PAPR Reduction in MIMO-OFDM via Power Efficient Transmit Waveform Shaping

ANTTI ARVOLA[1], **(Student Member, IEEE), SATYA KRISHNA JOSHI**[1], **(Member, IEEE),**
ANTTI TÖLLI[1], **(Senior Member, IEEE), AND DAVID GESBERT**[2], **(Fellow, IEEE)**
[1]Centre for Wireless Communications, University of Oulu, 90014 Oulu, Finland
[2]EURECOM, Campus SophiaTech, 06410 Biot, France

Corresponding author: Antti Arvola (antti.arvola@oulu.fi)

ABSTRACT In this paper we revisit the long-standing problem of peak-to-average power ratio minimization in MIMO-OFDM systems, with a new angle of approach on a well-known scheme. Utilizing the principles of tone reservation, we place dummy symbols, i.e., complex coefficients, on unused space-frequency resources with the aim of jointly minimizing the transmit signal PAPR and the self-power consumption of the dummy symbols. To solve this joint minimization, we propose three different algorithms exhibiting varying degrees of computational complexity and PAPR reduction performance. Our proposed framework utilizes the strict PAPR expression, i.e., we take into account the average transmit power of the antenna, to simultaneously reduce the PAPR on all antennas while keeping the self-power consumption of the scheme minimal. Our simulation results show that this optimization objective provides better worst-case PAPR reduction and dummy symbol power consumption performance compared to the peak power minimization objective widely utilized in the tone reservation literature. Finally, we propose a novel take on a well-known block-diagonalization algorithm by exploiting knowledge on dummy symbol allocations, resulting in high-gain data streams in downlink transmission.

INDEX TERMS ADMM, alternating directions method of multipliers, convex optimization, low complexity, MIMO-OFDM, PAPR, peak-to-average power ratio, tone reservation.

I. INTRODUCTION

Technologies leveraging multiple transmit and receive antennas (MIMO, multiple-input multiple-output) and orthogonal frequency division multiplexing (OFDM) have become a staple in the state-of-the-art cellular technologies. MIMO offers high data rates through beamforming and spatial multiplexing, along with increased diversity by better exploiting multipath propagation. Similarly, the frequency diversity of OFDM provides robustness against frequency selectivity, as it leverages high bandwidths that are divided into independent, narrow-band subcarriers [1]. This multiplexing of frequency resources also makes OFDM a very efficient multiple access scheme, which is one reason it has been adopted in the 3rd Generation Partnership Project 5G New Radio (NR) standard as a basis for both downlink (DL) and uplink (UL) communications [2].

The associate editor coordinating the review of this manuscript and approving it for publication was Jiankang Zhang.

OFDM, however, has a major drawback. Due to the constructive and destructive combining of complex-valued transmit data and beamformers with the inverse Fourier transform, the time-domain transmit waveform is known to exhibit a high dynamic range in signal power. Thus, the transmit signal suffers from high peak-to-average power ratio (PAPR), which causes problems especially in the design of radio frequency (RF) components [1].

High PAPR is detrimental to power amplifiers (PA) through high power consumption and low drain efficiency (ratio of RF output power to direct current input power [3]). To avoid distortions in the transmit signal, such as spectral broadening, the PA must have highly linear (i.e., low efficiency) amplification characteristics, and a back-off on the operating point is required to prevent the high signal peaks from being distorted by the PA saturation. Furthermore, high PAPR increases the power consumption of digital-to-analog converters [4]. Due to the adoption of OFDM as the basis for communications in the 5G NR standard [2], PAPR reduction is still considered a highly relevant problem, especially

in scenarios where power efficiency and consumption are important factors.

To combat the effects of PAPR, a rich body of research has accumulated over the years, and it is still an active research topic [1]. One of the best-known methods is amplitude clipping [5]–[7], a simple method which unfortunately causes clipping noise and signal distortions. The schemes for selected mapping [8], [9] and partial transmit sequences [10], [11] create low-PAPR permutations of the transmit signal but require transmission of side information or changes in the receiver structure. Active constellation extension (ACE), see for example [12], considers a distorted constellation for data symbols to obtain a low-PAPR signal.

Tone reservation (TR) [13]–[19] is a PAPR reduction scheme where subcarriers (tones) are reserved to transmit a peak-reducing signal along the data signal. The scheme does not induce distortions in the data signal but reduces spectral efficiency due to the reserved subcarriers. Many of the referenced TR techniques are single-antenna peak power minimization schemes, and do not translate well to MIMO systems as they ignore the antenna-specific average transmit power. This causes some antennas with lower average transmit power to have higher signal PAPR, as the peak reduction signal only targets the highest absolute peaks.

Methods similar to tone reservation but which optimize the whole transmit signal are proposed in [20], [21]. In [20], the average power is also considered (i.e., the strict PAPR, not just the peak power) through an error vector magnitude (EVM) based linear lower bound. The authors also propose an interior-point method -based algorithm, which requires computationally intensive Newton steps to be solved. The latter reference, [21], exploits the degrees of freedom of a massive MIMO system to jointly optimize precoding, modulation and peak reduction. Under a general MIMO setting, however, the scheme neglects antennas with lower average transmit power, which can still exhibit high PAPR. Furthermore, in [21], multiple system parameters (modulation, beamforming, PAPR) are optimized at once. For this reason, application of the scheme might require a significant overhaul of already existing systems.

For other PAPR reduction methods leveraging the spatial domain for signal shaping, see for example [22], [23], where the former considers a PAPR constrained power allocation strategy in single-carrier frequency division multiple access systems. The method provides good PAPR performance at high signal-to-noise ratio (SNR) but has prohibitively high computational complexity. The latter reference exploits leftover spatial modes from waterfilling for peak reduction, similar to our proposed case. However, their approach follows the peak power minimization principle, solved with both highly complex interior point methods and slow to converge descent methods.

In this article, we consider a PAPR reduction scheme that operates on the tone reservation principles. However, we utilize the term 'dummy symbols' instead of reserved tones, for two reasons. First, a reserved tone can be misunderstood to

mean that a whole subcarrier is reserved, which is misleading, as some spatial streams on the subcarrier could still carry data. Second, we argue the term 'dummy symbol' better describes the scheme's operating principle: use of optimized complex coefficients (symbols) that carry no information (dummy), applied on unused resources to affect the time-domain waveform.

Contributions: We propose a PAPR reduction framework that uses the principles of tone reservation to minimize the strict PAPR of all transmit antennas jointly with the self-power consumption of the scheme, which we call 'dummy symbol power reserve' (or 'power reserved for the dummy symbols'). Our approach has significant differences to the general peak power minimization used in the TR literature, as we also optimize the antenna-wise average transmit power. Thus, instead of using the l_∞-norm as the objective, we minimize the ratio between l_∞ and l_2 norms for all antennas, which improves the overall PAPR performance of a general MIMO-OFDM system by not only focusing on the absolute highest power peaks. This approach also makes the problem more difficult to solve due to its non-convexity. The prior art also does not consider the self-power consumption of the schemes explicitly, with the exception of [20], where it is proposed as an additional constraint. Implicitly, however, minimizing power peaks can be done by lowering the average transmit power, which has an effect on the self-power consumption. However, from the results it can be seen that considering the self-power consumption jointly with the PAPR is very beneficial in our approach.

Furthermore, unlike the prior art (see for example [21]), our approach is modular, i.e., it can operate on any given beamformers and frequency-domain symbol allocation, as long as there are some unused space-frequency resources (empty subcarriers on some streams) to exploit. The modularity of our scheme is beneficial in terms of implementation, as it can easily be applied over existing system configurations.

We derive three different algorithms with varying levels of PAPR reduction capability and computational complexity to minimize the proposed joint objective of strict PAPR and dummy symbol power reserve. We derive one baseline algorithm using successive convex approximation (SCA), solved via interior point methods, and two more computationally tractable low-complexity iterative schemes that are solved via the alternating directions method of multipliers (ADMM).

Our proposed PAPR reduction framework generalizes our previous work, see [24], [25]. As we previously applied the framework in uplink systems, in this article we consider a downlink scheme instead. However, it should be emphasized that the proposed PAPR reduction framework is specific to neither UL nor DL. We propose a novel approach on the well-known iterative block-diagonalization scheme [26], [27], that leverages knowledge about the dummy symbol placements to jointly optimize the data beamformers within a larger interference-free space. This leaves low-gain spatial modes for dummy symbol allocations, but as the dummy

symbols carry no information, this does not affect the rate performance.

Finally, we propose a simple rate maximization search algorithm to better observe the additional benefits obtained by our PAPR reduction scheme. The search algorithm finds, utilizing a MIMO-modification for the well-known Hughes-Hartogs bit and power loading algorithm (HH) [28], [29], the highest possible rate that can be achieved under a fixed peak power constraint over the antennas, symbol realization, and given system parameters. This search algorithm is novel but very straightforward, and we consider the main contribution of this paper to be in the low-complexity joint PAPR and dummy symbol power reserve minimization algorithm.

In summary, our contributions are as follows:

- We propose a novel PAPR reduction framework based on tone reservation principles, which minimizes the strict PAPR of the transmit signal instead of only the highest peak power. The proposed scheme is formulated as an optimization problem to find a PAPR reducing dummy symbol allocation. Our proposed scheme also jointly minimizes the dummy symbol power reserve, i.e., the self-power consumption of the scheme. Through simulations, our proposed scheme is shown to outperform the classical approach of peak power minimization in terms of worst-case PAPR across all antennas.
- We derive three algorithms to solve the optimization problem, with varying degrees of PAPR reduction capability and computational complexity.
- We propose a novel take on the well-known iterative block-diagonalization algorithm to provide high-gain streams for data, and low-gain streams for the dummy symbol allocation.

Organization: This article can be considered in two parts. First, in Sections II-IV we derive our proposed PAPR reduction framework in a general form. We begin with the time domain transmit signal for a generic MIMO-OFDM transmitter in Section II, including the dummy symbols, after which we define the joint PAPR and power reserve minimization problem in Section III. The algorithms to solve this problem are derived in Section IV. In the second part (Sections V-VII), we first propose a beamformer design in Section V that leverages a pre-determined dummy symbol resource allocation mask to jointly optimize the data beamformers and find the dummy symbol beamformers. Then, in Section VI we derive a simple rate maximization search algorithm that finds the highest transmit power possible (benefiting from reduced PAPR), subject to a given peak power constraint. Finally, the performance of the PAPR reduction framework is investigated through Monte Carlo simulations in Section VII.

Notation: We use general vector notation: a, $\mathbf{a}$ and $\mathbf{A}$ denote a scalar, a vector and a matrix, respectively. We use $\mathbb{R}$ and $\mathbb{C}$ to denote the real and complex sets. $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$ and $(\cdot)^*$ denote the transpose, Hermitian transpose and complex conjugate, respectively, while $\|\cdot\|_p$ denotes the l_p-norm and diag($\mathbf{a}$) denotes a diagonal matrix with the elements of $\mathbf{a}$ on the diagonal. Re[$\cdot$] denotes the real part, and $(\cdot)_+ = \max(\cdot, 0)$.

Finally, we denote iteration indices with bracketed superscripts.

II. TRANSMIT SIGNAL MODEL

We consider a generic MIMO-OFDM system where the transmitter is equipped with N_{T} antennas and transmits data (possibly to multiple users) on L parallel spatial streams. We assume that a total of N_{C} subcarriers are available. Let $p_{c,l} \geq 0$ and $\mathbf{V}_{c,l} \in \mathbb{C}^{N_{\mathrm{T}} \times 1}$ denote the power and direction of transmit precoder associated with l-th data stream on subcarrier c. Then, the transmitted frequency-domain signal vector in c-th subcarrier can be expressed as

$$x_c = \mathbf{V}_c P_c^{\frac{1}{2}} d_c, \tag{1}$$

where matrix $\mathbf{V}_c = [\mathbf{V}_{c,1}, \ldots, \mathbf{V}_{c,L}] \in \mathbb{C}^{N_{\mathrm{T}} \times L}$ is the transmit precoding matrix associated with c-th subcarrier, $P_c = \mathrm{diag}(p_{c,1}, \ldots, p_{c,L})$ is the power allocation matrix of L data streams on subcarrier c, and vector $d_c = [d_{c,1}, \ldots, d_{c,L}]^{\mathrm{T}}$ stacks the stream-wise data symbols transmitted on subcarrier c, normalized such that $\mathbb{E}[|d_{c,l}|^2] = 1, \forall c, l$.

