Distortion-Aware Power Allocation for Multi-Stream Distributed Massive MIMO System With Nonlinear Power Amplifier

BIN LIU[ID] (Member, IEEE), AND SOFIE POLLIN[ID] (Senior Member, IEEE)

Department of Electrical Engineering, KU Leuven, 3000 Leuven, Belgium

CORRESPONDING AUTHOR: B. LIU (e-mail: bin.liu@esat.kuleuven.be)

ABSTRACT Distributed MIMO systems leverage multiple access points (APs) distributed across a geographical area to enhance system performance and provide robust connectivity to users by transmitting independent data streams simultaneously, thus improving coverage and capacity. The nonlinear power amplifiers (PAs) employed at the APs can significantly distort the transmitted signal and degrade the spectral efficiency (SE), as PAs are more efficient when operating close to saturation. In this paper, we present an efficient power allocation framework in a multi-stream distributed massive MIMO system with nonlinear PAs, where APs send multiple spatial streams to the user equipment (UE). With the objective of maximizing the SE while adhering to total input power constraints, we first formulate the power allocation problem as nonlinear programming, which is first solved using the multiplier punitive function-based method. Moreover, we focus on the asymptotic analysis of the SE when multi-stream interference among APs is negligible. This can be achieved through ideal spatial multiplexing with interference nulling, assuming a sufficiently large antenna at the UE, or through interference-free frequency multiplexing. We examine the optimality of the water-filling method in power allocation and extend it by considering the PA's nonlinearity. Simulation results show that the proposed power allocation method can maximize the achievable rate while considering the PA nonlinearity, thereby preventing degradation in performance.

INDEX TERMS Distributed massive MIMO, multi-streams, nonlinear power amplifier, power allocation, AP selection.

I. INTRODUCTION

MASSIVE multiple-input multiple-output (MIMO) systems have been proven to drastically improve spectrum efficiency when the number of antennas increases, which is a key enabler for next-generation wireless communication systems [1], [2]. Compared with co-located massive MIMO, distributed massive MIMO has a great potential to improve spectral efficiency, coverage, and energy efficiency by harvesting the macro-diversity gain and reducing access distance [3]. In communication systems, the goal is to maximize the transmission rate under a given power budget. A more complex challenge arises in distributed MIMO systems as the focus shifts to allocating the available power resources effectively among multiple distributed access points (APs) [4].

A. MOTIVATION

Distributed MIMO systems leverage multiple APs geographically distributed across the area to enhance system performance and provide robust connectivity to users. By using the multi-stream transmission approach, each AP transmits independent data streams, exploiting spatial diversity, and improving coverage and capacity [5]. When considering a large number of antennas at the receiving user equipment (UE), it is possible to implement effective spatial multiplexing. When the transmitting distributed APs are separated well in the angular domain, spatial multiplexing can then be achieved even in line-of-sight channels.

However, the ideal spatial multiplexing is not always feasible in practice. If the channel is not orthogonal, or the number of antennas at the receiving UE is low, interference

will always remain between the spatial streams. Note also that ideal multiplexing could also be achieved in alternative implementation scenarios. A large number of antennas at the receiving UE, in which optimal receive combining can resolve the spatial streams. Another approach is a multi-connectivity over the separate band, where the multiple APs use another frequency band, so there is no multi-stream interference to be considered.

On the other hand, scaling up the massive MIMO system requires a large number of radio-frequency (RF) chains, increasing energy consumption [6]. Power amplifiers (PAs) are the most power-hungry RF components. More importantly, the PAs are nonlinear components and generate distortion when working in the saturation regime [7], [8], [9]. A typical solution to avoid PA saturation is using a back-off technique [10], which forces the PA to work below a certain power level. However, the back-off significantly reduces PA efficiency, as PAs are more efficient when close to the saturation point [11], [12]. There is a clear trade-off between PA efficiency and linearity. It is crucial to meticulously design the AP selection and power allocation to provide efficient transmission by considering PA nonlinearity.

B. RELATED WORK
Existing studies have been predominantly focused on AP selection and power allocation under the linear PA assumption. In [13], AP selection was optimized to reduce power consumption by turning off parts of APs, while satisfying the users' demands. In [14], the power allocation in the distributed MIMO system was cast as a max-min optimization problem, and an iterative max-min power optimization algorithm was developed to ensure downlink fairness. In [15], the joint energy efficiency optimization, including power allocation, user grouping, and antenna activation, was formulated as a mixed-integer nonlinear programming problem. The iterative successive convex approximation technique was applied to solve the problem sub-optimally. In [16], the power allocation under a sum power constraint was structured as a power control game. An iterative search algorithm was developed to maximize the sum rate. The joint power allocation and beamforming were formulated as a non-convex constrained optimization problem in [17], sub-optimally solved using the block coordinate descent method.

However, these works cannot be directly applied to the AP selection and power allocation problem studied in this paper. Specifically, these works optimized the AP selection and power allocation based on the linear PA assumption, aiming at maximizing the signal-to-noise ratio (SNR) or its related function, e.g., sum rate and energy efficiency. In practice, the amplification of PA is always *nonlinear*. When working in a saturation regime, PAs generate distortion – which is relevant to the input signal – and degrade system performance. Measurement-based study with the massive MIMO testbed also verified that nonlinear distortion appears and can coherently combine in space [18], [19].

Recent work [20] proposed a linear precoder considering PA nonlinearity to reduce the nonlinear distortion in a centralized MIMO system. An iterative precoding algorithm was proposed to maximize the signal-to-interference-noise and distortion ratio (SINDR). However, the iterative solver in [20] is computationally expensive, despite focusing on a centralized MIMO system with a single array. A linear precoder was proposed in [21] that allows the PA to work close to saturation while canceling the coherent combining of the third-order nonlinear distortion at the user location.

C. CONTRIBUTION
This paper investigates the AP selection and power allocation in the multi-stream distributed massive MIMO system with consideration of PA nonlinearity. As the distortion has its most significant impact on the massive MIMO systems under line-of-sight (LoS) channel conditions [21], [22]. In our analysis, we adopt this single-user LoS channel scenario, which offers tractable analytical models that yield valuable insights while representing a worst-case yet realistic situation. The single-user scenario is also extended to a multi-user optimization problem.

Our main contributions are summarized below.

- We formulate the power allocation problem for the multi-stream distributed massive MIMO system with nonlinear PAs as nonlinear programming. An iterative power allocation algorithm was developed based multiplier punitive function method to maximize the SE by optimizing the power allocation of APs.
- We analyze the asymptotic SE performance considering the nonlinearity of the PAs in the multi-stream distributed massive MIMO system.
- We further extend and develop the distortion-aware water-filling power allocation algorithm, solving the power allocation problem with lower complexity.

Simulation results show that the proposed power allocation can maximize the SE with awareness of the PA nonlinearity and prevent it from degrading performance, and the proposed distortion-aware water-filling power allocation algorithm can achieve near-optimal results with lower complexity.

The remainder of this paper is organized as follows. Section II introduces the system model of the multi-stream distributed massive MIMO. The multiplier punitive function-based power allocation algorithm is developed to maximize SE in Section III. The distortion-aware water-filling power allocation algorithm is presented in Section IV. Simulation results are provided in Section V, followed by conclusions in Section VI.

Notations: The following notations are used throughout this paper. a and A stand for a column vector and a matrix, respectively; $A_{i,j}$ is the entry in the ith row and jth column of A. The conjugate, transpose, conjugate transpose, determinant, and trace of A are represented by A^*, A^T, A^H, $\det(\cdot)$ and $\mathrm{tr}(\cdot)$, respectively; $|\cdot|$ and $\|\cdot\|$ denote modulus and Frobenius norm, respectively; $\mathrm{diag}(\cdot)$
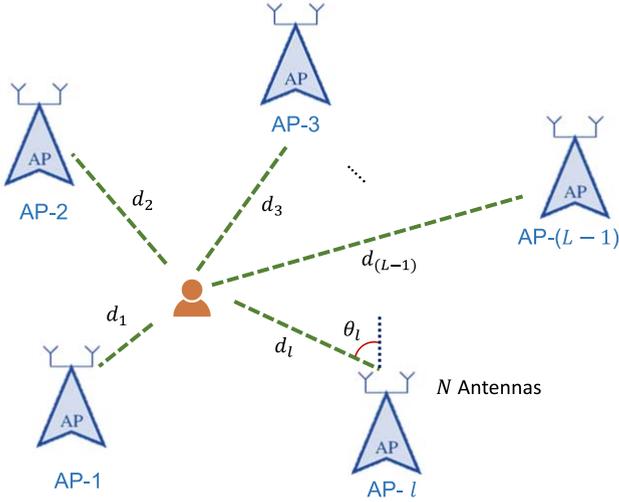
FIGURE 1. An illustration of the considered distributed massive MIMO system.

