{z}\|_2\right\}\mathbf{z}$ as it is well known.

In what follows, we assume $\|\mathbf{z}\|_\infty^2 > \alpha\|\mathbf{z}\|_2^2$ so that $L \ge 1$. We consider the following two cases separately: Case 1, where $\mathbf{z}_{\mathcal{I}^c} = \mathbf{0}_{(N-L)\times 1}$, and Case 2, where $\mathbf{z}_{\mathcal{I}^c} \ne \mathbf{0}_{(N-L)\times 1}$.

Case 1: If $\mathbf{z}_{\mathcal{I}^c} = \mathbf{0}_{N-L}$, then we cannot satisfy the PAR constraint by simply scaling the entries of $\mathbf{z}$, but we have to assign nonzero values to replace the zeros of $\mathbf{z}$. If $z_i = 0$, then any value of x_i could satisfy (C1) as long as $u_i + v + t = 0$. Therefore, if $z_i = 0, \forall i \in \mathcal{I}^c$, then $\{x_i\}_{i\in\mathcal{I}^c}$ is not uniquely determined by the optimization problem. Here, we assume that the minimizer $\mathbf{x}$ will be in the following form:

$$x_i = \begin{cases} \epsilon, & i \in \mathcal{I}^c \\ \frac{z_i}{u_i+v+t}, & i \in \mathcal{I}, \end{cases} \quad (15)$$

where $0 < \epsilon \le \sqrt{\alpha P'} = \frac{|z_i|}{u_i+v+t}, i \in \mathcal{I}$. From (15), we have

$$P' = (N-L)\epsilon^2 + L(\alpha P') \implies \epsilon = \frac{\sqrt{(1-\alpha L)P'}}{\sqrt{N-L}}. \quad (16)$$

From (C1), we have that $(u_i + v + t)\epsilon = 0, i \in \mathcal{I}^c$. As $u_i = 0, i \in \mathcal{I}^c$ by definition, this implies that $v + t = v + 1 - \alpha\sum_{i\in\mathcal{I}}u_i = 0$, hence $\sum_{i\in\mathcal{I}}u_i = (v+1)/\alpha$. Then, since $|x_i| = |z_i|/u_i = |z_j|/u_j, \forall i, j \in \mathcal{I}$, we have that

$$\sum_{i\in\mathcal{I}}u_i = \sum_{i\in\mathcal{I}}\frac{|z_i|}{|z_j|}u_j = \frac{\|\mathbf{z}_{\mathcal{I}}\|_1}{|z_j|}u_j = (v+1)/\alpha \quad (17)$$

$$\implies u_j = \frac{(v+1)|z_j|}{\alpha\|\mathbf{z}_{\mathcal{I}}\|_1}, j \in \mathcal{I}. \quad (18)$$

Therefore, for $j \in \mathcal{I}$, $P' = |z_j|^2/(\alpha u_j^2) = \alpha\|\mathbf{z}_{\mathcal{I}}\|_1^2/(v+1)^2$. Recall that $P' \le P$ from (C3); hence, (C2) implies that if $P' < P$, then $v = 0$. That is, if $\alpha\|\mathbf{z}_{\mathcal{I}}\|_1^2 < P$, then $P' = \alpha\|\mathbf{z}_{\mathcal{I}}\|_1^2$

(and $v = 0$); otherwise, $P' = P$ (and $v = \sqrt{\frac{\alpha}{P}}\|\mathbf{z}_{\mathcal{I}}\|_1 - 1$). Finally, $P' = \min\left\{\alpha\|\mathbf{z}_{\mathcal{I}}\|_1^2, P\right\}$, and the result is given by

$$x_i = \begin{cases} \frac{\sqrt{(1-\alpha L)P'}}{\sqrt{N-L}} & i \in \mathcal{I}^c \\ \frac{\sqrt{\alpha P'}}{|z_i|}z_i, & i \in \mathcal{I}. \end{cases} \quad (19)$$

Case 2: If $\mathbf{z}_{\mathcal{I}^c} \ne \mathbf{0}_{N-L}$, then $u_i + v + t > 0, \forall i$ by (C1), which allows us to write x_i by scaling z_i in the form below:

$$x_i = \begin{cases} \frac{z_i}{v+t}, & i \in \mathcal{I}^c \\ \frac{z_i}{u_i+v+t}, & i \in \mathcal{I}. \end{cases} \quad (20)$$

Intuitively, we divide large entries of $\mathbf{z}$ by a larger constant (compared to small entries) which reduces the dynamic range. From (20), it follows that

$$|x_i| = \sqrt{\alpha P'} = \frac{|z_i|}{u_i+v+t}, i \in \mathcal{I}. \quad (21)$$