To define the transmit waveform PAPR of the OFDM system, we need to find the time domain representation of transmitted signal (1). To do this, first let us compactly denote frequency-domain signal vectors $\{x_c\}_{c=1,\ldots,N_{\mathrm{C}}}$ as $X = [\mathbf{x}_1, \ldots, \mathbf{x}_{N_{\mathrm{C}}}]^{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{C}} \times N_{\mathrm{T}}}$, and equivalently express (1) as

$$X = \sum_{l=1}^{L} \tilde{P}_l^{\frac{1}{2}} \tilde{D}_l \tilde{\mathbf{V}}_l, \tag{2}$$

where we use the notation $\tilde{\mathbf{V}}_l \in \mathbb{C}^{N_{\mathrm{C}} \times N_{\mathrm{T}}}$ to denote a matrix obtained by stacking the stream-wise transmit precoders associated with l-th data stream over N_{C} subcarriers, i.e., $\tilde{\mathbf{V}}_l = [\mathbf{V}_{1,l}, \mathbf{V}_{2,l}, \ldots, \mathbf{V}_{N_{\mathrm{C}},l}]^{\mathrm{T}}$. Similarly,

$$\tilde{D}_l = \mathrm{diag}(d_{1,l}, d_{2,l}, \ldots, d_{N_{\mathrm{C}},l}),$$
$$\tilde{P}_l = \mathrm{diag}(p_{1,l}, p_{2,l}, \ldots, p_{N_{\mathrm{C}},l}).$$

Then, by defining the discrete Fourier transform (DFT) matrix, with α-times oversampling,[1] as $F \in \mathbb{C}^{N_{\mathrm{C}} \times \alpha N_{\mathrm{C}}}$ [5], the time domain representation of the considered MIMO-OFDM system can be expressed as

$$S^{\mathrm{data}} = F^{\mathrm{H}} X = F^{\mathrm{H}} \sum_{l=1}^{L} \tilde{P}_l^{\frac{1}{2}} \tilde{D}_l \tilde{\mathbf{V}}_l. \tag{3}$$

Note that the size of matrix S^{data} is $\alpha N_{\mathrm{C}} \times N_{\mathrm{T}}$, and hence each column of $\mathbf{S}^{\mathrm{data}}$ represents the samples of the time-domain transmit waveform from a single antenna.

To minimize the PAPR of the transmitted waveform, we shape the transmit waveform (3) by exploiting the unused space-frequency resources (i.e, any stream l on subcarrier c with no scheduled data). That is, dummy symbols (complex coefficients) are placed on empty subcarriers or

---

[1] An oversampling factor of $\alpha \geq 4$ should be used to provide sufficiently accurate discrete approximation of the continuous transmit signal [30].

non-interfering spatial modes.[2] This transmit waveform shaping mechanism can be achieved by updating expression (3) as

$$S = S^{\text{data}} + F^{\text{H}} \sum_{l=1}^{L} \text{diag}(m_{1,l}, \ldots, m_{N_C,l}) \tilde{\mathbf{V}}_l, \quad (4)$$

where $m_{c,l} \in \mathbb{C}$ is a dummy symbol that is placed on c-th subcarrier of data stream l. Note that dummy symbols $\{m_{c,l}\}$ are applied in the frequency domain, which are precoded with matrix $\{\tilde{\mathbf{V}}_l\}$ and then transformed to time domain representing using the DFT matrix F in expression (4) (i.e., by following the same steps as in the case of data symbols $\{\tilde{D}_l\}$).

For a compact representation of expression (4), let $\mathbf{m}_l = [m_{1,l}, \ldots, m_{N_C,l}]^{\text{T}}$ for all $l = 1, \ldots, L$, and also denote the n-th column of $\tilde{\mathbf{V}}_l$ by $\tilde{\mathbf{V}}_{l,n}$, then the n-th column (i.e., antenna n waveform) of S can be equivalently expressed as

$$\mathbf{s}_n = \mathbf{s}_n^{\text{data}} + F^{\text{H}} \sum_{l=1}^{L} \text{diag}(\tilde{\mathbf{V}}_{l,n}) \mathbf{m}_l, \quad (5)$$

where $\mathbf{s}_n^{\text{data}}$ is the n-th column of S^{data}. We assume that the data-carrying resource indices are known, and by corollary the indices (c, l) for which $p_{c,l} = 0$ are also known. This lets us obtain a compact reformulation of (5) where we consider only the indices of the resources for which it is possible to place non-zero dummy symbols. To do this, let us use $\tilde{\mathbf{m}} \in \mathbb{C}^{N_D \times 1}$ to denote a vector obtained by stacking $\{m_{c,l}\}$, among all $c = 1, \ldots, N_C$ and $l = 1, \ldots, L$, for which $p_{c,l} = 0$, i.e., a vector of N_D possible dummy symbols. Then, expression (5) can be compactly expressed as

$$\mathbf{s}_n = \mathbf{s}_n^{\text{data}} + \mathbf{A}_n \tilde{\mathbf{m}}, \quad (6)$$

where the matrix $\mathbf{A}_n$ is obtained by concatenating the columns of matrix $F^{\text{H}} \left[\text{diag}(\tilde{\mathbf{V}}_{1,n}), \ldots, \text{diag}(\tilde{\mathbf{V}}_{L,n}) \right]$ that correspond to indices (c, l) for which $p_{c,l} = 0$. Note that indices (c, l) select the c-th column of matrix $F^{\text{H}} \text{diag}(\tilde{\mathbf{V}}_{l,n})$.

III. PROBLEM FORMULATION

Our objective is to minimize the PAPR of an existing time-domain transmit signal in a power-efficient manner. This is done by minimizing the ratio between the transmit signal power peak and the average transmit power, given by $\|\mathbf{s}_n\|_\infty^2$ and $(1/\alpha N_C)\|\mathbf{s}_n\|_2^2$, respectively, for each antenna n. The PAPR minimization is achieved by exploiting unused subcarriers to transmit dummy symbols, which affect the final transmit waveform $\mathbf{s}_n, \forall n$, but which also require additional power. Therefore, in addition to finding dummy symbols that reduce the PAPR of the transmitted waveforms, we also aim to jointly minimize the power reserved for these symbols.

With expression (6), the crest factor (cf, i.e., linear domain equivalent of PAPR) of the signal transmitted by the n-th

---

[2] The non-interference consideration only needs to apply for possible dummy symbols interfering with the data transmission. As the dummy symbols do not carry information, they can still be interfered by the data signal. This can be exploited in beamformer optimization, see Section V.

antenna element can be expressed as

$$p^{\text{cf}}(\mathbf{s}_n) = \frac{\alpha N_C \|\mathbf{s}_n\|_\infty^2}{\|\mathbf{s}_n\|_2^2}. \quad (7)$$

The allocation of dummy symbols requires increased transmit power, similar to the data symbols. We call this additional power increase the dummy symbol power reserve, defined as a factor by which the already allocated data power is increased to enable the dummy symbols. To this end, let p^{data} denote the total allocated data power (see expression (1)), i.e,

$$p^{\text{data}} = \sum_{c=1}^{N_C} \sum_{l=1}^{L} p_{c,l}. \quad (8)$$

Then, the factor of increase is given by

$$p^{\text{res}}(\tilde{\mathbf{m}}) = \frac{p^{\text{data}} + \|\tilde{\mathbf{m}}\|_2^2}{p^{\text{data}}} = 1 + \frac{\|\tilde{\mathbf{m}}\|_2^2}{p^{\text{data}}}, \quad (9)$$

where $\|\tilde{\mathbf{m}}\|_2^2$ is the total extra transmission power due to the placement of dummy symbols.

Using (7) and (9) the joint PAPR and power reserve minimization can be cast as the following optimization problem:

$$\text{minimize} \quad \log(\max_n(p^{\text{cf}}(\mathbf{s}_n))) + \log(p^{\text{res}}(\tilde{\mathbf{m}})) \quad (10\text{a})$$

$$\text{subject to} \quad \mathbf{s}_n = \mathbf{s}_n^{\text{data}} + \mathbf{A}_n \tilde{\mathbf{m}}, \forall n \quad (10\text{b})$$

with variables $\{\tilde{\mathbf{m}}, \mathbf{s}_{n=1,\ldots,N_T}\}$. We use the max-operator in the objective to consider the worst-case PAPR among all antennas. We also utilize the log-transformation in the objective since power amplifier input-output characteristics are usually analyzed in logarithmic domain, where the PAPR can be straightforwardly summed to the input power. Similarly, the factor of increase on the power allocation, caused by the dummy symbols, can be summed on the input power, provided that the power reserve is translated into log-domain. This power summation objective is illustrated in Figure 1 for ideal amplifiers. In practice, the amplification curve is not linear (i.e., signals with low mean power but high PAPR get distorted even in the linear region), and the saturation happens gradually.

IV. ALGORITHM DERIVATION

By substituting the expressions (7) and (9) for $p^{\text{cf}}(\mathbf{s}_n)$ and $p^{\text{res}}(\tilde{\mathbf{m}})$, problem (10) can be written equivalently as

$$\text{minimize} \quad \log(t) + \log(r) \quad (11\text{a})$$

$$\text{subject to} \quad \frac{\alpha N_C \|\mathbf{s}_n\|_\infty^2}{\|\mathbf{s}_n\|_2^2} \leq t, \forall n,$$

$$\mathbf{s}_n = \mathbf{s}_n^{\text{data}} + \mathbf{A}_n \tilde{\mathbf{m}}, \forall n, \quad 1 + \frac{1}{p^{\text{data}}} \|\tilde{\mathbf{m}}\|_2^2 \leq r, \quad (11\text{b})$$

with variables $\{\tilde{\mathbf{m}}, t, r, \mathbf{s}_{n=1,\ldots,N_T}\}$. Problem (11) is non-convex due to the concave objective (11a) and the convex-over-convex PAPR constraint (11b). In this section, we propose and derive three different algorithms, with varying PAPR reduction performance and computational complexity, to find (sub-optimal) solutions for (11).
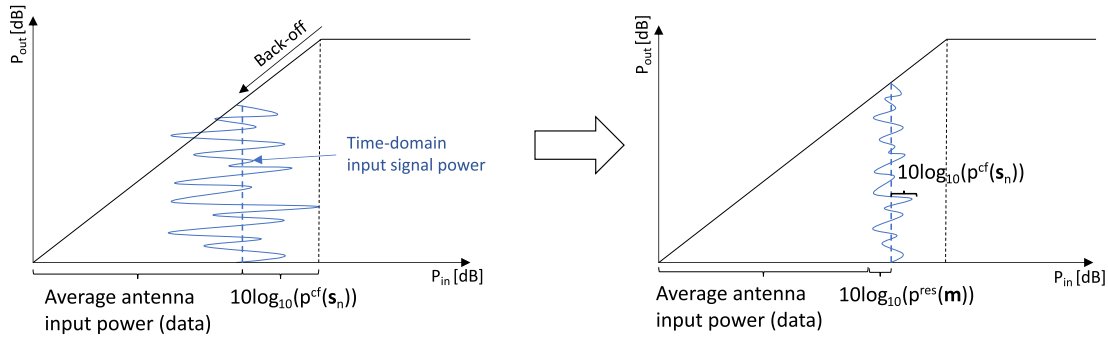
FIGURE 1. Illustration of the goal of problem (10), presented in terms of ideal power amplifier characteristics. The average antenna input power includes the effect of beamforming. Left: Unoptimized input signal with high PAPR, where sufficient back-off is taken that no saturation occurs. Right: Optimized input with some additional power allocated for dummy symbols; due to the low PAPR, back-off could be reduced.

The first algorithm utilizes first-order Taylor series approximations for the non-convex expressions of (11) and solves a linearized problem iteratively using successive convex approximation. The second algorithm ignores average antenna transmit powers and replaces the logarithmic relationship in (10a) with a trade-off coefficient. This solution structure corresponds to many tone reservation algorithms in the literature [13]–[19], [21], but our approach includes also the dummy symbol power reserve function in the objective, and the solution is obtained via ADMM. The last proposed algorithm combines the two approaches to find an iterative heuristic solution.

A. ITERATIVE ALGORITHM VIA SUCCESSIVE CONVEX APPROXIMATION

Problem (11) is non-convex due to the concave objective function (11a) and non-convex constraint (11b), which makes solving (11) computationally intractable.

To handle the non-convexities, let us first introduce slack variables $w_n, q_n, \forall n$, and extra (relaxed) constraints $\|\mathbf{s}_n\|_\infty \leq w_n, \forall n$, and $q_n \leq \|\mathbf{s}_n\|_2^2, \forall n$. This allows us to reformulate constraint (11b) as quadratic over linear, resulting in the following equivalent problem:

$$\text{minimize} \quad \log(t) + \log(r) \tag{12a}$$

$$\text{subject to} \quad \frac{\alpha N_C w_n^2}{q_n} \leq t, \forall n, \quad \|\mathbf{s}_n\|_\infty \leq w_n, \forall n,$$

$$\mathbf{s}_n = \mathbf{s}_n^{\text{data}} + \mathbf{A}_n \tilde{\mathbf{m}}, \forall n, \quad 1 + \frac{1}{p^{\text{data}}}\|\tilde{\mathbf{m}}\|_2^2 \leq r,$$

$$q_n \leq \|\mathbf{s}_n\|_2^2, \forall n \tag{12b}$$

with variables $\{\tilde{\mathbf{m}}, t, r\}$ and $\{w_n, q_n, \mathbf{s}_n\}_{n=1,\dots,N_T}$. Problem (12) is still non-convex due to the concave objective and constraints (12b), which are non-convex due to the convex R.H.S. terms. To obtain a problem which has a tractable solution, we need to iteratively approximate these non-convex parts using first-order Taylor series approximation.