TABLE 1. Abbreviations.

Abbreviation	Definition
AoA	angle of arrival
AoD	angle of departure
CPU	central processing unit
CSI	channel state information
AP	access point
LoS	line-of-sight
MIMO	multiple input multiple output
MMSE	minimum mean square error
MRC	maximum ratio combining
MRT	maximum ratio transmission
OFDM	orthogonal frequency-division multiplexing
PA	power amplifier
UE	user equipment
RF	radio frequency
SE	spectral efficiency
SNR	signal to noise ratio
SINDR	signal-to-interference-noise and distortion ratio
ULA	uniform linear array
WF	water-filling
ZF	zero-forcing

denotes the diagonalization operation; The expectation of a complex variable is noted $\mathbb{E}[\,\cdot\,]$. $\mathbb{R}$ and $\mathbb{C}$ represents real and complex number, respectively. $\mathcal{CN}(\mu, \boldsymbol{\Sigma})$ denotes the complex Gaussian distribution with mean μ and variance $\boldsymbol{\Sigma}$. Table 1 lists the abbreviation used in this paper.

II. MULTI-STREAM DISTRIBUTED MASSIVE MIMO SYSTEM
A. SYSTEM MODEL
Fig. 1 illustrates the considered multi-stream distributed massive MIMO system, where L APs are deployed within the considered region, Each AP, equipped with N antennas, is connected to a central processing unit (CPU) through a fronthaul link. This CPU centrally processes the data from

all APs, similar to the cell-free networking concept proposed before [23]. In the considered multi-stream distributed massive MIMO system, each AP implements effective spatial multiplexing, transmitting separate streams in parallel to the UE to maximize the SE. The referred distributed MIMO system is characterized by multiple antennas at each AP (large N), while it is often assumed that the number of antenna elements per AP is small or even only one (small N) in cell-free systems [24]. The UE has N_r antennas. The received signal vector at the UE can be expressed as follows

$$\mathbf{r}_k = \sum_{l=1}^{L} \mathbf{H}_{l,k} f(\mathbf{s}_l) + \mathbf{n}, \tag{1}$$

where $\mathbf{s}_l = [s_{l,1}, \ldots, s_{l,N}]^{\mathrm{T}} \in \mathbb{C}^N$ denotes the transmit signal vector of l-th AP. Assume $\mathbf{s}_l$ is a complex Gaussian distributed signal with zero mean and covariance matrix $\mathbf{C}_{\mathbf{s}_l}$. The function $f(\cdot):\mathbb{C} \to \mathbb{C}$ describes the input-output relationship in the amplification process. $\mathbf{n} \sim \mathcal{CN}(0, \sigma_0^2 \mathbf{I}_N)$ denotes the additive circularly symmetric complex Gaussian noise. $\mathbf{H}_l \in \mathbb{C}^{N_r \times N}$ represents the channel between the UE and the l-th AP.

This paper presents an efficient power allocation framework in a multi-stream distributed massive MIMO system with nonlinear PAs. As shown in [25], with the increasing number of users and multi-path components, the transmit signals become 'isotropically' radiated. In this sense, the distortion has its most significant impact on the massive MIMO systems for single-user under line-of-sight (LoS) channel conditions [21], [22]. In our analysis, we adopt this single-user LoS channel scenario, which offers tractable analytical models that yield valuable insights while representing a worst-case situation.

The LoS channel between the UE and the l-th AP is modeled as

$$\mathbf{H}_l = \sqrt{NN_r}\, \psi_l\, \mathbf{a}_r(\theta_l) \mathbf{a}_t^{\mathrm{H}}(\phi_l) \in \mathbb{C}^{N_r \times N}. \tag{2}$$

Here $\psi_l = \sqrt{d_l^{-\alpha}}$ represents the distance-dependent path loss, where d_l stands for the distance between the UE and the l-th AP, and α is the path loss exponent. θ_l and ϕ_l represent the angles of arrival (AoA) and angle of departure (AoD) of the signal from the l-th AP, respectively. We assume that both the transmitter and the receiver are equipped with uniform linear arrays (ULAs) with array responses, as given by

$$\mathbf{a}_r(\theta_l) = \sqrt{\frac{1}{N_r}}\left[1, e^{\frac{-j2\pi\Delta\sin(\theta_l)}{\lambda}}, \ldots, e^{\frac{-j2\pi\Delta(N_r-1)\sin(\theta_l)}{\lambda}}\right]^T, \tag{3}$$

$$\mathbf{a}_t(\phi_l) = \sqrt{\frac{1}{N}}\left[1, e^{\frac{-j2\pi\Delta\sin(\phi_l)}{\lambda}}, \ldots, e^{\frac{-j2\pi\Delta(N-1)\sin(\phi_l)}{\lambda}}\right]^T, \tag{4}$$

where λ is the signal wavelength. Δ represents the antenna spacing of the receiver and transmitter. The number of antennas at the receiver N_r needs to be large enough to eliminate multi-stream interference when relying on spatial filtering. Perfect channel state information is assumed for assessing the impact of non-linearities on the distributed

MIMO system performance. This simplifying assumption allows focusing on the impact of PA nonlinearity on the distributed MIMO system performance, as a common approach in many studies on the massive MIMO [21], [25], [26], [27], [28].

B. NONLINEAR POWER AMPLIFICATION

This paper employs a memoryless polynomial model of order $2M+1$ to characterize the nonlinear characteristics of PAs at the transmitter [29]. We assume the PAs in different antenna branches adhere to the same input-output relationship. The equivalent baseband output signal at the n-th PA of the l-th AP is then represented as follows:

$$x_{n,l} = f(s_{l,n}) \triangleq \sum_{m=0}^{M} \beta_{2m+1} |s_{l,n}|^{2m} s_{l,n}, \tag{5}$$

where β_{2m+1} are complex values in general, encapsulating both amplitude-to-amplitude modulation (AM/AM) and amplitude-to-phase modulation (AM/PM) distortions. The instantaneous amplitude gain of the PA is then given by

$$g_{l,n} \triangleq \frac{x_{l,n}}{s_{l,n}} = \sum_{m=0}^{M} \beta_{2m+1} |s_{l,n}|^{2m}. \tag{6}$$

The amplified signal vector at the l-th AP can be written as

$$\mathbf{x}_l = \mathbf{G}_l \mathbf{s}_l, \tag{7}$$

where the diagonal matrix $\mathbf{G}_l = \text{diag}(g_{l,1}, \ldots, g_{l,N})$ represents the instantaneous amplitude gain of each PA at the l-th AP.