From (20) and (21), we have

$$P' = \sum_{i=1}^N |x_i|^2 = \frac{1}{(v+t)^2}\sum_{i\in\mathcal{I}^c}|z_i|^2 + L(\alpha P'). \quad (22)$$

Note that we must have $L < 1/\alpha$, since $P' > 0$ and $\mathbf{z}_{\mathcal{I}^c} \ne \mathbf{0}_{N-L}$; we will revisit this fact later. From (21) and (22), we have the following:

$$P' = \frac{|z_i|^2}{\alpha(u_i+v+t)^2} = \frac{\|\mathbf{z}_{\mathcal{I}^c}\|_2^2}{(1-\alpha L)(v+t)^2}, i \in \mathcal{I} \quad (23)$$

$$\implies u_i = (\beta_i - 1)(v+t), i \in \mathcal{I}, \quad (24)$$

where $\beta_i \triangleq \frac{\sqrt{1-\alpha L}|z_i|}{\sqrt{\alpha}\|\mathbf{z}_{\mathcal{I}^c}\|_2}$ for $i \in \mathcal{I}$. From (23), we obtain

$$v + t = \frac{\|\mathbf{z}_{\mathcal{I}^c}\|_2}{\sqrt{(1-\alpha L)P'}}. \quad (25)$$

Using (21) and (25), we can rewrite (20) as

$$x_i = \begin{cases} \frac{\sqrt{(1-\alpha L)P'}}{\|\mathbf{z}_{\mathcal{I}^c}\|_2}z_i, & i \in \mathcal{I}^c \\ \frac{\sqrt{\alpha P'}}{|z_i|}z_i, & i \in \mathcal{I}. \end{cases} \quad (26)$$

Now, from (21), we express u_i in terms of u_j as follows:

$$\frac{|z_i|}{u_i+v+t} = \frac{|z_j|}{u_j+v+t}, i, j \in \mathcal{I} \quad (27)$$

$$\implies u_j = \frac{|z_j|}{|z_i|}u_i + (v+t)\left(\frac{|z_j|}{|z_i|} - 1\right), i, j \in \mathcal{I}. \quad (28)$$

Using (28), we write t in terms of u_i, $i \in \mathcal{I}$, as follows:

$$t = 1 - \alpha\sum_{j\in\mathcal{I}}\left(\frac{|z_j|}{|z_i|}u_i + (v+t)\left(\frac{|z_j|}{|z_i|} - 1\right)\right) \quad (29)$$

$$= 1 - \alpha\gamma_i u_i - \alpha(\gamma_i - L)(v+t), \quad (30)$$

where we defined $\gamma_i \triangleq \frac{\|\mathbf{z}_{\mathcal{I}}\|_1}{|z_i|}$ for $i \in \mathcal{I}$. Inserting (24) in (30) followed by simplifications, we obtain

$$v + t = \frac{(1+v)\|\mathbf{z}_{\mathcal{I}^c}\|_2}{\sqrt{1-\alpha L}(\sqrt{1-\alpha L}\|\mathbf{z}_{\mathcal{I}^c}\|_2 + \sqrt{\alpha}\|\mathbf{z}_{\mathcal{I}}\|_1)}. \quad (31)$$

Inserting (31) in (23), we have

$$P' = \left(\sqrt{1-\alpha L}\|\mathbf{z}_{\mathcal{I}^c}\|_2 + \sqrt{\alpha}\|\mathbf{z}_{\mathcal{I}}\|_1\right)^2/(1+v)^2. \quad (32)$$

Recall that $P' \leq P$ from (C3); hence, (C2) implies that if $P' < P$, then $v = 0$. That is, if $\left(\sqrt{1-\alpha L}\|\mathbf{z}_{\mathcal{I}^c}\|_2 + \sqrt{\alpha}\|\mathbf{z}_{\mathcal{I}}\|_1\right)^2 < P$, then $P' = \left(\sqrt{1-\alpha L}\|\mathbf{z}_{\mathcal{I}^c}\|_2 + \sqrt{\alpha}\|\mathbf{z}_{\mathcal{I}}\|_1\right)^2$ (and $v = 0$); otherwise, $P' = P$ (and $v = \left(\sqrt{1-\alpha L}\|\mathbf{z}_{\mathcal{I}^c}\|_2 + \sqrt{\alpha}\|\mathbf{z}_{\mathcal{I}}\|_1\right)^2/P - 1$). Therefore, $P' = \min\left\{\left(\sqrt{1-\alpha L}\|\mathbf{z}_{\mathcal{I}^c}\|_2 + \sqrt{\alpha}\|\mathbf{z}_{\mathcal{I}}\|_1\right)^2, P\right\}$. Inserting (32) in (26) yields the desired final result.