The first-order Taylor series approximations on the objective (12a) and constraints (12b) around a fixed local point $\{\hat{t}, \hat{r}, \hat{\mathbf{s}}_{n=1,\dots,N_T}\}$ are given by

$$\log(t) + \log(r) \leq \log(\hat{t}) + \log(\hat{r}) + \frac{1}{\hat{t}}(t - \hat{t}) + \frac{1}{\hat{r}}(r - \hat{r}) \tag{13}$$

$$\|\mathbf{s}_n\|_2^2 \geq \|\hat{\mathbf{s}}_n\|_2^2 + 2\text{Re}[\hat{\mathbf{s}}_n^H(\mathbf{s}_n - \hat{\mathbf{s}}_n)], \forall n. \tag{14}$$

Replacing (12a) and the R.H.S. of (12b) with the above approximations, we can write a convex approximation of problem (11) as

$$\text{minimize} \quad \frac{t}{\hat{t}} + \frac{r}{\hat{r}} \tag{15a}$$

$$\text{subject to} \quad \frac{\alpha N_C w_n^2}{q_n} \leq t, \forall n, \quad \|\mathbf{s}_n\|_\infty \leq w_n, \forall n,$$

$$1 + \frac{1}{p^{\text{data}}}\|\tilde{\mathbf{m}}\|_2^2 \leq r,$$

$$q_n \leq \|\hat{\mathbf{s}}_n\|_2^2 + 2\text{Re}[\hat{\mathbf{s}}_n^H(\mathbf{s}_n - \hat{\mathbf{s}}_n)], \forall n, \tag{15b}$$

$$\mathbf{s}_n = \mathbf{s}_n^{\text{data}} + \mathbf{A}_n \tilde{\mathbf{m}}, \forall n, \tag{15c}$$

with variables $\{\tilde{\mathbf{m}}, t\}$ and $\{w_n, q_n, \mathbf{s}_n\}_{n=1,\dots,N_T}$. Problem (15) is convex and can be solved via interior point methods. The solution is then used to update the local point $\{\hat{t}, \hat{r}, \hat{\mathbf{s}}_{n=1,\dots,N_T}\}$, after which a new solution that is a closer approximation to a solution of (11) can be found. This is repeated until the difference between objective values from one iteration to the next falls under some threshold $\epsilon^{\text{SCA}} > 0$, or the number of iterations reaches a given maximum j^{SCA}. The objective value of (15) can be guaranteed to decrease between SCA iterations due to the linear upper bound approximation [31]. However, since the original problem is non-convex, convergence to the globally optimal solution cannot be guaranteed. The joint PAPR and dummy symbol power reserve minimization algorithm is summarized in Algorithm 1.

B. LOW-COMPLEXITY PEAK POWER MINIMIZATION VIA ADMM

Finding a solution to problem (11) using SCA and interior point methods is computationally taxing, and thus, not practical. Therefore, finding a low-complexity algorithm with comparable PAPR reduction performance is necessary.

Another approach to tackling the non-convexity of problem (11) is to only consider the numerator of the non-convex PAPR constraint (11b), i.e., peak power minimization. This principle of tackling PAPR minimization has been considered in the literature as it is much simpler than utilizing the strict PAPR expression. Here, we derive the peak power minimization algorithm that also accounts for the dummy symbol

---

Algorithm 1 Joint PAPR and Dummy Symbol Power Reserve Minimization Algorithm

---

1: Input: $\{\mathbf{s}_n^{\text{data}}, \mathbf{A}_n\}_{n=1,\dots,N_T}, p^{\text{data}}$.
2: Set: $\epsilon^{\text{SCA}}, j^{\text{SCA}}, \hat{t} = \max_n(\alpha N_C \|\mathbf{s}_n^{\text{data}}\|_\infty^2 / \|\mathbf{s}_n^{\text{data}}\|_2^2), \hat{r} = 1, \hat{\mathbf{s}}_n = \mathbf{s}_n^{\text{data}}, \forall n, j = 1, O^{(0)} = \infty$.
3: **while** $|O^{(j)} - O^{(j-1)}| > \epsilon^{\text{SCA}}$ **and** $j < j^{\text{SCA}}$ **do**
4:     Solve problem (15).
5:     Updates: $O^{(j)} \leftarrow$ (15a), $\hat{t} \leftarrow t, \hat{r} \leftarrow r, \hat{\mathbf{s}}_n \leftarrow \mathbf{s}_n, \forall n$.
6:     $j = j + 1$.
7: **end while**
8: Output: $\tilde{\mathbf{m}}$.

---

power reserve, as a multi-objective optimization problem that can be solved using ADMM [32].

Peak power minimization ignores the relationship between the time-domain signal peak and the antenna average transmit power. It also ignores the logarithmic relationship between the dummy symbol power reserve $p^{\text{res}}(\tilde{\mathbf{m}})$ and the crest factor $p^{\text{cf}}(\mathbf{s}_n), \forall n$. Therefore, we can consider the objective (10a) without the logarithms, but we require a trade-off coefficient δ to find a good balance between the two functions. Furthermore, as the average transmit powers of the antennas are ignored, we can denote

$$\bar{\mathbf{s}} = [\mathbf{s}_1^T \dots \mathbf{s}_{N_T}^T]^T \in \mathbb{C}^{\alpha N_T N_C \times 1},$$
$$\bar{\mathbf{A}} = [\mathbf{A}_1^T \dots \mathbf{A}_{N_T}^T]^T \in \mathbb{C}^{\alpha N_T N_C \times N_D},$$
$$\bar{\mathbf{s}}^{\text{data}} = [(\mathbf{s}_1^{\text{data}})^T \dots (\mathbf{s}_{N_T}^{\text{data}})^T]^T \in \mathbb{C}^{\alpha N_T N_C \times 1}.$$

With this notation, the peak power minimization problem can be formulated as

$$\underset{\tilde{\mathbf{m}}, \bar{s}}{\text{minimize}} \quad \|\bar{s}\|_\infty + \frac{\delta}{p^{\text{data}}} \|\tilde{\mathbf{m}}\|_2^2 \tag{16a}$$

$$\text{subject to} \quad \bar{s} = \bar{s}^{\text{data}} + \bar{A}\tilde{\mathbf{m}}. \tag{16b}$$

We use ADMM to solve this problem. We start by writing the complex augmented Lagrangian as [33]

$$\mathcal{L}_\rho(\bar{\mathbf{s}}, \tilde{\mathbf{m}}, \boldsymbol{\omega})$$
$$= \|\bar{\mathbf{s}}\|_\infty + \frac{\delta}{p^{\text{data}}} \|\tilde{\mathbf{m}}\|_2^2$$
$$\quad + 2\text{Re}[\boldsymbol{\omega}^H(\bar{\mathbf{s}} - \bar{\mathbf{s}}^{\text{data}} - \bar{\mathbf{A}}\tilde{\mathbf{m}})] + \rho\|\bar{\mathbf{s}} - \bar{\mathbf{s}}^{\text{data}} - \bar{\mathbf{A}}\tilde{\mathbf{m}}\|_2^2$$
$$= \|\bar{\mathbf{s}}\|_\infty + \frac{\delta}{p^{\text{data}}} \|\tilde{\mathbf{m}}\|_2^2 + \rho\|\bar{\mathbf{s}} - \bar{\mathbf{s}}^{\text{data}} - \bar{\mathbf{A}}\tilde{\mathbf{m}} + \frac{1}{\rho}\boldsymbol{\omega}\|_2^2, \tag{17}$$

where $\rho > 0$ is the ADMM penalty parameter and $\boldsymbol{\omega} \in \mathbb{C}^{\alpha N_C N_T \times 1}$ is the vector of dual variables corresponding to constraint (16b). The ADMM steps are [32, Ch. 3]:

$$\bar{\mathbf{s}}^{(k+1)} = \underset{\bar{\mathbf{s}}}{\arg\min} \, \mathcal{L}_\rho(\bar{\mathbf{s}}, \tilde{\mathbf{m}}^{(k)}, \boldsymbol{\omega}^{(k)}) \tag{18}$$

$$\tilde{\mathbf{m}}^{(k+1)} = \underset{\tilde{\mathbf{m}}}{\arg\min} \, \mathcal{L}_\rho(\bar{\mathbf{s}}^{(k+1)}, \tilde{\mathbf{m}}, \boldsymbol{\omega}^{(k)}) \tag{19}$$

$$\boldsymbol{\omega}^{(k+1)} := \boldsymbol{\omega}^{(k)} + \rho(\bar{\mathbf{s}}^{(k+1)} - \bar{\mathbf{s}}^{\text{data}} - \bar{\mathbf{A}}\tilde{\mathbf{m}}^{(k+1)}) \tag{20}$$

Following the approach in [21, Sec. IV], step (18) can be written as an unconstrained scalar optimization problem and

easily solved via one-dimensional search, for example, the well-known golden section search. To see this, we denote the constant vector $\mathbf{z}^{(k)} = \bar{\mathbf{s}}^{\text{data}} + \bar{\mathbf{A}}\tilde{\mathbf{m}}^{(k)} - (1/\rho)\boldsymbol{\omega}^{(k)}$, and write

$$\bar{\mathbf{s}}^{(k+1)} = \underset{\bar{\mathbf{s}}}{\arg\min} \big(\|\bar{\mathbf{s}}\|_\infty + \rho\|\bar{\mathbf{s}} - \mathbf{z}^{(k)}\|_2^2\big)$$
$$= \underset{\bar{\mathbf{s}}}{\arg\min} \big(\|\bar{\mathbf{s}}\|_\infty + \rho \sum_{i=1}^{\alpha N_T N_C} |z_i^{(k)} - \bar{s}_i|^2\big)$$
$$= \underset{\bar{\mathbf{s}}}{\arg\min} \big(\max_i |\bar{s}_i| + \rho \sum_{i=1}^{\alpha N_T N_C} (|z_i^{(k)}| - |\bar{s}_i|)^2\big). \tag{21}$$

The second equality of (21) follows from the definition of the l_2-norm. The third equality is obtained by matching the element-wise phases of vectors $\bar{\mathbf{s}}$ and $\mathbf{z}^{(k)}$ (the phases of $\bar{\mathbf{s}}$ do not affect the l_∞-norm and can be freely chosen). After matching the phases, we only need to optimize and match the element-wise magnitudes of $\bar{\mathbf{s}}$. These can be found via scalar optimization by solving

$$\gamma = \underset{\hat{\gamma}}{\arg\min} \big(\hat{\gamma} + \rho \sum_{i=1}^{\alpha N_T N_C} (|z_i^{(k)}| - \hat{\gamma})_+^2\big), \tag{22}$$

and then set $|\bar{s}_i| = \min(|z_i^{(k)}|, \gamma), \forall i$, i.e., we apply element-wise clipping. To see this, i.e., to get from (21) to (22), we first denote $\max_i |\bar{s}_i| = \hat{\gamma}$. Then, for all i, we can set

$$|z_i^{(k)}| - |\bar{s}_i| = \begin{cases} 0, & \text{if } |z_i^{(k)}| \le \hat{\gamma} \\ |z_i^{(k)}| - \hat{\gamma}, & \text{if } |z_i^{(k)}| > \hat{\gamma}, \end{cases}$$

which characterizes the perfect match in magnitudes if $|z_i^{(k)}| \le \hat{\gamma}$, and the remaining residual otherwise. Another way to write this is $\max\{0, |z_i^{(k)}| - \hat{\gamma}\} = (|z_i^{(k)}| - \hat{\gamma})_+, \forall i$.

Next, we move on to find a solution for step (19). The step has a closed-form solution, as it is a minimization of an unconstrained convex function, i.e., the zero-gradient condition gives the minimum. Taking the gradient of the augmented Lagrangian with respect to $\tilde{\mathbf{m}}$:

$$\frac{\partial \mathcal{L}_\rho(\bar{\mathbf{s}}^{(k+1)}, \tilde{\mathbf{m}}, \boldsymbol{\omega}^{(k)})}{\partial \tilde{\mathbf{m}}^*}$$
$$= \frac{\delta}{p^{\text{data}}} \tilde{\mathbf{m}} - \rho\bar{\mathbf{A}}^H(\bar{\mathbf{s}}^{(k+1)} - \bar{\mathbf{s}}^{\text{data}} - \bar{\mathbf{A}}\tilde{\mathbf{m}} + \frac{1}{\rho}\boldsymbol{\omega}^{(k)}) = \mathbf{0}$$
$$\iff \tilde{\mathbf{m}} = (\frac{\delta}{p^{\text{data}}}\mathbf{I} + \rho\bar{\mathbf{A}}^H\bar{\mathbf{A}})^{-1}\bar{\mathbf{A}}^H(\boldsymbol{\omega}^{(k)} - \rho(\bar{\mathbf{s}}^{\text{data}} - \bar{\mathbf{s}}^{(k+1)})) \tag{23}$$

By iteratively solving the variables $\bar{\mathbf{s}}, \tilde{\mathbf{m}}$ and $\boldsymbol{\omega}$, we can find a solution for problem (16). The iterations are stopped once a convergence criterion is met, or a maximum number of iterations k^{peak} is reached.