By applying Bussgang's theorem [30], the PAs output signal can be represented as a linearly amplified version of the input signal $\mathbf{s}_l$ perturbed by the nonlinear distortion $\mathbf{e}_l$, as given by

$$\mathbf{x}_l = \overline{\mathbf{G}}_l \mathbf{s}_l + \mathbf{e}_l \in \mathbb{C}^N, \tag{8}$$

where $\overline{\mathbf{G}}_l$ denotes the average linear gain of the PA's amplification process. $\mathbf{e}_l = [e_{l,1}, \ldots, e_{l,N}]$ collects the nonlinear distortion across all PAs at the l-th AP. According to Bussgang's theorem [31], the distortion generated is uncorrelated with the input signal of the PA, i.e., $\mathbb{E}\{s_{l,n}^* e_{l,n}\} = 0$. We assume that the antenna branches are perfectly isolated from each other, and coupling is negligible. Hence, $\overline{\mathbf{G}}_l$ is a diagonal matrix as [29],

$$\overline{\mathbf{G}}_l = \text{diag}(\overline{g}(P_1), \ldots, \overline{g}(P_{l,n})), \tag{9}$$

where $\overline{g}(P_{l,n}) = \sum_{m=0}^{M} \beta_{2m+1}(m+1)! P_{l,n}^m$ represents the linear amplification gain. $P_{l,n} = \mathbb{E}\{|s_{l,n}|^2\} = [\mathbf{C}_{\mathbf{s}_l}]_{n,n}$ is the average power of the input signal at the n-th PA of the l-th AP. As described in [29], the sketch of the proof can be found in the Appendix. By comparing this with the relationship $\mathbf{x}_l = \mathbf{G}_l \mathbf{s}_l$ as given (6), we can derive an expression for the nonlinear distortion as

$$\mathbf{e}_l = (\mathbf{G}_l - \overline{\mathbf{G}}_l) \mathbf{s}_l, \tag{10}$$

where the nonlinear distortion vector $\mathbf{e}_l$ is a zero-mean complex random vector with the covariance matrix, as given by [30]

$$\mathbf{C}_{\mathbf{e}_l} = \sum_{m=1}^{M} \mathbf{\Gamma}_m \mathbf{C}_{\mathbf{s}_l} \odot |\mathbf{C}_{\mathbf{s}_l}|^{2m} \mathbf{\Gamma}_m^H, \tag{11}$$

where $\mathbf{\Gamma}_m = \text{diag}(\gamma_m(P_{l,1}), \ldots, \gamma_m(P_{l,N}))$ and

$$\gamma_m(P_{l,n}) = \sqrt{\frac{1}{m+1} \sum_{q=m}^{M} \beta_{2m+1} \binom{q}{m}(q+1)! P_{l,n}^{(q-m)}}. \tag{12}$$

III. SUM ACHIEVABLE RATE OPTIMIZATION

In this section, we study the optimal power allocation among APs to maximize spectral efficiency in the considered distributed MIMO spatial multiplexing system. In this system, each multi-antenna AP transmits an individual spatial stream to a multi-antenna UE. We assume each AP applies the maximum ratio transmission (MRT), while at the UE side, combining is performed using maximum ratio combining (MRC). The objective is to maximize spectral efficiency by optimizing the power allocation among APs under the total power budget. ZF and MMSE rely on accurate and up-to-date CSI for effective interference cancellation and noise suppression. Ensuring that each distributed node has accurate information about the channel state of other nodes is non-trivial. Thanks to the simplicity and robustness of MRT [32], this paper investigates the power allocation in the multi-stream distributed massive MIMO system under the given MRT scheme.

By applying MRT, the transmit signal vector of l-th AP is specified by $\mathbf{s}_l = \mathbf{a}_t(\phi_l) u_l$, where u_l denotes the transmitted symbol of the l-th AP. In particular, the transmit signal at the n-th AP is given by $s_{l,n} = a_{l,n}^* u_l$, where $a_{l,n}$ is the precoding weight at the n-th antenna when MRT is applied at the l-th AP. The transmit power per antenna of the l-th AP is then given by

$$\rho_l \triangleq \mathbb{E}\{|s_{l,n}|^2\} = P_l/N, \ \forall n \tag{13}$$

When the number of receive antennas at UE N_r is large, the arrival angles from different APs can be well-separated. After performing MRC at the UE, the received signal from the l-th AP can be expressed as

$$\hat{y}_{l,k} = \mathbf{a}_r^H(\theta_l) \mathbf{r}$$
$$= \underbrace{\overline{g}_l u_{l,k}}_{\text{amplified signal}} + \underbrace{e_l}_{\text{distortion}}$$
$$+ \underbrace{\sum_{l' \neq l} v_{l,l'}(\overline{g}_{l'} u_{l'})}_{\text{multi-stream interference}} + \underbrace{\mathbf{a}_r^H(\theta_l) \mathbf{n}}_{\text{effective noise}}, \tag{14}$$

which is obtained by leveraging Bussgang's theorem, and the received signal from the l-th AP can be represented as a linearly amplified version of the signal symbol u_l contaminated with the nonlinear distortion e_l and interference term. $v_{l,l'} = \mathbf{a}_r^H(\theta_l) \mathbf{a}_r(\theta_{l'})$ denotes the channel correlation,

which causes multi-stream interference due to the imperfect channel orthogonality. $\bar{g}_l$ denotes the average linear amplitude amplification gain of the l-th AP as defined in (9) and is updated as

$$\bar{g}(\rho_l) = \sqrt{\frac{NN_r}{d_l^\alpha}} \sum_{m=0}^{M} \beta_{2m+1}(m+1)! \rho_l^m, \qquad (15)$$

and e_l is the nonlinear distortion generated at the l-th AP, which is a zero-mean complex random variable with variance

$$\bar{g}_e(\rho_l) \triangleq \mathbb{E}(|e_l|^2) = \frac{NN_r}{d_l^\alpha} \sum_{m=1}^{M} |\gamma_m(\rho_l)|^2 \rho_l^{2m+1}. \qquad (16)$$

When Gaussian symbols are transmitted, the SE of the l-th AP can be expressed as

$$R_l(\rho_l) = \log_2 \left(1 + \frac{\rho_l |\bar{g}(\rho_l)|^2}{\bar{g}_e(\rho_l) + \sum_{l' \neq l} |v_{l,l'}|^2 \bar{g}(\rho_{l'})|^2 \rho_{l'} + \sigma_0^2} \right). \qquad (17)$$

Gaussian signals are frequently employed as a modeling tool in communication theory and signal processing literature [33], [34]. It allows for analytical solutions and closed-form expressions that facilitate a clear understanding of system behavior. While we acknowledge that real-world signals may exhibit more complex statistical properties, the use of Gaussian signals provides a baseline for theoretical analysis and comparison with existing literature [35], [36], [37].

OFDM symbols can't be perfectly modeled as Gaussian signals in general due to the fact that the combined effect of these individual subcarriers in creating the overall OFDM symbol doesn't always precisely follow a Gaussian distribution. However, the complex envelope of a bandlimited OFDM signal tends toward exhibiting characteristics similar to a Gaussian random process under certain conditions, for instance, when the signal bandwidth is sufficiently large, and the subcarrier spacing is small enough to meet certain criteria. In such scenarios, the central limit theorem might apply, leading the complex envelope of the OFDM signal to exhibit properties akin to a Gaussian random process [38].

On the other hand, the distorted signal is not Gaussian in general due to the nonlinearity of the PA, and obtaining the spectral efficiency expression is not straightforward in our study [20]. A common approach to evaluate the spectral efficiency using the so-called "auxiliary-channel lower bound" to derive a lower bound on the spectral efficiency expression as in [39]. By starting with a well-understood and analytically tractable signal model, we can clearly quantify the impact of other factors, such as PA nonlinearity, on performance.