Determining the Index Set $\mathcal{I}$: So far, we have assumed that we knew $\mathcal{I}$ (and thus L); we will now explain how to determine $\mathcal{I}$. We will first consider the constraints that the set $\mathcal{I}$ should satisfy in Case 2. (i) For the condition (C3) to hold, the following must be satisfied:

$$\max_{i\in\mathcal{I}^c}|x_i| = \max_{i\in\mathcal{I}^c}|z_i|/(v+t) \leq \sqrt{\alpha P'}. \qquad (33)$$

Inserting (31) and (32) in (33) yields

$$\max_{i\in\mathcal{I}^c}|z_i| \leq \sqrt{\tfrac{\alpha}{1-\alpha L}}\|\mathbf{z}_{\mathcal{I}^c}\|_2. \qquad (34)$$

(ii) From (21), we have that $u_i = \frac{|z_i|}{\sqrt{\alpha P'}} - (v+t)$, $i \in \mathcal{I}$. Then, for $u_i > 0$, $\forall i \in \mathcal{I}$ to hold, the following must be satisfied:

$$\min_{i\in\mathcal{I}}|z_i| > (v+t)\sqrt{\alpha P'} = \sqrt{\tfrac{\alpha}{1-\alpha L}}\|\mathbf{z}_{\mathcal{I}^c}\|_2. \qquad (35)$$

Combining (34) and (35) gives us the following:

$$\max_{i\in\mathcal{I}^c}|z_i| \leq \sqrt{\tfrac{\alpha}{1-\alpha L}}\|\mathbf{z}_{\mathcal{I}^c}\|_2 < \min_{i\in\mathcal{I}}|z_i|. \qquad (36)$$

Note that (36) is also trivially satisfied if $\mathbf{z}_{\mathcal{I}_c} = \mathbf{0}_{N-L}$, which corresponds to Case 1.

Now, recall that $\mathcal{I}$ is empty when $\mathbf{z}$ already satisfies the PAR constraint; and note that $\mathcal{I}$ indexes the entries of $\mathbf{z}$ whose magnitude is too large to satisfy the PAR constraint and will be clipped to a fixed upper bound value in the corresponding entries of $\mathbf{x}$. Since $\mathbf{x}$ has the minimum distance to $\mathbf{z}$ while satisfying the PAR constraint, the set $\mathcal{I}$ (thus L) should be as small as possible while the KKT conditions are satisfied, and also defined uniquely in order to avoid an ambiguity about which entries to clip. Therefore, we begin our search for L by initializing $|\mathcal{I}| = L = 1$ in our solution procedure and increment L by 1 until $\mathcal{I}$ is determined uniquely and (36) is satisfied. Recall that $L < 1/\alpha$ from (22), which implies that we will test (36) for at most $\lfloor 1/\alpha \rfloor$ values of L before finding the correct choice, where $\lfloor \cdot \rfloor$ denotes rounding towards $-\infty$. Testing for the condition in (36) requires calculating the magnitude of the entries of $\mathbf{z}$ and sorting them once regardless of how many L's we try; hence, the complexity remains $O(N\log N)$.

Finally, we make an important observation: Computing $\mathrm{proj}_{\mathcal{D}}(\mathbf{z})$ is equivalent to performing $\mathrm{proj}_{\mathcal{D}_1}\left(\mathrm{proj}_{\mathcal{D}_2}(\mathbf{z})\right)$. As a proof sketch, we note that (i) $\mathrm{proj}_{\mathcal{D}_2}(\mathbf{z})$ is a special case of our derivation since $\mathrm{proj}_{\mathcal{D}}(\mathbf{z}) = \mathrm{proj}_{\mathcal{D}_2}(\mathbf{z})$ for $P = \infty$, and (ii) it is well-known that $\mathrm{proj}_{\mathcal{D}_1}(\mathbf{x}') = \min\left\{1, \sqrt{P}/\|\mathbf{x}'\|_2\right\}\mathbf{x}'$. In Sec. III-B2, we computed $\mathbf{x}' = \mathrm{proj}_{\mathcal{D}_2}(\mathbf{z})$ in Steps 2-to-4 and $\mathrm{proj}_{\mathcal{D}_1}(\mathbf{x}')$ in Step 5 to emphasize the separability of the PAR and PINC constraints.

REFERENCES

[1] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[2] R. van Nee and R. Prasad, *OFDM for wireless multimedia communications.* Artech House Publ., 2000.

[3] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Comm.*, vol. 12, no. 2, pp. 1536–1284, Apr. 2005.

[4] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1567–1592, Mar. 2013.