To measure convergence, we use the residuals of the primal and dual feasibility conditions of (16), given by [32, Ch. 3.3]

$$(r^{\text{primal}})^{(k)} = \|\bar{\mathbf{s}}^{(k)} - \bar{\mathbf{s}}^{\text{data}} - \bar{\mathbf{A}}\tilde{\mathbf{m}}^{(k)}\|_2, \tag{24}$$
$$(r^{\text{dual}})^{(k)} = \|\rho\bar{\mathbf{A}}(\tilde{\mathbf{m}}^{(k)} - \tilde{\mathbf{m}}^{(k-1)})\|_2. \tag{25}$$

As the ADMM iterations proceed, these residuals converge to zero. To set a practical stopping criterion, we can stop iterating once the residuals fall below a tolerance threshold ϵ^{peak}. The joint peak power and dummy symbol power reserve minimization algorithm is summarized in Algorithm 2.

Algorithm 2 Joint Peak Power and Dummy Symbol Power Reserve Minimization Algorithm

1: Input: $\{\mathbf{s}_n^{\text{data}}, \mathbf{A}_n\}_{n=1,\dots,N_T}$, p^{data}, ϵ^{peak}, k^{peak}, ρ, δ.
2: Set: $\tilde{\mathbf{m}} = \mathbf{0}$, $\boldsymbol{\omega} = \mathbf{0}$, $k = 0$.
3: **while** $\|(r^{\text{primal}})^{(k)}, (r^{\text{dual}})^{(k)}\|_\infty > \epsilon^{\text{peak}}$ **and** $k < k^{\text{peak}}$ **do**
4:     Calculate $\bar{\mathbf{s}}^{(k+1)}$ with (21), $\tilde{\mathbf{m}}^{(k+1)}$ with (23), and update $\boldsymbol{\omega}^{(k+1)}$ with (20).
5:     $k = k + 1$.
6: **end while**
7: Output: $\tilde{\mathbf{m}}$.

C. ITERATIVE RELATIVE PEAK POWER MINIMIZATION VIA ADMM

While the peak power minimization framework of Section IV-B provides a low-complexity solution to reduce the peaks of the transmit waveform signal, the PAPR reduction performance of the algorithm is less effective than the SCA approach. This is due to the fact that the peak power minimization algorithm ignores the average antenna-specific power. Thus, while the peaks of the transmit signal are reduced, the antennas with low mean power can still have substantially high PAPR. In the following, we derive a low-complexity heuristic approach that accounts for the average power, too.

Similar to the peak power minimization, we manipulate the denominator of the non-convex PAPR constraint (11b). Instead of ignoring it completely, we fix the denominator to be a constant that is iteratively updated between each SCA iteration. Starting from problem (11), we denote the denominator of (11b) as $(1/\alpha N_C)\|\mathbf{s}_n\|_2^2 = (1/\alpha N_C)\|\mathbf{s}_n^{\text{data}} + \mathbf{A}_n\tilde{\mathbf{m}}\|_2^2 = \hat{\beta}_n$, $\forall n$, and, by applying first-order Taylor series approximation on the objective, see (13), we can write the relative peak power minimization problem as

$$\text{minimize} \quad \max_n\left(\frac{\|\mathbf{s}_n\|_\infty^2}{\hat{\beta}_n \hat{t}}\right) + \frac{\|\tilde{\mathbf{m}}\|_2^2}{p^{\text{data}}\hat{r}} \quad (26\text{a})$$

$$\text{subject to} \quad \mathbf{s}_n = \mathbf{s}_n^{\text{data}} + \mathbf{A}_n\tilde{\mathbf{m}}, \forall n, \quad (26\text{b})$$

with variables $\{\tilde{\mathbf{m}}, \mathbf{s}_{n=1,\dots,N_T}\}$. Note that we also rolled back the epigraph formulation of (11) to reduce the amount of optimization variables. Next, we apply a change of variables $\tilde{\mathbf{s}}_n = \hat{\beta}_n^{-\frac{1}{2}}\mathbf{s}_n$, $\forall n$, and formulate an equivalent problem as

$$\text{minimize} \quad \max_n\left(\frac{\|\tilde{\mathbf{s}}_n\|_\infty^2}{\hat{t}}\right) + \frac{\|\tilde{\mathbf{m}}\|_2^2}{p^{\text{data}}\hat{r}} \quad (27\text{a})$$

$$\text{subject to} \quad \tilde{\mathbf{s}}_n = \frac{1}{\sqrt{\hat{\beta}_n}}(\mathbf{s}_n^{\text{data}} + \mathbf{A}_n\tilde{\mathbf{m}}), \forall n, \quad (27\text{b})$$

with variables $\{\tilde{\mathbf{m}}, \tilde{\mathbf{s}}_{n=1,\dots,N_T}\}$. This problem is convex around a local point $\{\hat{t}, \hat{r}, \hat{\beta}_{n=1,\dots,N_T}\}$ and can be solved via ADMM.

The complex augmented Lagrangian of (27) is given by [33]

$$\mathcal{L}_\sigma(\tilde{\mathbf{m}}, \{\tilde{\mathbf{s}}_n, \boldsymbol{v}_n\}_{n=1,\dots,N_T}) = \frac{\max_n \|\tilde{\mathbf{s}}_n\|_\infty^2}{\hat{t}} + \frac{\|\tilde{\mathbf{m}}\|_2^2}{p^{\text{data}}\hat{r}}$$

$$+ \sigma\sum_{n=1}^{N_T}\|\tilde{\mathbf{s}}_n - \frac{1}{\sqrt{\hat{\beta}_n}}(\mathbf{s}_n^{\text{data}} + \mathbf{A}_n\tilde{\mathbf{m}}) + \frac{1}{\sigma}\boldsymbol{v}_n\|_2^2, \quad (28)$$

where $\sigma > 0$ is the ADMM penalty parameter, and $\boldsymbol{v}_n \in \mathbb{C}^{\alpha N_C \times 1}$, $\forall n$, are the dual variable vectors corresponding to constraints (27b). Similar to (17), the dual variable vectors have been moved inside the l_2-norm. The ADMM steps are given by [32, Ch. 3]

$$\tilde{\mathbf{s}}_n^{(k+1)} = \underset{\tilde{\mathbf{s}}_n}{\arg\min}\ \mathcal{L}_\sigma(\tilde{\mathbf{m}}^{(k)}, \{\tilde{\mathbf{s}}_n, \boldsymbol{v}_n^{(k)}\}_{n=1,\dots,N_T}), \forall n \quad (29)$$

$$\tilde{\mathbf{m}}^{(k+1)} = \underset{\tilde{\mathbf{m}}}{\arg\min}\ \mathcal{L}_\sigma(\tilde{\mathbf{m}}, \{\tilde{\mathbf{s}}_n^{(k+1)}, \boldsymbol{v}_n^{(k)}\}_{n=1,\dots,N_T}) \quad (30)$$

$$\boldsymbol{v}_n^{(k+1)} := \boldsymbol{v}_n^{(k)} + \sigma(\tilde{\mathbf{s}}_n^{(k+1)} - \frac{1}{\sqrt{\hat{\beta}_n}}(\mathbf{s}_n^{\text{data}} + \mathbf{A}_n\tilde{\mathbf{m}}^{(k+1)})) \quad (31)$$

ADMM steps (29) and (30) can be simplified in a similar manner to Section IV-B. Starting with (29), let us denote the constant vectors $\tilde{\mathbf{z}}_n^{(k)} = (\hat{\beta}_n)^{-\frac{1}{2}}(\mathbf{s}_n^{\text{data}} - \mathbf{A}_n\tilde{\mathbf{m}}^{(k)}) - \sigma^{-1}\boldsymbol{v}_n^{(k)}$, $\forall n$. We also note that $\max_n \|\tilde{\mathbf{s}}_n\|_\infty^2 = \max_n \max_i |\tilde{s}_{i,n}|^2 = \max_{i,n}|\tilde{s}_{i,n}|^2$.

Due to the max-operator over all n, the vectors $\tilde{\mathbf{s}}_n$, $\forall n$, have to be jointly optimized. To this end, ADMM step (29) is re-written as

$$\tilde{\mathbf{s}}_{n=1,\dots,N_T}^{(k+1)} = \underset{\tilde{\mathbf{s}}_{n=1,\dots,N_T}}{\arg\min}\left(\max_{i,n}\frac{|\tilde{s}_{i,n}|^2}{\hat{t}} + \sigma\sum_{i=1}^{\alpha N_C}\sum_{n=1}^{N_T}|\tilde{z}_{i,n}^{(k)} - \tilde{s}_{i,n}|^2\right), \quad (32)$$

which can be solved via scalar optimization and clipping, following the same chain of arguments as in (21).

Step (30) is an unconstrained minimization of a convex function, and has a closed-form solution via setting the gradient to zero. To simplify the notation, let us denote $\mathbf{q}_n^{(k)} = \tilde{\mathbf{s}}_n^{(k+1)} - (\hat{\beta}_n)^{-\frac{1}{2}}\mathbf{s}_n^{\text{data}} + \sigma^{-1}\boldsymbol{v}_n^{(k)}$. Then, taking the gradient of the Lagrangian (28) with respect to $\tilde{\mathbf{m}}$:

$$\frac{\partial \mathcal{L}_\sigma(\tilde{\mathbf{m}}, \mathbf{q}_{n=1,\dots,N_T}^{(k)})}{\partial\tilde{\mathbf{m}}^*}$$

$$= \frac{1}{p^{\text{data}}\hat{r}}\tilde{\mathbf{m}} - \sigma\sum_{n=1}^{N_T}\frac{1}{\sqrt{\hat{\beta}_n}}\mathbf{A}_n^H\left(\mathbf{q}_n^{(k)} - \frac{1}{\sqrt{\hat{\beta}_n}}\mathbf{A}_n\tilde{\mathbf{m}}\right) = \mathbf{0}$$

$$\iff \tilde{\mathbf{m}} = \left(\frac{1}{p^{\text{data}}\hat{r}}\mathbf{I} + \sigma\sum_{n=1}^{N_T}\frac{1}{\hat{\beta}_n}\mathbf{A}_n^H\mathbf{A}_n\right)^{-1}$$

$$\times\left(\sigma\sum_{n=1}^{N_T}\frac{1}{\sqrt{\hat{\beta}_n}}\mathbf{A}_n^H\mathbf{q}_n^{(k)}\right). \quad (33)$$

By iteratively solving the variables $\tilde{\mathbf{m}}, \{\tilde{\mathbf{s}}_n, \boldsymbol{v}_n\}_{n=1,\dots,N_T}$, we can find a solution for problem (26). The iterations are stopped once a convergence criteria is met, or a maximum number of iterations $k^{\text{rel-peak}}$ is reached.

Similar to Algorithm 2, we use the residuals of the primal and dual feasibility conditions, given by [32, Ch. 3.3]

$$(r^{\text{primal}})^{(k)} = \sum_{n=1}^{N_{\text{T}}} \|\tilde{\mathbf{s}}_n^{(k)} - \frac{1}{\sqrt{\hat{\beta}_n}}(\mathbf{s}_n^{\text{data}} + \mathbf{A}_n\tilde{\mathbf{m}}^{(k)})\|_2, \quad (34)$$

$$(r^{\text{dual}})^{(k)} = \sum_{n=1}^{N_{\text{T}}} \|\sigma\mathbf{A}_n(\tilde{\mathbf{m}}^{(k)} - \tilde{\mathbf{m}}^{(k-1)})\|_2, \quad (35)$$

to measure the convergence. We assume the algorithm is sufficiently converged once both of the residuals fall below a set tolerance threshold $\epsilon^{\text{rel-peak}}$. However, unlike Algorithm 2, we also require an outer loop for the SCA updates, i.e., the ADMM solution has to be found for each SCA step. However, by using the solutions of the previous SCA step to initialize the subsequent ADMM call, the number of iterations is significantly reduced for quick convergence. The joint relative peak power and dummy symbol power reserve minimization algorithm is presented in Algorithm 3.

D. DISCUSSION ON COMPUTATIONAL COMPLEXITY

Algorithm 1 requires solving the convex optimization problem (15) around each local point $\{\hat{t}, \hat{r}, \hat{\mathbf{s}}_{n=1,...,N_{\text{T}}}\}$, which can be done with interior point methods. This solution structure requires solving Newton steps, which include the inversion of a matrix with size proportional to the number of primal and dual optimization variables. The number of primal variables in (15) can be reduced to $N_{\text{D}} + N_{\text{T}}$ by rolling back the epigraph formulation and backward substituting constraints (15b) and (15c). The number of dual variables is then N_{T}, and the computational complexity of solving one Newton step is in the order of $\mathcal{O}((N_{\text{D}} + 2N_{\text{T}})^3)$. Running Algorithm 1 requires solving multiple SCA steps, each requiring multiple Newton steps, which becomes practically infeasible for any large number of dummy symbols N_{D}.