In the considered multi-stream distributed massive MIMO system, each AP sends data streams independently in a multiplexing approach. Under the total power budget, the

---

Algorithm 1 Multiplier Punitive Function-Based Power Allocation Algorithm

---

1: Input: total input power constraint P_{total}, the channel $\mathbf{H} = [\mathbf{H}_1, \cdots, \mathbf{H}_L]$. Max iteration K, error tolerance ϵ, step size υ
2: Initialize: iteration $k \leftarrow 0$, $\zeta^{(0)} \leftarrow 0$, $\boldsymbol{\rho}^{(0)} \leftarrow \mathbf{0}$.
3: $\quad R_{\text{sum}}^{(0)} \leftarrow \sum_{l=1}^{L} R_l(\boldsymbol{\rho}^{(0)})$
4: **for** $k = 1 \cdots K$ **do**
5: $\quad$ Calculate $\hat{\boldsymbol{\rho}}$ that maximizes $F(\boldsymbol{\rho}_k, \zeta_k)$ by using the gradient descent method;
6: $\quad R_{\text{sum}}^{(\text{tem})} \leftarrow \sum_{l=1}^{L} R_l(\hat{\boldsymbol{\rho}})$
7: $\quad$ **if** $R_{\text{sum}}^{(\text{tem})} > R_{\text{sum}}^{(k-1)}$ **then**
8: $\quad\quad \boldsymbol{\rho}^{(k)} \leftarrow \hat{\boldsymbol{\rho}}, \zeta^{(k)} \leftarrow \zeta^{(k-1)}$
9: $\quad$ **else**
10: $\quad\quad \boldsymbol{\rho}^{(k)} \leftarrow \boldsymbol{\rho}^{(k-1)}, \zeta^{(k)} \leftarrow \frac{1}{2}\zeta^{(k-1)}$
11: $\quad$ **end if**
12: **end for**
13: $\boldsymbol{\rho}^{\text{opt}} \leftarrow \boldsymbol{\rho}^{(K)}$

---

optimization problem is formulated to maximize the SE by allocating the input power of each AP, as given by

$$\max_{\boldsymbol{\rho}} \quad \sum_{l=1}^{L} R_l \qquad (18a)$$

$$\text{s.t.} \quad \sum_{l=1}^{L} \rho_l \leq P_{\text{total}}/N, \qquad (18b)$$

where $\boldsymbol{\rho} = [\rho_1, \ldots, \rho_L]$ is the power allocation vector for each AP. In particular, $\rho_l = 0$ denotes the l-th AP is not chosen. The total transmit power over all selected APs is constrained by the total power budget P_{total}.

With the multiplier punitive function-based method [40], [41], the optimization problem (18) can be solved by constructing the augmented Lagrange function as

$$F(\boldsymbol{\rho}, \zeta) = \sum_{l=1}^{L} R_l - \zeta(P_{\text{total}}/N - \sum_{l=1}^{L} \rho_l), \qquad (19)$$

where ζ is the Lagrange multiplier. Algorithm 1 summarizes the procedure of the multiplier punitive function-based power allocation algorithm. The multiplier punitive function-based method (Algorithm 1) employs the multiplier as a factor to adjust or penalize, which ensures the method operates within the specified power boundaries.

Complexity: The proposed Algorithm 1 first uses the gradient descent method to obtain the optimal solution of the augmented Lagrange function. The optimization problem (18) has L optimization variables. Denote the number of iterations of the gradient descent method as I_{GD}. and then the complexity of the gradient descent algorithm is $\mathcal{O}(I_{GD}L)$. Let I_{MP} denote the number of iterations of the multiplier punitive function-based method, and thus the complexity of the proposed Algorithm 1 is $\mathcal{O}(I_{MP}I_{GD}L)$.

The proposed framework can be extended to the multiuser case using a simplified and sub-optimal approach, where we

assume that each AP divides the total power equally among all users. Consequently, the power allocation is simplified to optimize P_l/K. While our approach offers a straightforward solution, we acknowledge the importance of considering all resources, including user-AP association, beamforming, and power allocation for each user, as well as fairness policies for multi-user scenarios, in future research.

In particular, by applying MRT and allocating power equally to all K users, i.e., P_l/K. The transmit signal vector of the l-th AP is then specified by $\mathbf{s}_l = \mathbf{A}_t(\boldsymbol{\theta}_l)\mathbf{u}_l$, where $\mathbf{u}_l = [u_{l,1}, \ldots, u_{l,K}] \in \mathbb{C}^K$ denotes the transmitted symbols for K users at the l-th AP. $\mathbb{E}\{|\mathbf{u}_{l,k}|^2\} = P_l/K$. $\mathbf{A}_t = [\mathbf{a}_t(\theta_{l,1}), \ldots, \mathbf{a}_t(\theta_{l,K})] \in \mathbb{C}^{N \times K}$, where $\boldsymbol{\theta}_l = [\theta_{l,1}, \ldots, \theta_{l,K}] \in \mathbb{C}^K$ represent the angle of departure (AoD) of the signal from the l-th AP to each user.

After performing MRC at the k-th UE, the received signal from the l-th AP can be expressed as

$$
\begin{aligned}
\hat{y}_{l,k} &= \mathbf{a}_r^{\mathrm{H}}(\theta_{l,k})\mathbf{r} \\
&= \underbrace{\bar{g}_l u_{l,k}}_{\text{amplified signal}} + \underbrace{e_{l,k}}_{\text{distortion}} + \underbrace{\sum_{k' \neq k} \chi_{k,k',l}(\bar{g}_l u_{l,k'})}_{\text{multi-user interference}} \\
&\quad + \underbrace{\sum_{l' \neq l} v_{l,l',k}(\bar{g}_{l'} u_{l',k})}_{\text{multi-stream interference}} + \underbrace{\mathbf{a}_r^{\mathrm{H}}(\theta_{l,k})\mathbf{n}}_{\text{effective noise}},
\end{aligned}
\tag{20}
$$

which is obtained by leveraging Bussgang's theorem, and the received signal from the l-th AP can be represented as a linearly amplified version of the symbol vector $\mathbf{u}_l$ contaminated with the nonlinear distortion e_l and interference term. $v_{l,l',k} = \mathbf{a}_r^{\mathrm{H}}(\theta_{l,k})\mathbf{a}_r(\theta_{l',k})$ denotes the channel correlation, which causes multi-stream interference due to the imperfect channel orthogonality. $\chi_{k,k',l} = \mathbf{a}_t^{\mathrm{H}}(\theta_{l,k})\mathbf{a}_t(\theta_{l,k'})$ denotes the channel correlation between different UEs at the AP l. When i.i.d. Gaussian symbols are transmitted, the SE of the k-th user from l-th AP can be expressed as (21), shown at the bottom of the page. The transmit power per antenna of the l-th AP is then given by $\rho_l \triangleq \mathbb{E}\{\|s_{l,n}\|^2\} = P_l/N$, $\forall n$. It should be noted that $e_{l,k}$ is not Gaussian in general due to the nonlinearity of the PA, and obtaining the spectral efficiency expression is not straightforward [19]. A common approach to evaluate the spectral efficiency using the so-called "auxiliary-channel lower bound" to derive a lower bound on the spectral efficiency expression as [39].

In the considered multi-user multi-stream distributed massive MIMO system, each AP sends data streams to each UE independently in a simple multiplexing approach consisting of maximum ratio transmission and reception. Using this approach, the power allocation problem is significantly simplified and can be decoupled from the beamforming

design. The power control then allocates power while considering both interference between users and streams, and distortion by the PA. Under the total power budget, the optimization problem is formulated to maximize the SE by allocating the input power of each AP, as given by

$$
\max_{\boldsymbol{\rho}} \quad \sum_{l=1}^{L} \sum_{k=1}^{K} R_{l,k}(\rho_l) \tag{22a}
$$

$$
\text{s.t.} \quad \sum_{l=1}^{L} \rho_l \leq P_{\text{total}}/N, \tag{22b}
$$

where $R_l = \sum_{k=1}^{K} R_{l,k}(\rho_l)$. $\boldsymbol{\rho} = [\rho_1, \ldots, \rho_L]$ is the power allocation vector for each AP. In particular, $\rho_l = 0$ denotes the l-th AP is not chosen. The total transmit power over all selected APs is constrained by the total power budget P_{total}. Then, the multiplier punitive function-based method can be similarly applied to obtain the optimal power allocation over each array, as given in (19). The proposed approach optimizes the power allocation over the different APs for multiple users, assuming the power per AP is equally divided between all the users K.

IV. DISTORTION-AWARE WATER-FILLING POWER ALLOCATION

In this section, we focus our analysis on the case when inter-stream interference is negligible and will show by simulation that this analysis gives a good bound on the results. Specifically, we assume that the antenna number of the UE, i.e., N_r, is large and that the angles of arrival from each AP are well separated. The streams from different APs distributed in different directions to the UE can be separated using a spatial filter. We can then neglect such multi-stream interference and focus on only the distortion from the amplification process in the multi-AP power allocation.