[5] H. Bao, J. Fang, Q. Wan, Z. Chen, and T. Jiang, "An ADMM approach for PAPR reduction for large-scale MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7407–7418, Aug. 2018.

[6] S. Kant, M. Bengtsson, B. Göransson, G. Fodor, and C. Fischione, "Robust PAPR reduction in large-scale MIMO-OFDM using three-operator ADMM-type techniques," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Oct. 2021, pp. 616–622.

[7] S. Kant, M. Bengtsson, G. Fodor, B. Göransson, and C. Fischione, "EVM mitigation with PAPR and ACLR constraints in large-scale MIMO-OFDM using TOP-ADMM," *IEEE Trans. Wireless Commun.*, vol. 21, no. 11, pp. 9460–9481, May 2022.

[8] S. Gökceli, T. Riihonen, T. Levanen, and M. Valkama, "Machine learning based tuner for frequency-selective PAPR reduction," *IEEE Trans. Veh. Technol.*, pp. 1–6, 2022, early access.

[9] A. Kalinov, R. Bychkov, A. Ivanov, A. Osinsky, and D. Yarotsky, "Machine learning-assisted PAPR reduction in massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 537–541, Mar. 2021.

[10] H. Jang, S. Jang, Y. Park, J. Jung, J. Lee, and S. Choi, "SeqNet: Data-driven PAPR reduction via sequence classification," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM) Workshop*, Dec. 2021, pp. 1–6.

[11] C. Studer and E. G. Larsson, "PAR-aware large-scale multi-user MIMO-OFDM downlink," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 303–313, Feb. 2013.

[12] H.-S. Cha, H. Chae, K. Kim, J. Jang, J. Yang, and D. K. Kim, "Generalized inverse aided PAPR-aware linear precoder design for MIMO-OFDM system," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1363–1366, Aug. 2014.

[13] S. Wang, Y. Li, and J. Wang, "Convex optimization based downlink precoding for large-scale MIMO," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2014, pp. 218–223.

[14] C. Studer, T. Goldstein, W. Yin, and R. G. Baraniuk, "Democratic representations," Apr. 2015. [Online]. Available: http://arxiv.org/abs/1401.3420

[15] Z. Guo, Y. Yılmaz, and X. Wang, "Transmitter-centric channel estimation and low-PAPR precoding for millimeter-wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 64, no. 7, pp. 2925–2938, Jul. 2016.

[16] R. Zayani, H. Shaiek, and D. Roviras, "PAPR-aware massive MIMO-OFDM downlink," *IEEE Access*, vol. 7, pp. 25 474–25 484, Feb. 2019.

[17] H. Bao, J. Fang, Z. Chen, H. Li, and S. Li, "An efficient Bayesian PAPR reduction method for OFDM-based massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4183–4195, Jun. 2016.

[18] M. Yao, M. Carrick, M. M. Sohul, V. Marojevic, C. D. Patterson, and J. H. Reed, "Semidefinite relaxation-based PAPR-aware precoding for massive MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2229–2243, Mar. 2019.

[19] T. Liu, M. T. Hoang, Y. Yang, and M. Pesavento, "A parallel optimization approach on the infinity norm minimization problem," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2019, pp. 1–5.

[20] L. Hua, Y. Wang, Z. Lian, Y. Su, and Z. Xie, "LADMM-based PAPR-aware precoding for massive MIMO-OFDM downlink systems," *IEEE Trans. Veh. Technol.*, pp. 1–11, Sep. 2022.

[21] S. Taner and C. Studer, "$\ell^p - \ell^q$-norm minimization for joint precoding and peak-to-average-power ratio reduction," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Oct. 2021, pp. 437–442.

[22] O. Ginat, "The method of alternating projections," Sep. 2018. [Online]. Available: http://arxiv.org/abs/1809.05858

[23] T. Goldstein, C. Studer, and R. G. Baraniuk, "A field guide to forward-backward splitting with a FASTA implementation," Nov. 2014. [Online]. Available: http://arxiv.org/abs/1411.3406

[24] J. A. Tropp, I. S. Dhillon, R. W. Heath Jr., and T. Strohmer, "Designing structured tight frames via an alternating projection method," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 188–209, Jan. 2005.

[25] 3GPP, "5G; NR; base station (BS) radio transmission and reception," May 2019, TS 38.104 version 15.5.0 Rel. 15.

[26] C. Studer and G. Durisi, "Quantized massive MU-MIMO-OFDM uplink," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2387–2399, Jun. 2016.

[27] K. Kreutz-Delgado, "The complex gradient operator and the CR-calculus," Jun. 2009. [Online]. Available: https://arxiv.org/abs/0906.4835v1