As low-complexity alternatives, Algorithms 2 and 3 require solving the problems (16) and (26), respectively, which are done with ADMM instead of interior point methods. This solution structure requires iteratively solving ADMM steps, multiple times for Algorithm 3 due to the SCA structure. We can consider the computational complexity of solving the ADMM steps used in Algorithm 2, i.e., (21), (23), and (20), and similar consideration will hold for Algorithm 3.

Solving step (21) is a simple one-dimensional search followed by a clipping operation, and step (20) is just a subgradient update. The most complexity seems to be in step (23), which requires the inversion of a symmetric matrix. For the inversion, we can distinguish two cases: when the dummy symbols are assigned on a single stream, or on multiple streams. For the single stream allocation, the non-diagonal term $\bar{\mathbf{A}}^{\text{H}}\bar{\mathbf{A}}$ can be easily simplified to a diagonal matrix, and thus, the inversion is trivial. Then, the most computationally intensive calculations are the evaluation of the matrix-vector products $\bar{\mathbf{A}}\tilde{\mathbf{m}}$ and the part of (23) following the matrix inverse. These products can easily be shown to be Fourier-transforms, with computational complexity in the

Algorithm 3 Joint Relative Peak Power and Dummy Symbol Power Reserve Minimization Algorithm

1: **Input:** $\{\mathbf{s}_n^{\text{data}}, \mathbf{A}_n\}_{n=1,...,N_{\text{T}}}$, p^{data}, ϵ^{SCA}, $\epsilon^{\text{rel-peak}}$, j^{SCA}, $k^{\text{rel-peak}}$, σ.
2: Set: $O^{(0)} = \infty$, $j = 1$, $\hat{t} = \max_n(\alpha N_{\text{C}}\|\mathbf{s}_n^{\text{data}}\|_\infty^2/\|\mathbf{s}_n^{\text{data}}\|_2^2)$, $\hat{r} = 1$, $\tilde{\mathbf{m}} = \mathbf{0}$, $\boldsymbol{\nu}_{n=1,...,N_{\text{T}}} = \mathbf{0}$.
3: **while** $|O^{(j-1)} - O^{(j)}| > \epsilon^{\text{SCA}}$ **and** $j < j^{\text{SCA}}$ **do**
4:      Update $\hat{\beta}_n = (1/\alpha N_{\text{C}})\|\mathbf{s}_n^{\text{data}} + \mathbf{A}_n\tilde{\mathbf{m}}\|_2^2$, $\forall n$.
5:      Set $k = 0$.
6:      **while** $\|(r^{\text{primal}})^{(k)}, (r^{\text{dual}})^{(k)}\|_\infty > \epsilon^{\text{rel-peak}}$ **and** $k < k^{\text{rel-peak}}$ **do**
7:          Calculate $\tilde{\mathbf{s}}_n^{(k+1)}$, $\forall n$, with (32), $\tilde{\mathbf{m}}^{(k+1)}$ with (33), and update $\boldsymbol{\nu}_n^{(k+1)}$, $\forall n$ with (31).
8:          $k = k + 1$.
9:      **end while**
10:      Calculate $t = \max_n \|\tilde{\mathbf{s}}_n\|_\infty^2$, and $r = (1 + \frac{1}{p^{\text{data}}}\|\tilde{\mathbf{m}}\|_2^2)$.
11:      Updates: $O^{(j)} \leftarrow$ (27a), $\hat{t} \leftarrow t$, $\hat{r} \leftarrow r$.
12:      $j = j + 1$.
13: **end while**
14: **Output:** $\tilde{\mathbf{m}}$.

order $\mathcal{O}(\alpha N_{\text{C}}\log(\alpha N_{\text{C}}))$, if fast Fourier transform is implemented. If the dummy symbols are allocated on multiple streams, the non-diagonal term in the inversion can easily be shown to be a (Hermitian) striped matrix. If we denote the number of allocated dummy symbol streams on subcarrier c with L_c^{dummy}, then the number of non-zero elements on each row of the striped matrix is less than or equal to $L^{\text{dummy-max}} = \max_c(L_c^{\text{dummy}})$. Thus, the inversion can be undertaken by solving a sparse system of equations. We can also reorder the rows and columns of the striped matrix, to obtain a block-diagonal matrix with N_{C} blocks of size $L^{\text{dummy-max}} \times L^{\text{dummy-max}}$ at most, which can be inverted separately. The same arguments hold also for Alg. 3.

The steps are solved iteratively until a desired level of convergence, or a maximum iteration count, is reached. Therefore, the number of required ADMM iterations depends on the choice of ϵ^{peak}, k^{peak} for Algorithm 2 (or $\epsilon^{\text{rel-peak}}$, $k^{\text{rel-peak}}$ for Algorithm 3). The convergence rate of ADMM also depends on the penalty parameter ρ (or σ), which is a design parameter. High values for the penalty parameter set emphasis on the fast convergence of the primal residual, and low values put focus on the dual residual. In our simulations, we found that the dual residual tends to converge much slower, so low values for the penalty parameter were chosen.

Figure 2 presents a numerical example simulated under an arbitrary channel realization with $N_{\text{D}} = 64$ dummy symbols, providing insight into the required number of iterations (the full details of the system setup are given in Section VII). The top figure represents the decay of the primal and dual residuals, (34) and (35), for the first SCA step, up to a tolerance threshold of 10^{-3}. The bottom figure represents the number of ADMM iterations required for each SCA step. It can be seen that in the beginning, as the SCA point can significantly
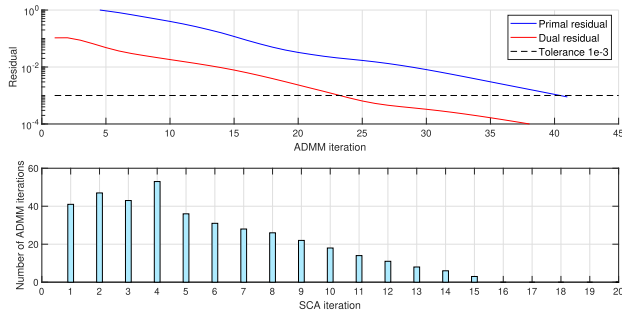
FIGURE 2. Numerical example on required number of iterations for Alg. 3. Top: Decay of primal and dual residuals during ADMM iterations. Bottom: Required number of ADMM iterations per SCA step.

change, ADMM requires more iterations to converge to the threshold. Then, as the SCA local point gets closer to a local optimum, the ADMM converges quicker. It should be noted that the threshold of 10^{-3} is overly strict for this application, and in practical applications a lower tolerance level (10^{-2} to 10^{-1}) would suffice. Furthermore, for the first SCA steps, a lower limit can be set on the maximum number of ADMM iterations to speed up the early (coarse) convergence to a local optimum. Finally, it is obvious that the number of ADMM iterations increases as the number of antennas and dummy symbols increase, as finding a consensus for (larger) $\tilde{\mathbf{m}}$ that works the best for all N_T antennas gets more difficult.

Motivated by the numerical example, we should also highlight that setting a good initial point can reduce the number of SCA steps. For Algorithm 1, this requires finding a good initial dummy symbol vector $\tilde{\mathbf{m}}$, i.e., selecting over space $\mathbb{C}^{N_D}$, which is used to calculate $\{\hat{t}, \hat{r}, \hat{\mathbf{s}}_{n=1,\dots,N_T}\}$. In contrast, for Algorithm 3, the initial point $\{\hat{t}, \hat{r}, \hat{\beta}_{n=1,\dots,N_T}\}$ consists of only $2 + N_T$ real variables. Furthermore, the initial $\{\hat{t}, \hat{r}\}$ should be set dynamically based on the averages of previous solutions, especially in time-correlated channels where the transmission parameters can be assumed to change slowly. Judging by the bottom graph of Figure 2, this would also significantly reduce the required ADMM iterations, as only the last SCA steps would be required.

V. DUMMY SYMBOL AWARE BEAMFORMER DESIGN FOR DOWNLINK MULTI-USER MIMO

The joint PAPR and dummy symbol power reserve minimization framework operates on any arbitrary beamformers and data signals, provided there are free space-frequency resources for dummy symbol allocation. In our previous work [24], [25], we considered an uplink scenario where joint transmission between cooperating users was considered, assuming low angular separation between the users, resulting in low-gain streams that could be exploited through dummy symbol allocations. In this article, we consider a downlink scenario and the corresponding beamformer design.

We consider a multi-user MIMO-OFDM system where a base station with N_T antennas serves K users with N_r antennas each, on N_C subcarriers. The number of antennas can be generalized, but for simplicity, we consider a case where all the users have identical configurations. User k is scheduled with

$L_{c,k}^{\text{data}}$ data streams on subcarrier c. Furthermore, we assume that the number of spatial modes for dummy symbol allocation $L_c^{\text{dummy}}, \forall c$, is known (for example imposed by a standard or defined by the network operator), and the users are scheduled with equal amount of space-frequency resources based on this knowledge. Thus, the total number of streams is $\sum_{k=1}^{K} L_{c,k}^{\text{data}} + L_c^{\text{dummy}} = L, \forall c$.

Finally, as we utilize the Hughes-Hartogs bit and power loading algorithm [28], [29] to construct the transmit signal, a strict no-interference constraint is set in order to decouple beamformer design and power allocation. However, this constraint is not necessary in general.

Our approach utilizes a novel take on an iterative block-diagonalization algorithm [34], [26, Sec. IV], [35] to obtain non-interfering transmit and receiver beamformers.[3] The first novel aspect comes from the exploitation of known dummy symbol allocations. As the resources allocated for dummy symbols can have arbitrarily low gain, this allows us to optimize the data precoders with increased dimension (i.e., project the data into a larger dimensional interference nullspace), simply by ignoring the dummy symbol precoders at this point. Thus, we propose scheduling the users with less than maximum spatial modes, and using the remaining modes to minimize the PAPR through dummy symbol allocations. This scheduling scheme can be easily justified, for example, in scenarios where the user-specific channel matrix is poorly conditioned, i.e., the majority of the achieved rate is obtained by using only the best eigenmode for every user.[4]

The remaining spatial modes on the transmit side can be exploited to reduce the PAPR via transmitting dummy symbols. Furthermore, the beamforming gains of these excess spatial modes can be made negligibly small (in expense of maximizing the gains of the data-carrying spatial modes), as dummy symbols are not meant to carry information. This is done by projecting the dummy symbol precoders into the nullspace of the already optimized data. This is the second novel aspect of our approach. Next, we'll briefly reproduce the iterative block-diagonalization process to find the transmit and receive beamformers, described in more detail in [26, Sec. IV].

The detected signal of user k on subcarrier c is given by

$$\hat{d}_{c,k} = \mathbf{U}_{c,k}^{H} \mathbf{H}_{c,k} \mathbf{V}_c P_c^{\frac{1}{2}} d_c + \mathbf{U}_{c,k}^{H} n_{c,k}, \quad (36)$$

where $\mathbf{U}_{c,k} \in \mathbb{C}^{N_r \times L_{c,k}^{\text{data}}}$ is the user-specific receive beamformer, $\mathbf{H}_{c,k} \in \mathbb{C}^{N_r \times N_T}$ is an arbitrary channel matrix between the base station and user k, and $n_{c,k}$ is the additive white Gaussian noise vector.

---

[3]The beamformers are calculated jointly at the base station and therefore would need to be signaled to the users. However, as the block-diagonalization algorithm constructs beamformers with equal performance to zero-forcing, in practice the receiver could utilize minimum mean square error (MMSE) detectors for better results without the need for extra signaling [26].

[4]In non-correlated channels, the allocation of data and dummy symbols is a trade-off between the benefits of PAPR minimization and throughput, as the dummy symbols can reserve resources better suited for data transmission.

The iterative block-diagonalization algorithm is based on a series of consecutive nullspace projections that eliminate interference between streams. The transmit and receive beamformers of the data streams are optimized separately for all subcarriers c, exploiting the L_c^{dummy} spatial modes allocated for the dummy symbols as additional free dimensions of the interference-free subspace, which results in increased beamforming gain for the data streams (which can freely interfere the streams carrying dummy symbols).