When multi-stream interference is negligible, the SE achieved from the l-AP is given by

$$
\bar{R}_l(\rho_l) = \log_2\left(1 + \frac{\rho_l|\bar{g}(\rho_l)|^2}{\bar{g}_e(\rho_l) + + \sigma_0^2}\right). \tag{23}
$$

In the following, we develop a distortion-aware water-filling power allocation algorithm, which can obtain approximately optimal power allocation at lower complexity. The key idea is that we extend the water-filling method by incorporating PA nonlinearity into the power allocation. Specifically, we show that the SE increases with the input power before the PA enters the saturation region, though it is a non-increasing (non-monotonic) function of the input power in general. The well-known water-filling power allocation method can guarantee its optimality in capacity

$$
R_{l,k}(\rho_l) = \log_2\left(1 + \frac{\rho_l|\bar{g}(\rho_l)|^2}{\bar{g}_e(\rho_l) + \sum_{k' \neq k}|\chi_{k,k',l}\bar{g}(\rho_l)|^2\rho_l + \sum_{l' \neq l}|v_{l,l',k}\bar{g}(\rho_{l'})|^2\rho_{l'} + \sigma_0^2}\right). \tag{21}
$$

during PA works in its linear region [42]. We also show that there is an optimal average input power for PA of l-th AP, denoted by ρ_l^*, when PA nonlinearity is taken into account. The optimal value of input power for PA ρ_l^* is independent of total power constraints and has a closed-form expression. Therefore, the optimal power allocation takes the minimum power allocation from the water-filling method and its local optimal value of average input power for PA of l-th AP ρ_l^*.

First, we consider power allocation without the total power constraints. The sufficient condition is when the maximum SE of the l-th AP is given by

$$\frac{\partial \overline{R}_l}{\partial \rho_l} = \frac{1}{D_{\text{SINR}}^2} \left[\left(|\overline{g}(\rho_l)|^2 + 2\rho_l \overline{g}(\rho_l) \frac{\partial \overline{g}(\rho_l)}{\partial \rho_l} \right) D_{\text{SINR}} \right.$$
$$\left. - \frac{\partial g_e(\rho_l)}{\partial \rho_l} S_{\text{SINR}} \right] = 0, \quad (24)$$

where $S_{\text{SINR}} = \rho_l |\overline{g}(\rho_l)|^2$ and $D_{\text{SINR}} = \overline{g}_e(\rho_l) + \sigma_0^2$ denote the numerator and denominator of the signal-to-interference-noise and distortion ratio (SINDR) in (23), respectively. The local optimal average input power for PA of l-th AP, i.e., ρ_l^*, can be obtained by solving the first-order condition in (24).

Different from the observations in most distributed MIMO works under linear PA assumption, the increase of transmit power cannot always lead to a higher SE due to the nonlinear amplification, which shows the SE is a non-monotonically increasing function of the input power. In the following, we provide an asymptotic analysis of the SE for a massive MIMO system with nonlinear PAs and characterize the floor region of SE performance, as stated in Theorem 1.

Theorem 1: For the AP l with a finite number of antennas, when the input power $\rho_l \to \infty$, the SE approaches to

$$\lim_{\rho_l \to \infty} \overline{R}_l = \log_2 \left(1 + \frac{\left[\beta_{2M+1}(M+1)! \right]^2}{\sum_{m=1}^{M} \frac{1}{m+1} \left[\beta_{2m+1} \binom{M}{m}(M+1)! \right]^2} \right), \quad (25)$$

and when $\rho_l \to 0$, the SE approaches to

$$\lim_{\rho_l \to 0} \overline{R}_l = \beta_1^2 \frac{\rho_l N N_r}{\sigma_0^2 \ln 2}. \quad (26)$$

Proof: When the input power goes into infinite, i.e., $\rho_l \to \infty$, we have

$$\lim_{\rho_l \to \infty} S_{\text{SINR}} = \left[\beta_{2M+1}(M+1)! \right]^2 \rho_l^{2M+1}, \quad (27)$$

$$\lim_{\rho_l \to \infty} D_{\text{SINR}} = \sum_{m=1}^{M} \frac{1}{m+1} \left[\beta_{2m+1} \binom{M}{m}(M+1)! \right]^2 \rho_l^{2M+1}. \quad (28)$$

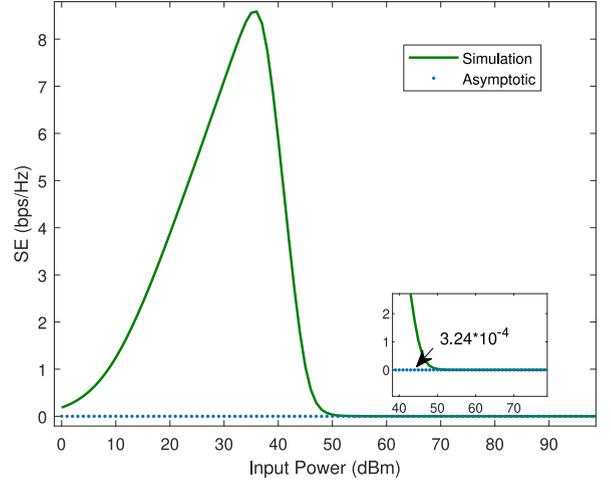With asymptotic linear and nonlinear power gains in (27) and (28), the SE of the l-th AP in (23) approaches to

FIGURE 2. Asymptotic SE when input power increases to infinite.

$$\lim_{\rho_l \to \infty} \overline{R}_l$$
$$= \log_2 \left(1 + \frac{\left[\beta_{2M+1}(M+1)! \right]^2}{\sum_{m=1}^{M} \frac{1}{m+1} \left[\beta_{2m+1} \binom{M}{m}(M+1)! \right]^2} \right). \quad (29)$$

When in the low SNR regime, the input power approaches 0, i.e., $\rho_l \to 0$, the SE $\overline{R}_l$ approaches to

$$\lim_{\rho_l \to 0} \overline{R}_l$$
$$= \lim_{\rho_l \to 0} \log_2 \left(1 + \frac{N N_r \left[\sum_{m=0}^{M} \beta_{2m+1}(m+1)! \right]^2 \rho_l^{2m+1}}{\sum_{m=1}^{M} |\gamma_m(\rho_l)|^2 \rho_l^{2m+1} N N_r + (d_l)^\alpha \sigma_0^2} \right)$$
$$= \beta_1^2 \frac{\rho_l N N_r}{\sigma_0^2 \ln 2}, \quad (30)$$

which is achieved by using $\ln(1 + x) \sim x$. ∎

The asymptotic SE $\overline{R}_l$ when input power ρ_l increases to 0 and infinite is plotted in Fig. 2.

The saturation point of PA is determined by the physical characteristics and limitations of the amplifier itself, which is reflected by the coefficients in the polynomial model. The optimal input power and maximum SE correspond to the coordinates of the peak in the SE curve in Fig. 2. With these observations, we extend the well-known water-filling power allocation with the awareness of PA nonlinearity. Specifically, we first derive the optimal input power for PA of each AP ρ_l^* without imposing the total power constraints. Since the water-filling power allocation preserves its optimality when $\rho_l < \rho_l^*$, we obtain the power allocation with total power constraints, denoted by ρ_l^{wf}. Consequently, the selected input power of PA should choose the minimum between ρ_l^{wf} and ρ_l^*. The detailed process is described in the following.