First, user-specific effective channels are constructed using the left eigenvectors of the channel matrix as $\bar{\mathbf{H}}_{c,k} = \mathbf{U}_{c,k}^{\text{H}}\mathbf{H}_{c,k} \in \mathbb{C}^{L_{c,k}^{\text{data}} \times N_{\text{T}}}$, where the columns of $\mathbf{U}_{c,k}$ are chosen to correspond to the highest $L_{c,k}^{\text{data}}$ eigenvalues of the channel matrix. Then, considering user k, the interfering users' effective channels are concatenated as

$$\tilde{\bar{\mathbf{H}}}_{c,k} = [\bar{\mathbf{H}}_{c,1}^{\text{T}} \ldots \bar{\mathbf{H}}_{c,k-1}^{\text{T}} \; \bar{\mathbf{H}}_{c,k+1}^{\text{T}} \ldots \bar{\mathbf{H}}_{c,K}^{\text{T}}]^{\text{T}}, \quad (37)$$

that we can use to find the last $L_{c,k}^{\text{data}}$ right eigenvectors of $\tilde{\bar{\mathbf{H}}}_{c,k}$ to form an orthogonal basis for its nullspace $\tilde{\mathbf{V}}_{c,k}^{\text{null}}$, i.e., an interference-free basis. Next, by projecting the channel matrix of user k into this interference-free domain, i.e., $\hat{\mathbf{H}}_{c,k} = \mathbf{H}_{c,k}\tilde{\mathbf{V}}_{c,k}^{\text{null}}$, we can obtain the left and right eigenvectors corresponding to the $L_{c,k}^{\text{data}}$ strongest eigenvalues as $\hat{\mathbf{U}}_{c,k}^{\text{data}}$ and $\hat{\mathbf{V}}_{c,k}^{\text{data}}$, respectively, from the singular value decomposition of $\hat{\mathbf{H}}_{c,k}$. The interference-free transmit and receive beamformers for user k are given by $\mathbf{V}_{c,k} = \tilde{\mathbf{V}}_{c,k}^{\text{null}}\hat{\mathbf{V}}_{c,k}^{\text{data}} \in \mathbb{C}^{N_{\text{T}} \times L_{c,k}^{\text{data}}}$ and $\mathbf{U}_{c,k} = \hat{\mathbf{U}}_{c,k}^{\text{data}}$, respectively. This process is applied for all users in parallel, and to further increase the beamforming gain, the receive beamformer can be used to initialize subsequent iterations following the same process.

Next, we utilize a strategy similar to successive block-diagonalization [27] to obtain the dummy stream precoders, which are projected into the nullspace of the already optimized data. This way, the dummy streams will not interfere with the data streams, although the opposite can happen. However, as the dummy symbols are ignored in the receiver, this interference does not affect the system performance.

The dummy stream precoders can be found, on subcarrier c, from the last $L_c^{\text{dummy}} > 0$ right eigenvectors of

$$[\bar{\mathbf{H}}_{c,1}^{\text{T}} \ldots \bar{\mathbf{H}}_{c,K}^{\text{T}}]^{\text{T}} = \bar{\mathbf{U}}_c \bar{\boldsymbol{\Sigma}}_c [\bar{\mathbf{V}}_c^{\text{data}} \; \bar{\mathbf{V}}_c^{\text{null}}]^{\text{H}}, \quad (38)$$

denoted as $\bar{\mathbf{V}}_c^{\text{dummy}} \in \mathbb{C}^{N_{\text{T}} \times L_c^{\text{dummy}}}$. Using this notation, the complete final transmit beamformer is given by the concatenation

$$\mathbf{V}_c = [\mathbf{V}_{c,1} \ldots \mathbf{V}_{c,K} \; \bar{\mathbf{V}}_c^{\text{dummy}}] \in \mathbb{C}^{N_{\text{T}} \times L}, \forall c. \quad (39)$$

The iterative block-diagonalization algorithm is summarized in Algorithm 4.

VI. RATE MAXIMIZATION SEARCH ALGORITHM

There are multiple benefits to reduced PAPR, for example, less strict linearity requirements and reduced signal distortion of power amplifiers when using a high input power operating

---

Algorithm 4 Iterative Block-Diagonalization Algorithm for Data and Dummy Beamformer Optimization

1: Input: $\mathbf{H}_{c,k}, L_{c,k}^{\text{data}}, L_c^{\text{dummy}}, \forall c, k$.
2: Set: $i^{\text{MAX}}, i = 1$.
3: Select $L_{c,k}^{\text{data}}$ left eigenvectors for initial $\mathbf{U}_{c,k}, \forall c, k$.
4: **repeat** $\{\forall c, k\}$
5:     Calculate effective channels: $\bar{\mathbf{H}}_{c,k} = \mathbf{U}_{c,k}^{\text{H}}\mathbf{H}_{c,k}$.
6:     Using (37), obtain interference-free basis $\tilde{\mathbf{V}}_{c,k}^{\text{null}}$.
7:     Find the interference-free domain $\hat{\mathbf{H}}_{c,k} = \mathbf{H}_{c,k}\tilde{\mathbf{V}}_{c,k}^{\text{null}}$.
8:     Find $L_{c,k}^{\text{data}}$ left and right eigenvectors corresponding to the highest eigenvalues of $\hat{\mathbf{H}}_{c,k}$.
9:     Find beamformers: $\mathbf{V}_{c,k} = \tilde{\mathbf{V}}_{c,k}^{\text{null}}\hat{\mathbf{V}}_{c,k}^{\text{data}}, \mathbf{U}_{c,k} = \hat{\mathbf{U}}_{c,k}^{\text{data}}$.
10:     $i = i + 1$.
11: **until** $i = i^{\text{MAX}}$
12: Find $\bar{\mathbf{V}}_c^{\text{dummy}}$ using (38), $\forall c$.

---

point, where the power efficiency is maximized. The reduced backoff requirement to combat the signal distortion, a benefit obtained with reduced PAPR, could be translated to increased transmit SNR, providing increased cell size.

In this paper, we also consider this additional SNR increase in terms of increasing the transmit data rate under a given peak power constraint P^{tx}. This means allocating as much power to the data power budget p^{data} as possible, while ensuring that the antenna-specific peak power remains below P^{tx} (for example, the dashed black line in Figure 1).

We utilize the well-known Hughes-Hartogs bit and power loading algorithm [28], [29], to construct the transmit signal. The HH algorithm iteratively allocates bits on subcarriers in a greedy manner based on the channel gains. The algorithm provides an optimal bit allocation for a given bit error rate (BER) and total power budget, provided a strict no-interference constraint is imposed between the streams on any given subcarrier. Now, using the iterative block-diagonalization algorithm of Section V, we can obtain beamformers that remove interference.

Our implementation of the HH algorithm follows [29], modified to operate in a multi-antenna system with parallel spatial streams. The algorithm iteratively allocates b^{inc} bits on stream l, subcarrier c, that requires the least amount of incremental power, given by

$$p_{c,l}^{\text{inc}} = (2^{b^{\text{inc}}} - 1)2^{b_{c,l}} \frac{5}{8} \frac{|\tilde{\mathbf{U}}_{c,l}^{\text{H}}\mathbf{H}_{c,k}\mathbf{V}_{c,l}|^2}{N_0 B} \ln\left(\frac{1}{5\Gamma_{c,l}}\right), \quad (40)$$

where $b_{c,l}$ is the current bit allocation on subcarrier c of stream l, $\tilde{\mathbf{u}}_{c,l}$ is the l-th column of matrix $[\mathbf{U}_{c,1} \ldots \mathbf{U}_{c,K}] \in \mathbb{C}^{N_{\text{r}} \times \sum_k L_{c,k}^{\text{data}}}$, N_0 is the noise variance, B is the subcarrier bandwidth and $\Gamma_{c,l}$ denotes the subcarrier and stream-wise BER target. The bits are allocated using (40) until all the power available for data allocation is used. Once the bit and power loading is complete, we assign quadrature amplitude modulated (QAM) symbols for the subcarriers

---

Algorithm 5 Rate Maximization Search Algorithm

1: **Input:** $\{\mathbf{U}_{c,k}, \mathbf{H}_{c,k}\}, \forall c, k, \mathbf{V}_c, \forall c, \Gamma_{c,l}, \forall c, l, B, P^{\text{tx}}$.
2: **Set:** $P^{\text{data}} = P^{\text{tx}}/2$.
3: **while** $P_n^{\text{peak}} < P^{\text{tx}}, \forall n$ **do**
4:      Use (40) to iteratively allocate $\{p_{c,l}, b_{c,l}\}, \forall c, l$.
5:      Assign symbols according to $b_{c,l}, \forall c, l$ to find $\mathbf{d}_c, \forall c$.
6:      PAPR reduction using any of Alg. 1-3.
7:      Evaluate (41), $\forall n$, update data power budget with
     $P^{\text{data}} := P^{\text{data}} + P^{\text{tx}} - \max_n(P_n^{\text{peak}})$.
8: **end while**
9: **while** $P_n^{\text{peak}} > P^{\text{tx}}$, for any n **do**
10:      Remove $p_{\bar{c},\bar{l}}^{\text{inc}}$ from $p_{\bar{c},\bar{l}}$, where indices $\{\bar{c}, \bar{l}\}$ correspond
     to the latest allocation.
11:      Remove b^{inc} from $b_{\bar{c},\bar{l}}$.
12:      PAPR reduction using any of Alg. 1-3.
13:      Evaluate (41), $\forall n$.
14: **end while**

---

according to the bit rate, and obtain the frequency-domain data vectors $\mathbf{d}_c, \forall c$.

Our aim is to maximize the transmitted data rate by alternating between bit and power loading and PAPR minimization. After jointly minimizing the PAPR and the power reserve required by the dummy symbols, we compare the antenna-specific peak power to the peak power constraint P^{tx}, and allocate the difference to the data power budget. The antenna-specific peak power is given by

$$P_n^{\text{peak}} = 10 \log(\|\mathbf{s}_n\|_\infty^2), \forall n. \tag{41}$$

Then, if $P_n^{\text{peak}} < P^{\text{tx}}, \forall n$, we allocate the minimum difference to the data power budget as $P^{\text{data}} := P^{\text{data}} + P^{\text{tx}} - \max_n(P_n^{\text{peak}})$, where $P^{\text{data}} = 10 \log((1/N_c) p^{\text{data}})$ is the average SNR. The bit and power loading and PAPR minimization are repeated using the new data power budget.

Once the antenna-wise peak power constraint is violated, i.e., $P_n^{\text{peak}} > P^{\text{tx}}$, for some n, we backstep individual bit allocations and re-optimize the PAPR, until $P_n^{\text{peak}} < P^{\text{tx}}, \forall n$. This provides the highest possible data rate for Hughes-Hartogs bit and power loading, under a given peak power constraint and QAM-symbol realization. The rate maximization search algorithm is summarized in Algorithm 5.

We acknowledge that the proposed rate maximization search algorithm is not necessarily appealing in practice, as it requires constructing the transmit signal multiple times, while continuously re-optimizing the PAPR. For the purposes of this paper, the algorithm is used to simulate the performance increase that is possible to be obtained using the PAPR reduction framework. In a more practical scenario, the data power budget would be immediately set according to statistics on past PAPR reduction capabilities, such as 99-percentile results for $P^{\text{cf}} + P^{\text{res}}$, without utilizing a search algorithm.

VII. NUMERICAL RESULTS

We investigate the performance of the proposed joint PAPR and power reserve minimization framework using Monte Carlo simulations by first simulating the complementary cumulative distribution function (CCDF) for the different PAPR reduction algorithms and comparing them to the OFDM baseline. Afterward, we investigate the effect that the number of dummy symbols N_D has on the 99-percentile PAPR reduction performance and the achieved rate when using the different algorithms. Finally, we compare the PAPR-rate performance of the algorithms for various peak power constraints.

We compare our proposed algorithms to other TR techniques in the literature [13]–[19], [21], [23], in terms of worst-case PAPR performance. As many of the existing schemes are designed for single-antenna systems and might not readily translate to the multi-antenna case, we will base our comparison on the operating principles, i.e., between the optimization objectives of peak power and strict PAPR minimization. This is done through (16) by setting $\delta = 0$, which corresponds to peak power minimization (i.e., the general literature approach) translated to a general MIMO framework. Furthermore, as [20] uses a linear approximation of the average transmit power, without updating the local point, we use a $j^{\text{SCA}} = 1$ setting of Algorithm 1 with $\delta = 0$ to provide the performance of a similar system. We also compare our algorithms to active constellation extension (solving [12, Eq. (6)]), as it has a similar operating principle as tone reservation.

Straight comparison to other, non-TR based techniques is difficult due to differing objectives and design constraints, such as allowed performance deficiencies. For example, amplitude clipping [5]–[7] can provide significant PAPR reduction if clipping noise and signal distortions can be tolerated. Our approach causes only negligible noise and distortions due to the orthogonal resources of the dummy symbols. Furthermore, methods that trade EVM or BER for PAPR reduction (such as [20], [21]) have poor performance if the BER target is strict. Our proposed approach has a negligible impact on EVM/BER. Thus, setting strict constraints for signal distortions and BER, i.e., mere selection of system parameters, can result in superiority of one method over another. Thus, we have limited our comparisons to methods with similar operating principles.

A. CHANNEL MODELS

For the rate maximization simulations, we generate the channel matrix in two ways. As the iterative block-diagonalization algorithm to find the beamformers is well-justified in scenarios where the channels are correlated, we generate the channel matrix by using a MIMO version of the multipath channel model for uniform linear arrays [36]. The MIMO modification is obtained by utilizing a rank-1 channel matrix for each path instead of a vector [37]. The multipath

channel is given by

$$\mathbf{H}_{c,k} = \frac{\eta_k}{\sqrt{M}} \sum_{m=1}^{M} a_{\mathrm{r}}(\theta_{k,m}) a_{\mathrm{t}}(\theta_{k,m})^{\mathrm{H}} e^{j\phi_{k,m}}, \forall c, \qquad (42)$$

where η_k is the pathloss to user k, M denotes the number of independent (and identically distributed) paths, $\phi_{k,m}$ is the uniformly distributed random phase noise of the path, and the vector $a_{\mathrm{t}}(\theta_{k,m})$ is the array signature vector at the transmitter, with angle of departure $\theta_{k,m}$.

For simplicity, we consider equidistant users with $\eta_k = 1, \forall k$, and assume that the random phase rotation $\phi_{k,m}$ provides enough averaging to use the same angle $\theta_{k,m}$ for both array vectors, without loss of generality. The antenna spacing and carrier wavelength are normalized to 1. The number of multipaths is set to $M = 50$.

For comparison, we also consider uncorrelated channels, where all of the channel coefficients are random, independent and identically distributed as $h_{c,k,i,j} \sim \mathcal{CN}(0, 1), \forall c, k, i, j$.

B. COMMON SYSTEM PARAMETERS

We consider a base station with N_{T} antennas serving $K = 2$ users each equipped with $N_{\mathrm{r}} = 2$ antennas on $N_{\mathrm{C}} = 64$ subcarriers. The users are on a circle around the base station, with a 40 degree angle between the users. Both users have a uniformly distributed angular spread of 20 degrees. The noise power and subcarrier bandwidth are set to $N_0 = 1$ and $B = 1$, respectively. The BER-target for Hughes-Hartogs bit and power loading is set to $\Gamma_{c,l} = 10^{-3}, \forall c, l$, and we consider 4-, 16-, and 64-QAM alphabets, i.e., $b^{\mathrm{inc}} = 2$, for the bit loading.[5]

For the SCA and ADMM algorithms, we set the convergence thresholds $\epsilon^{\mathrm{SCA}} = \epsilon^{\mathrm{peak}} = \epsilon^{\mathrm{rel-peak}} = 10^{-3}$, and the maximum number of iterations as $j^{\mathrm{SCA}} = 30$, $k^{\mathrm{peak}} = k^{\mathrm{rel-peak}} = 1000$, unless otherwise defined. The number of dummy symbols N_{D} varies, but the symbols are placed using a comb-2 structure. We set the ADMM penalty parameter $\rho = 0.01$ for Algorithm 2 and $\sigma = 0.05$ for Algorithm 3, which were found to provide good convergence properties.

C. SIMULATIONS

We investigate the PAPR reduction performance of the proposed joint PAPR and dummy symbol power reserve minimization algorithms. The base station has $N_{\mathrm{T}} = 6$ antennas, with a transmit SNR of $P^{\mathrm{data}} = 20$ dB, and the number of dummy symbols is fixed to $N_{\mathrm{D}} = 64$. We simulate the algorithm performance for minimizing $P^{\mathrm{cf}} + P^{\mathrm{res}}$ over 10000 channel instances, plotted as a complementary cumulative distribution function in Figure 3. To provide sufficiently accurate results, the oversampling factor was set to $\alpha = 4$.

---

[5]Our proposed framework can also straightforwardly handle higher order alphabets, as they only affect the data signal $\mathbf{S}^{\mathrm{data}}$. The use of higher order modulations results in higher non-optimized PAPR, which translates to slightly increased numerical values for the PAPR results. However, the effect would be similar for both our proposed scheme and the comparison cases.
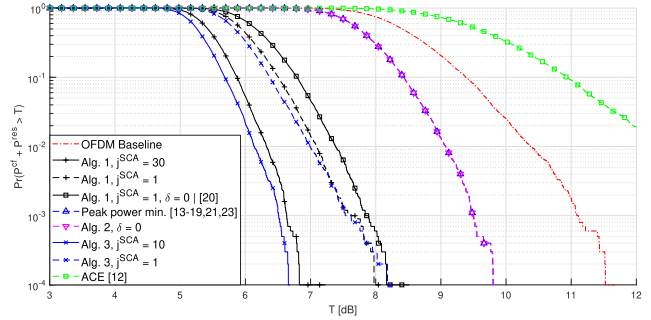

FIGURE 3. Complementary cumulative distribution function for the joint worst-case PAPR and dummy symbol power reserve ($P^{\mathrm{cf}} + P^{\mathrm{res}}$) with different PAPR reduction algorithms.

From Figure 3 we can see the powerful PAPR reduction capabilities of the proposed algorithms, especially Algorithm 3, i.e., the iterative relative peak power minimization. Even for a single SCA iteration, the algorithm performs better than the peak power minimization schemes, with comparable performance to single iteration Algorithm 1. With additional SCA iterations it even beats Algorithm 1, with significantly reduced computational complexity, as discussed in Section IV-D. As the only principal difference between these algorithms is in the handling of the denominator (11b), we conclude that the better performance is caused by a better choice for the fixed SCA local point, i.e., fixing the denominator at each SCA step instead of using a linear approximation. For Algorithm 2, note that we have used $\delta = 0$, which was found to provide the best results in terms of $P^{\mathrm{cf}} + P^{\mathrm{res}}$. Finally, when compared to active constellation extension [12], our proposed scheme performs significantly better in terms of worst-case PAPR. This is due to the peak power minimization objective and our use of high order constellations. Also, due to the power requirement of ACE, the worst-case PAPR can be worse than the OFDM baseline, where no additional power is added to the signal. We investigate the effect of δ next.

It is of interest to see how the performance behaves for $\delta = \{0, 0.1, 0.5, 1, 5, 10\}$, and for comparison, we also provide results for Algorithm 1, where we have modified the objective (15a) to $t/\hat{t} + \delta \|\tilde{\mathbf{m}}\|_2^2 / (p^{\mathrm{data}}\hat{r})$. As Algorithm 3 operates on the same principle as Algorithm 1, the results are assumed to behave similarly, and are thus omitted to avoid cluttering. The results are presented in Figure 4.

It can be seen that when using the strict PAPR expression, i.e., Algorithms 1 and 3, it is best to have equal priority for both the PAPR and dummy symbol power reserve objectives. However, for the peak power minimization problem, the smaller values of δ provide better results. This is intuitive, since minimizing the signal peaks can be done also by minimizing the average power of the signal. Thus, the objective function of minimizing of the dummy symbol power reserve, which adds to the average signal power, is already present in the peak power objective. The same effect does not hold when utilizing the strict PAPR expression, where minimizing PAPR can be done by minimizing the peaks or increasing
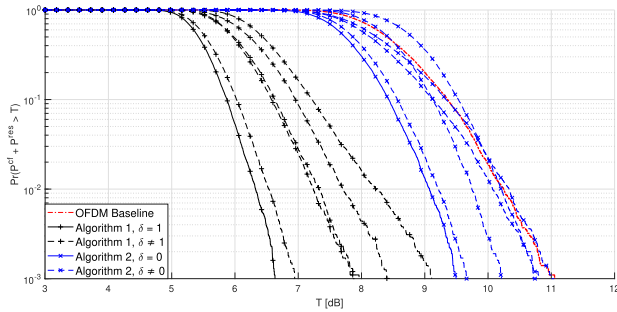
FIGURE 4. Complementary cumulative distribution function for $P^{cf} + P^{res}$, effect of the trade-off parameter δ on the performance of Algorithms 1 and 2.

TABLE 1. Average increase on transmit power required by the dummy symbols, and the 99-percentile PAPR achieved, for different values of δ.

	Algorithm 1		Algorithm 2	
δ	P^{res} [dB]	P^{cf} [dB]	P^{res} [dB]	P^{cf} [dB]
0	2.5395	4.3253	2.5345	6.5291
0.1	2.2324	4.2830	2.4010	6.8162
0.5	1.4236	4.6132	1.9296	7.8643
1	0.9410	5.0754	1.5137	8.7095
5	0.1872	7.0025	0.4761	9.8265
10	0.0768	8.1469	0.2357	9.8222

the average power. It is evident that setting $\delta = 0$, i.e., the approach taken in the literature, provides the best results for Algorithm 2. However, our proposal for Algorithm 2 still provides the possibility of trading between the power peaks and the allocated dummy symbol power (i.e., transmit SNR), which is not considered in the literature, with low computational complexity. This trade off would provide benefits in, for example, low-SNR applications where the signal power peaks are already low enough to not be distorted by PA saturation due to the low average transmit power. Then, a bigger part of the power budget could be allocated for data instead of the dummy symbols and the subsequent loss in peak reduction performance can be tolerated at the PA. Table 1 illustrates the average additional transmit power required for the dummy symbols, and the 99-percentile PAPR that is achieved, for different values of δ. We can observe that the dummy symbol power requirement can be quite significant with lower values of δ, i.e., the region of best performance for Algorithm 2.

Figures 3 and 4 are simulated using $N_D = 64$ dummy symbols. Next, we investigate how the number of dummy symbols affects the algorithm performance. We consider the 99-percentile $P^{cf} + P^{res}$, calculated from results simulated over 1000 channel iterations. We set $\delta = 0$ for Algorithm 2. The oversampling factor was set to $\alpha = 1$ for quicker simulations, as it affects the behavior of all algorithms similarly. Figure 5 presents the results.

It can be seen that increasing the number of dummy symbols has diminishing returns. Algorithm 3 has the best performance overall, offering significant PAPR reduction even at low dummy symbol counts. Algorithm 2 has the worst performance, possibly even lower than the OFDM baseline, at low dummy symbol counts. This is caused by the increase
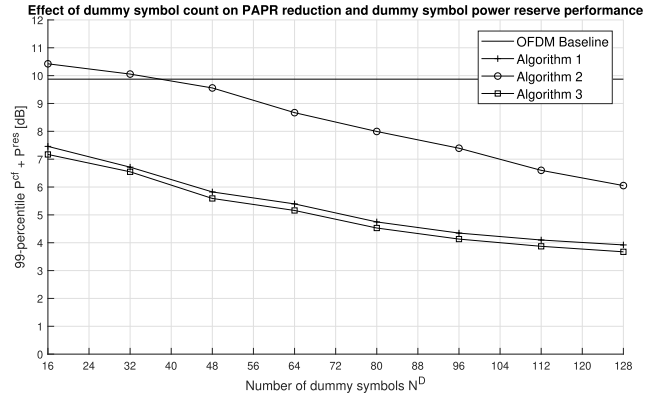

FIGURE 5. PAPR reduction algorithm performance as a function of the number of dummy symbols.

in P^{res}, combined with the low PAPR reduction capability when only few resources are allocated for the algorithm to operate with. As the number of dummy symbols is increased, the algorithms based on the strict PAPR expression begin to saturate, whereas the performance of Algorithm 2 could still benefit from more dummy symbols. It should be noted that $N_D = 128$ corresponds to two full streams allocated for dummy symbols only, which in this setup could be done without affecting the data allocation.

Finally, we investigate the rate performance of the different proposed PAPR and dummy symbol power reserve minimization algorithms using Algorithm 5. The simulation setup has strict constraints due to our use of Hughes-Hartogs bit and power allocation in the signal construction. Thus, we limit our investigation into differences between the algorithms under the same setup instead of simulating a realistic scenario.

We consider a system where the base station has $N_T = 4$ antennas and serves $K = 2$ users with $N_r = 2$ antennas each. Thus, there is a trade-off in space-frequency resources between dummy and data symbols. The results are simulated over 1000 channel realizations, and since we are only interested in the performance differences between the algorithms, the oversampling factor was set to $\alpha = 1$, and we set a peak power constraint $P^{tx} = 20$ dB for Algorithm 5. Here, we also consider the effect of correlated channels to show the benefits of having low-gain streams to exploit.

Figure 6 presents the achieved average rate as a function of the number of dummy symbols, with separate plots for correlated (solid line) and uncorrelated channels (dashed line). We can immediately make two main observations: first, the trade-off in resources between data and dummy symbols, which is evident in the case of uncorrelated channels around the point $N_D = 64$. Allocating more dummy symbols reserves high gain subcarriers from data allocation, and the power increase provided by PAPR minimization is unable to compensate, thus reducing the rate. The second observation is that the rate reduction is not yet present in correlated channels, where the low-gain streams can be allocated for dummy symbols to provide additional power boost, thus increased rate, with minimal penalty.
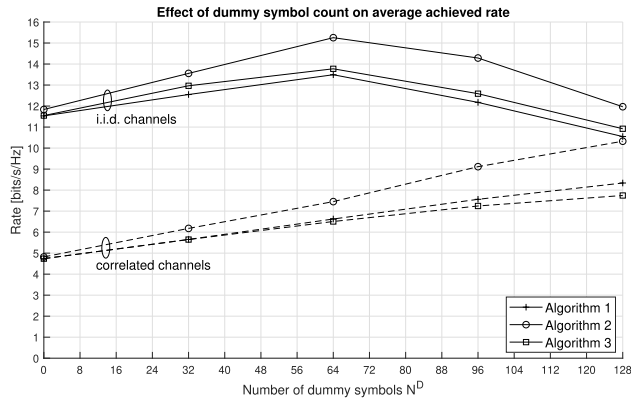
FIGURE 6. Average rate achieved with Algorithm 5 as a function of the number of dummy symbols N_D under correlated and uncorrelated channels for different PAPR minimization algorithms.
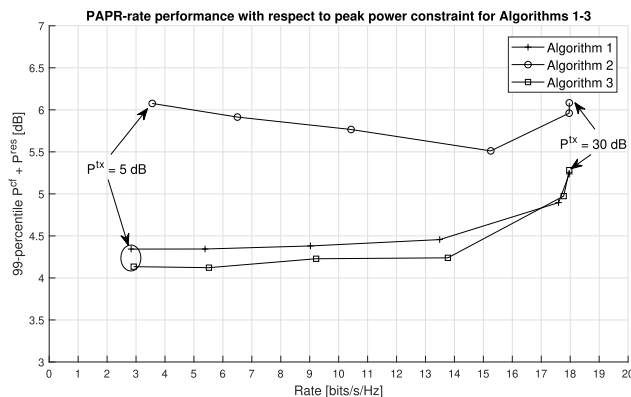

FIGURE 7. PAPR-rate performance of different PAPR minimization algorithms for different peak power constraints.