The local optimal power ρ_l^* is obtained by solving the first-order condition in (24). As the partial derivative is always equal to zero in the floor region, ρ_l^* should take the

minimum value ρ satisfying that (24) is zero. Then, we apply the water-filling power allocation when $\rho_l < \rho_l^*$, where the PA works in the linear region, where the signal-to-noise ratio dominates the SE performance. In this case, the SE can be approximated by

$$\overline{R}_{\text{sum}} = \sum_{l=1}^{L} \log_2\left(1 + \frac{N\beta_1^2\lambda_l^2\rho_l}{\sigma_0^2}\right), \quad (31)$$

where $\overline{R}_{\text{sum}} = \sum_{l=1}^{L} \overline{R}_l$. $\lambda_l = \sqrt{NN_r/d_l^\alpha}$ is the single singular value of channel $\mathbf{H}_l$. Under the total power constraints, the input power of the l-th AP, determined by the water-filling power allocation, is given by

$$\rho_l^{\text{wf}} = \frac{1}{N}\left(\mu - \frac{\sigma_0^2}{\beta_1^2\lambda_l^2}\right)^+ = \frac{1}{N}\left(\mu - \frac{d_l^\alpha\sigma_0^2}{\beta_1^2NN_r}\right)^+, \quad (32)$$

where $(a)^+ = \max\{0, a\}$ and

$$\mu = \frac{P_{\text{total}} + \sum_{l=1}^{L} \frac{d_l^\alpha\sigma_0^2}{\beta_1^2NN_r}}{L}. \quad (33)$$

To avoid the PA working in deep saturation, the proposed distortion-aware water-filling (DA-WF) power allocation should take the minimum number of ρ_l^{wf} and ρ_l^*, as given by

$$\rho_l^{\text{DA-WF}} = \min(\rho_l^*, \rho_l^{\text{wf}}). \quad (34)$$

Namely, if $\rho_l^{\text{wf}} \leq \rho_l^*$, PA operates in the linear region, and the SNR performance dominates the SE performance. In this case, the utilization of water-filling power allocation is deemed optimal. On the other hand, if $\rho_l^{\text{wf}} > \rho_l^*$, applying water-filling power allocation could potentially drive the PA into saturation, causing distortion. In such cases, opting for ρ_l^* becomes imperative to mitigate distortion effects. Moreover, if the power allocation of any AP l satisfies $\rho_l^{\text{wf}} > \rho_l^*$, it will be excluded, and the remaining APs will reassign the power employing the water-filling methodology, thereby redistributing the available power allocation. Algorithm 2 summarizes the procedure of the proposed distortion-aware water-filling power allocation algorithm. The power constraint of the proposed distortion-aware water-filling power allocation algorithm (Algorithm 2) is ensured by the nature of the water-filling mechanism.

Complexity: The complexity of Algorithm 2 is from the water-falling computing step in (32). The water-filling algorithm typically converges in several iterations, denoted by I_{WF}. In each iteration, the computational complexity can be approximated as $O(L\log(L))$ since the sorting operation is the dominant factor. Hence, the overall complexity of the water-filling algorithm can be $\mathcal{O}(I_{WF}(L\log(L)))$, which is lower compared to the multiplier punitive function-based optimization method in Algorithm 1.

V. SIMULATION RESULTS

In this section, we present simulation results for a distributed MIMO system. We consider a distributed MIMO system with a single UE having $N_r = 10$ receiving antennas, and

---

Algorithm 2 Distortion-Aware Water-Filling Power Allocation Algorithm

1: Input: total input power constraint P_{total}, the channel between the UE to all AP, $\mathbf{H} = [\mathbf{H}_1, \cdots, \mathbf{H}_L]$.
2: Set $l \leftarrow 0$
3: **while** $l < L$ **do**
4:     Compute $\overline{R}_l$ using (23)
5:     Find local optimal average input power for ρ_l^* by solving the first-order condition in (24).
6: **end while**
7: Find power allocation under constraints with the water-filling method
8:     $\rho_l^{\text{wf}} = \frac{1}{N}\left(\mu - \frac{d_l^\alpha\sigma_0^2}{NN_r}\right)^+$
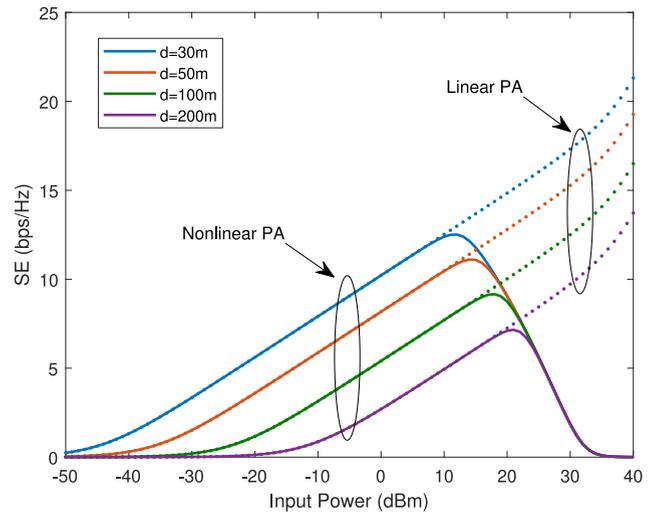9: Output $\rho^{\text{opt}} = \min(\rho_l^*, \rho_l^{\text{wf}})$.

FIGURE 3. SE under different input power (ρ_l) for the AP with different distances to the user. Dash line and solid line represent the spectral efficiency under linear and nonlinear PA assumptions, respectively.

each AP is equipped with $N = 20$ antennas unless specified otherwise. The simulation parameters are as follows: the path loss exponent $\alpha = 5$, and the PA model parameters are $\beta_1 = 2.96$, $\beta_3 = 0.1418e^{-J2.816}$, and $\beta_5 = 0.003e^{J0.39}$ [20]. We generate the line-of-sight channels between the UE and APs for each given scenario and evaluate the SE by using the Shannon formula, as shown in (17).

Fig. 3 plots the spectral efficiency of a single AP as a function of total input power. Under the linear PA assumption, the spectral efficiency increases with the input power (dashed line). However, as revealed in Section IV, the spectral efficiency is a non-monotonically increasing function of the input power due to the presence of nonlinearity distortion. When the input power is low, the spectral efficiency is dominated by the signal-to-noise ratio (SNR), and distortion is negligible compared to the noise level. In the lower SNR region, the performance of a distributed MIMO system coincides with the analysis under the linear
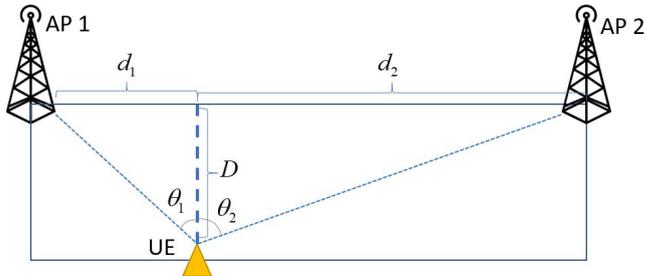
**FIGURE 4.** Illustration of the locations of UE and APs.

PA assumption. As the input power increases, the spectral efficiency decreases when distortion dominates performance rather than noise. It is also observed that the optimal input power changes with distance or path loss. The AP with a large distance to the UE has a higher optimal input power (ρ_l^*), where a higher input power is needed to compensate for path loss.

We evaluate the proposed algorithm under the considered two-AP scenario, where a UE is located in the middle of two APs, as illustrated in Fig. 4. Two benchmarks are chosen for comparison with the proposed Algorithm 1 and Algorithm 2:

- Water-Filling power allocation (WF), which is demonstrated as the capacity-optimal power allocation with a total power constraint. The power allocation to each AP can be obtained by (32).
- Equal power allocation (Equal), which allocates the power to each AP equally;

We consider that a UE moves between two arrays with a spacing of 400 m as illustrated in Fig. 4, and plot the sum spectral efficiency of the UE when moving from AP-1 to AP-2, as illustrated in Fig. 5(a). A large distortion can be observed when the UE is close to one array since all power is allocated to that array. Both the equal and water-filling power allocation suffer a great loss in spectral efficiency. For the same input power constraint, the WF and equal power allocation method allocate more power to the closet AP of the UE, which can push PAs into a deeper saturation regime. In this way, more distortion is generated due to PA saturation, thus degrading the performance. It is observed that the proposed Algorithm 2 achieves approximate optimal results with lower complexity, particularly when the multi-stream interference is negligible. Algorithm 1 demonstrates higher SE compared to the other schemes. Both Algorithms 1 and 2 can deal with distortion from the PA, but not with the interference that is left because of the non-ideal MRT. Algorithm 1 can deal with interference better in the high-power region, while Algorithm 2 only optimizes for distortion and creates strong multi-stream interference, as shown in Fig. 5(a).

Another observation is the gap between the upper bound between Algorithm 1 (as well as Algorithm 2) can be attributed to the imperfect removal of multi-stream interference when $N_r = 10$. We simulate the upper bound
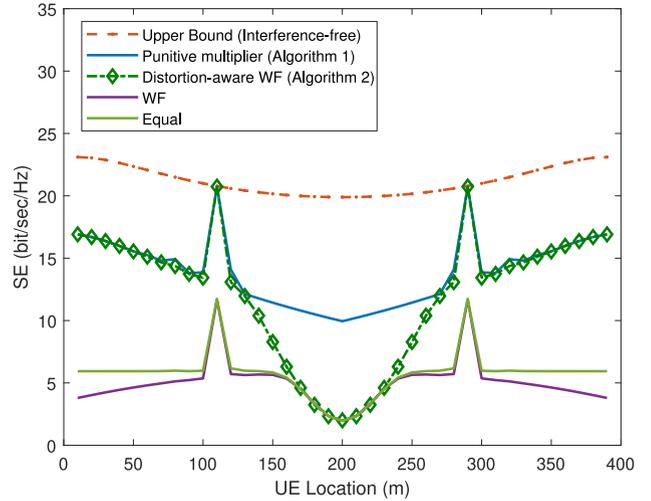
(a) Interference-considered scenario
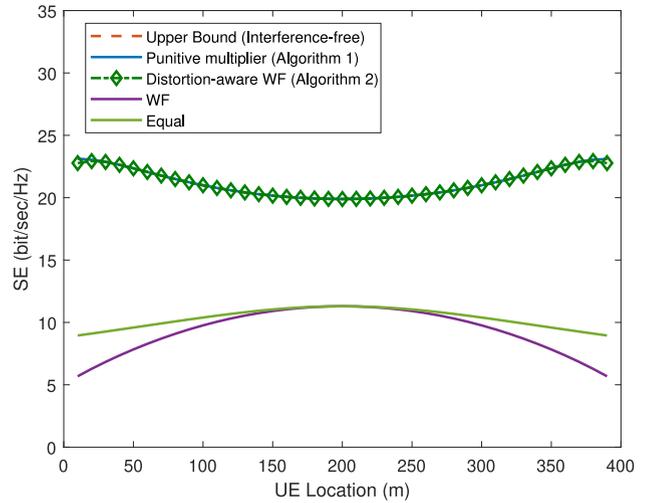
(b) Interference-free scenario

FIGURE 5. Spectral efficiency when UE moves between two APs: total power constraints $P_{\text{total}} = 43$ dBm, $N = 20$, $N_r = 10$, $D = 60$ m. (a) with interference consideration; (b) under the interference-free assumption.

of SE performance, assuming perfect removal of multi-stream interference by the spatial filter at the receiver side, i.e., UE. This scenario is achievable when the number of receiver antennas (N_r) is sufficiently large or through multi-band operation at different APs. The level of multi-stream interference is influenced by the geometric relationship between the APs and the UE, which determines the channel correlation in LoS channel conditions. It can also be observed that the proposed Algorithms 1 and 2 achieve a high SE when UE locates around d_1 (or d_2) = 110 m. In the LoS channel, the multi-streams interference from two APs is minimized when the channels between UE and two APs are orthogonal ($v_{1,2} = \mathbf{a}_r^H(\theta_1)\mathbf{a}_r(\theta_2) \rightarrow 0$), which leads to a higher SE (as depicted by peaks in Fig. 5(a)). It is evident that, in this scenario, the proposed algorithm achieves performance close to the upper bound performance.
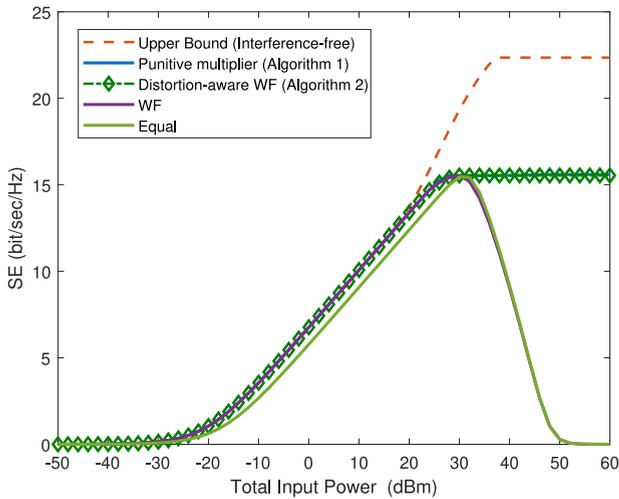
FIGURE 6. Spectral efficiency of UE versus total power constraints P_{total}: two APs with distance $d_1 = 50$ m and $d_2 = 350$ m, $D = 60$ m.

FIGURE 7. Convergence performance versus the number of iterations for the Algorithms 1 and 2.

On the other hand, it is found that both equal power allocation and water-filling (WF) power allocation suffer a significant loss in SE when the user is close to one of the APs. For the same input power constraint, the WF and equal power allocation method allocate more power to the closest AP of the UE without awareness of PA's nonlinearity, which can push PAs into deeper saturation regimes. In this way, more distortion is generated due to PA saturation, degrading the SE performance. Fig. 5 plots the spectral efficiency of the UE under the interference-free assumption with the same settings as Fig. 5(a). The proposed Algorithm 1 and Algorithm 2 can reach optimal (upper bound) performance when the impact of multi-stream interference is negligible. It is also equal and WF power allocation suffers a great loss in SE performance due to the distortion.

Fig. 6 plots the SE of the UE under different total power constraints P_{total}, where the UE is located at distances $d_1 = 50$ m and $d_2 = 350$ m from the two APs. The gap between the proposed algorithm and the upper bound performance is due to non-ideal combining at the UE side, since the multi-stream interference is not ideally mitigated when the antenna number is not sufficiently large, e.g., $N_r = 10$. It can be observed in Fig. 6 that the proposed Algorithm 2 achieves nearly optimal results with lower complexity. Furthermore, the WF method demonstrates comparable performance in the low input power regime, which validates the motivation behind utilizing the WF method within the linear working regime of the PA when designing the distortion-aware water-filling algorithm (Algorithm 2). With awareness of distortion, the proposed algorithm will avoid allocating more power after reaching a certain power threshold. In this sense, the PA can be prevented from working in deep saturation.

Fig. 7 plots the average convergence performance of Algorithms 1 and 2, respectively, under the same settings used as Fig. 6 when $P_{total} = 30$ dBm. It is observed
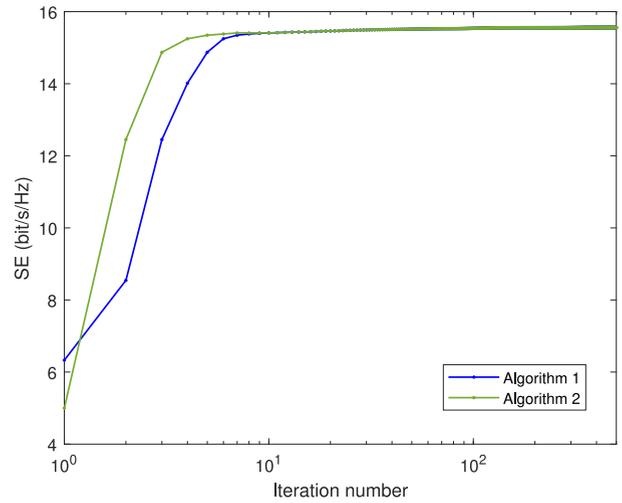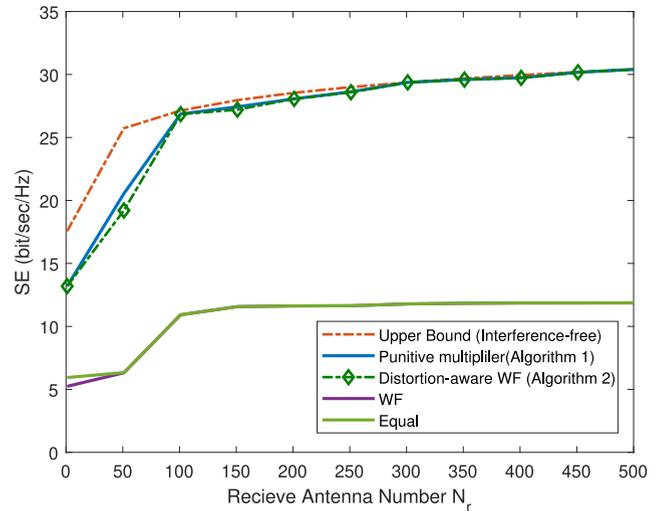

FIGURE 8. Spectral efficiency under different antenna number N_r: distance to the UE $d_1 = 50$ m and $d_2 = 350$ m, $D = 60$ m, $P_{total} = 43$ dBm.

that Algorithm 2 achieves fast convergence compared to Algorithm 1, which requires more iterations due to the nature of the multiplier punitive function-based method.