Comparing the different algorithms, the peak power minimization algorithm (Algorithm 2) performs the best in terms of achieved rate, and the algorithms utilizing the strict PAPR expression (Algorithms 1,3) have a lower but comparable performance. The higher rate performance of Algorithm 2 is explained by the objective of (16), i.e., the focus to minimize the highest signal peak (see (41)), which is the limiting factor when attempting to allocate as much power for data under a given common peak power constraint P^{tx} for all antennas.

Despite the highest performance in terms of rate, the PAPR performance of Algorithm 2 is still worse due to the high PAPR of the antennas with lower average transmit power. To illustrate this, Figure 7 plots the 99-percentile PAPR and achieved rate for different algorithms for $N_D = 64$. In the figure, the points correspond to different peak power constraints, $P^{tx} = \{5, 10, 15, 20, 25, 30\}$ dB, set over all antennas.

It can be seen that Algorithm 2, operating under the peak power minimization principle, has the highest rates (rightward direction), but at the same time, the highest PAPR (upward direction). For a fixed rate, the lowest PAPR is obtained with Algorithm 3. The worse behavior for the peak power minimization at lower peak power constraints is explained by P^{res}, which corresponds to the factor of increase on the allocated data power. Thus, as the data power is very

low, this factor can be very high, even more significant than the reduction in P^{cf}.

VIII. CONCLUSION

We proposed a joint PAPR and dummy symbol power reserve minimization framework to combat the high PAPR of MIMO-OFDM systems. Our proposed method operates on the strict PAPR expression, which also accounts for the PAPR of antennas with lower average transmit power. We were motivated by the fact that the peak power minimization principle, common in the existing literature, does not necessarily translate well to multi-antenna systems as it can ignore the antennas with lower than maximum average transmit power. We also account for the self-power requirement of the scheme, which has not been widely considered.

We derived low-complexity iterative algorithms to provide solutions for the proposed problem, and also proposed an approach to beamforming, where we leveraged knowledge of dummy symbol allocations to provide higher gain data streams at the cost of lower gain streams for dummy symbol allocations. Our simulation results show the benefits of using the strict PAPR expression instead of peak power minimization, in terms of worst-case PAPR over all antennas. We also see that the dummy symbol power reserve should be accounted for if strict PAPR expression is used, however, for peak power minimization it is not necessary. Finally, we see from the results that the proposed PAPR reduction scheme is very beneficial in terms of rate, especially under correlated channels where low gain streams are common.

REFERENCES

[1] 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, Nov. 2013.
[2] *Physical Layer; General Description (Release 16)*, 3GPP, document 38.201, Dec. 2019.
[3] F. H. Raab, P. Asbeck, S. Cripps, P. Kenington, Z. Popovic, N. Pothecary, J. Sevic, and N. Sokal, "Power amplifiers and transmitters for RF and microwave," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 3, pp. 814–826, Mar. 2002.
[4] S. Singh, S. M. Kumar, and H. Mruthyunjaya, "Effect of peak-to-average power ratio reduction on the multicarrier communication system performance parameters," *Int. J. Electr. Comput. Eng.*, vol. 4, no. 12, pp. 779–786, 2009.
[5] X. Li and L. J. Cimini, "Effects of clipping and filtering on the performance of OFDM," in *Proc. IEEE 47th Veh. Technol. Conf.*, vol. 3, May 1997, pp. 1634–1638.
[6] J. Armstrong, "Peak-to-average power reduction for OFDM by repeated clipping and frequency domain filtering," *Electron. Lett.*, vol. 38, no. 5, pp. 246–247, Feb. 2002.
[7] F. Peng and W. Ryan, "On the capacity of clipped OFDM channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2006, pp. 1866–1870.
[8] R. W. Bäuml, R. F. H. Fischer, 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] R. F. H. Fischer and M. Hoch, "Directed selected mapping for peak-to-average power ratio reduction in MIMO OFDM," *Electron. Lett.*, vol. 42, no. 22, pp. 1289–1290, Oct. 2006.
[10] L. J. Cimini and N. R. Sollenberger, "Peak-to-average power ratio reduction of an OFDM signal using partial transmit sequences," *IEEE Commun. Lett.*, vol. 4, no. 3, pp. 86–88, Mar. 2000.
[11] 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.

[12] B. S. Krongold and D. L. Jones, "PAR reduction in OFDM via active constellation extension," *IEEE Trans. Broadcast.*, vol. 49, no. 3, pp. 258–268, Sep. 2003.

[13] B. S. Krongold and D. L. Jones, "An active-set approach for OFDM PAR reduction via tone reservation," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 495–509, Feb. 2004.

[14] J.-C. Chen and C.-P. Li, "Tone reservation using near-optimal peak reduction tone set selection algorithm for PAPR reduction in OFDM systems," *IEEE Signal Process. Lett.*, vol. 17, no. 11, pp. 933–936, Nov. 2010.

[15] L. Wang and C. Tellambura, "Analysis of clipping noise and tone-reservation algorithms for peak reduction in OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 57, no. 3, pp. 1675–1694, May 2008.

[16] H. Li, T. Jiang, and Y. Zhou, "An improved tone reservation scheme with fast convergence for PAPR reduction in OFDM systems," *IEEE Trans. Broadcast.*, vol. 57, no. 4, pp. 902–906, Dec. 2011.

[17] S. Janaaththanan, C. Kasparis, and B. G. Evans, "A gradient based algorithm for PAPR reduction of OFDM using tone reservation technique," in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, May 2008, pp. 2977–2980.

[18] T. Jiang, C. Ni, C. Xu, and Q. Qi, "Curve fitting based tone reservation method with low complexity for PAPR reduction in OFDM systems," *IEEE Commun. Lett.*, vol. 18, no. 5, pp. 805–808, May 2014.

[19] C. Ni, Y. Ma, and T. Jiang, "A novel adaptive tone reservation scheme for PAPR reduction in large-scale multi-user MIMO-OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 5, no. 5, pp. 480–483, Oct. 2016.

[20] A. Aggarwal and T. H. Meng, "Minimizing the peak-to-average power ratio of OFDM signals using convex optimization," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 3099–3110, Aug. 2006.

[21] 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.

[22] V. Tervo, A. Tolli, and T. Matsumoto, "PAPR constrained power allocation for multicarrier transmission in multiuser SIMO communications," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2458–2473, Apr. 2016.

[23] H. Zhang and D. L. Goeckel, "Peak power reduction in closed-loop MIMO-OFDM systems via mode reservation," *IEEE Commun. Lett.*, vol. 11, no. 7, pp. 583–585, Jul. 2007.

[24] A. Arvola, A. Tolli, and D. Gesbert, "A user cooperative beamforming approach to PAPR reduction in MIMO-OFDM uplink," in *Proc. 50th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2016, pp. 691–695.

[25] A. Arvola, A. Tolli, and D. Gesbert, "Rate maximization via PAPR reduction in MIMO-OFDM uplink: A user cooperation approach," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[26] A. Tolli, M. Codreanu, and M. Juntti, "Cooperative MIMO-OFDM cellular system with soft handover between distributed base station antennas," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1428–1440, Apr. 2008.

[27] Q. H. Spencer and A. L. Swindlehurst, "A hybrid approach to spatial multiplexing in multiuser MIMO downlinks," *EURASIP J. Wireless Commun. Netw.*, vol. 2004, no. 2, pp. 236–247, Dec. 2004.

[28] D. Hughes-Hartogs, "Ensemble modem structure for imperfect transmission media," U.S. Patent 4 833 706, May 23, 1989.

[29] Q. Sun, L. Li, A. Tolli, M. Juntti, and J. Mao, "Optimal energy efficient bit and power loading for multicarrier systems," *IEEE Commun. Lett.*, vol. 18, no. 7, pp. 1194–1197, Jul. 2014.

[30] M. Sharif, M. Gharavi-Alkhansari, and B. H. Khalaj, "On the peak-to-average power of OFDM signals based on oversampling," *IEEE Trans. Commun.*, vol. 51, no. 1, pp. 72–78, Jan. 2003.

[31] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Oper. Res.*, vol. 26, no. 4, pp. 681–683, 1978.

[32] S. Boyd, N. Parikh, and E. Chu, *Distributed Optimization and Statistical Learning Via the Alternating Direction Method of Multipliers*. Norwell, MA, USA: Now Publishers, 2011.

[33] L. Li, X. Wang, and G. Wang, "Alternating direction method of multipliers for separable convex optimization of real functions in complex variables," *Math. Problems Eng.*, vol. 2015, pp. 1–14, Nov. 2015.

[34] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[35] B. Farhang-Boroujeny, Q. Spencer, and L. Swindlehurst, "Layering techniques for space-time communication in multi-user networks," in *Proc. IEEE 58th Veh. Technol. Conf.*, vol. 2, Oct. 2003, pp. 1339–1343.

[36] A. F. Molisch, *Wireless Communications*. Hoboken, NJ, USA: Wiley, 2007.

[37] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

ANTTI ARVOLA (Student Member, IEEE) received the M.Sc. (Tech.) degree in electrical engineering from the University of Oulu, in 2015, where he is currently pursuing the Ph.D. degree. He is also a Ph.D. Researcher with the Centre of Wireless Communications (CWC), University of Oulu. His research interests include PAPR reduction in MIMO-OFDM systems, signal processing with low-resolution quantization in sub-THz transmitters and receivers, and resource and interference management problems.

SATYA KRISHNA JOSHI (Member, IEEE) received the M.Eng. degree in telecommunications from the Asian Institute of Technology, Thailand, in 2007, and the Ph.D. degree from the University of Oulu, Oulu, Finland, in 2018. From 2018 to 2019, he was a Postdoctoral Researcher with the University of Oulu. He is currently a Research Specialist at Nokia Bell Labs. His research interests include application of optimization techniques for signal processing and MIMO communications.

ANTTI TÖLLI (Senior Member, IEEE) received the Dr.Sc. (Tech.) degree in electrical engineering from the University of Oulu, Oulu, Finland, in 2008. From 1998 to 2003, he worked at Nokia Networks as a Research Engineer and a Project Manager both in Finland and Spain. In May 2014, he was granted a five year (2014–2019) Academy Research Fellow post by the Academy of Finland. During the academic year 2015–2016, he visited at EURECOM, Sophia Antipolis, France. From August 2018 to June 2019, he was visiting at the University of California at Santa Barbara, Santa Barbara, CA, USA. He is currently a Professor with the Centre for Wireless Communications (CWC), University of Oulu. He has authored numerous papers in peer-reviewed international journals and conferences and several patents all in the area of signal processing and wireless communications. His research interests include radio resource management and transceiver design for broadband wireless communications with a special emphasis on distributed interference management in heterogeneous wireless networks. From 2017 to 2021, he served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING.

DAVID GESBERT (Fellow, IEEE) received the Ph.D. degree from TelecomParis, France, in 1997. From 1997 to 1999, he was with the Information Systems Laboratory, Stanford University. He was the Founding Engineer of Iospan Wireless Inc., a Stanford spin off pioneering MIMO-OFDM (currently Intel). He is currently working as the Director of EURECOM, Sophia Antipolis, France (www.eurecom.fr). Before joining EURECOM, in 2004, he was with the Department of Informatics, University of Oslo, as an Adjunct Professor. He has published about 350 articles and 25 patents, seven of them winning IEEE best paper awards. He is a Board Member of the OpenAirInterface (OAI) Software Alliance. In 2015, he has been awarded an ERC Advanced Grant. In 2020, he was awarded funding by the French Interdisciplinary Institute on Artificial Intelligence for a Chair in the area of AI for the future IoT. In 2021, he received the Grand Prix in Research from IMT-French Academy of Sciences. He has been the Technical Program Co-Chair of ICC2017 and has been named a Thomson-Reuters Highly Cited Researchers in computer science.

• • •