Fig. 8 plots the SE under varying numbers of antennas at the receiver (UE). It can be observed that as the number of received antennas at the UE becomes sufficiently large, for example, $N_r > 200$, the impact of multi-stream interference diminishes. The increased number of antennas enhances the channel orthogonality among different APs, enabling the simultaneous transmission and reception of independent data streams. Moreover, it is noteworthy that the SE increases with the number of received antennas (N_r), as this leads to a higher achievable channel gain. However, the SE achieved through WF and equal power allocation approaches do not scale proportionally with the number of received antennas (N_r). Without considering the PA nonlinearity, WF and equal power allocation methods can drive the power amplifiers
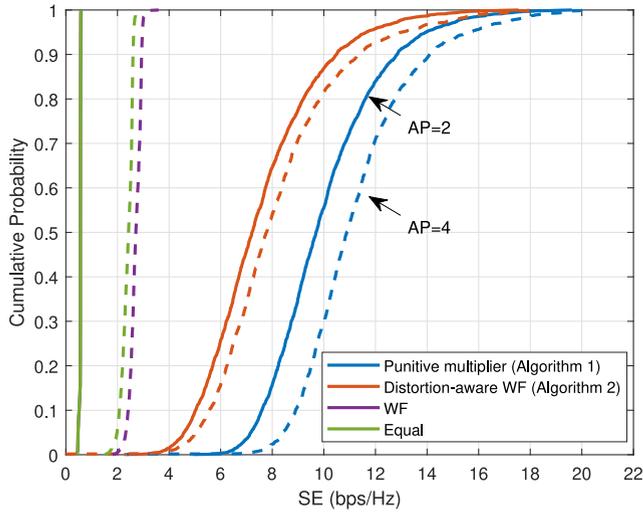
FIGURE 9. Cumulative distribution function of spectral efficiency when different number of APs with randomly generated location: total power constraints $P_{total} = 43$ dBm, $N = 20$, $N_r = 100$.

FIGURE 10. Cumulative distribution function of spectral efficiency of two users when different number of APs with randomly generated location: total power constraints $P_{total} = 43$ dBm, $N = 20$, $N_r = 100$.

into saturation, increasing distortion. Consequently, the SE encounters limitations in growth, even in the presence of considerable channel gain.

Fig. 9 presents the cumulative distribution function (CDF) of the SE for the UE as the number of deployed APs varies from 2 to 4. We generated 2,000 realizations by randomly placing the APs with an average distance of 400 m from the UE. The resulting CDF of SE is depicted in Fig. 9. It is observed that the SE increases when more APs are available in the considered area. This occurs because the UE is more likely to be served by APs closer to it. Moreover, as the number of APs increases, the improved SE performance with water filling and equal power allocation can be attributed to the reduced power allocation to each AP, resulting in lower distortion and enhanced SE. Furthermore, it is evident that the performance gap between the proposed Algorithm1 and Algorithm 2 becomes more pronounced as the number of APs increases. This discrepancy arises because Algorithm 1 handles interference more effectively, while Algorithm 2 primarily optimizes for distortion, making it challenging to cope with strong multi-stream scenarios as the number of APs rises.

Fig. 10 illustrates the cumulative distribution function (CDF) of the SE of two UEs, while varying the number of deployed APs from 2 to 4. Through 2,000 random placements of APs and UEs with an average separation of 400 m, the resulting CDF of SE is captured in Fig. 10. Similar to Fig. 9, it is observed that the performance gap between the proposed Algorithm 1 and Algorithm 2 becomes more pronounced as the number of APs and UEs increases. This distinction arises due to the more effective interference handling by Algorithm 1, whereas Algorithm 2 primarily focuses on distortion optimization, posing challenges in handling robust multi-stream and multi-user scenarios. Moreover, a less steep CDF curve of SE
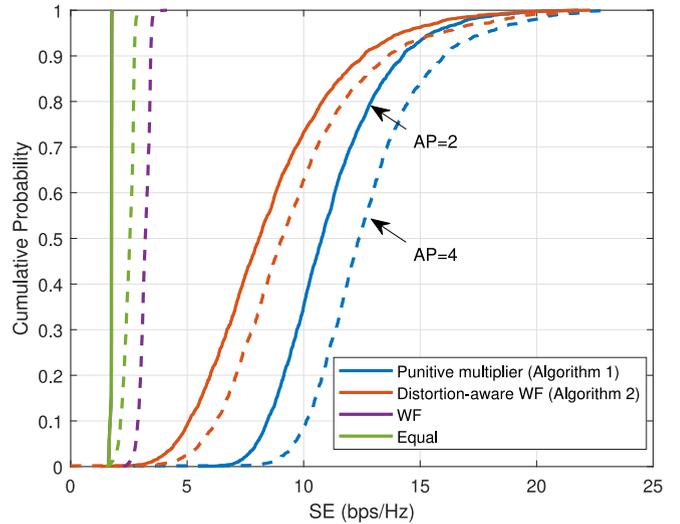
is observed with the increase in the number of UEs. This indicates that SE exhibits significant variability under different UEs and APs deployment, and is more pronounced when the number of UEs increases under our considered settings.

VI. CONCLUSION

In this paper, we studied AP selection and power allocation in the multi-stream distributed massive MIMO system with nonlinear PAs. With a power constraint, the power allocation for the SE maximization is formulated as nonlinear programming, which is first solved by a multiplier punitive function-based algorithm. We further study the special case when multi-stream interference among APs is negligible, and extend the well-known water-filling power allocation method by considering PA's nonlinearity, solving the power allocation problem at lower complexity. Extensive simulations corroborate that the proposed multiplier punitive function-based power allocation methods can maximize the SE with awareness of the PA nonlinearity, and the proposed distortion-aware water-filling power allocation can achieve near-optimal results with lower complexity in the interference-limited case.

This paper focuses on the impact of PA nonlinearity in spectral efficiency performance analysis and power allocation of a multi-stream distributed massive MIMO system. The perfect channel state information is assumed to assess the performance bounds of the system. While our current work primarily concentrates on PA nonlinearity, we acknowledge the significance of channel estimation imperfections in the performance analysis. Future iterations of research should consider the impact of imperfect channel estimation and non-coherent transmissions with PA nonlinearity in distributed MIMO systems. Further studies could delve into the development of adaptive algorithms that dynamically

adjust power allocation and bit-loading based on real-time channel conditions, ensuring improved performance in practical deployments. Furthermore, we recognize the significance of incorporating all resources, such as user-AP association, beamforming, and per-user power allocation, alongside implementing fairness policies for multi-user scenarios, as key aspects for future research.

APPENDIX
PROOF OF LINEAR POWER GAIN

Proof: Following the methodology in [29], we give the sketch proof of linear gain. The average linear amplification gain can be written by using the Bussgang theorem

$$
\begin{aligned}
\bar{g}_{l,n} &= \frac{\mathbb{E}\{x_{l,n}s_{l,n}^*\}}{\mathbb{E}\{|x_{l,n}|^2\}} \overset{(a)}{=} \frac{1}{P_{l,n}}\mathbb{E}\left\{\sum_{m=0}^{M}\beta_{2m+1}\,|s_{l,n}|^{2m+2}\right\} \\
&= \frac{1}{P_{l,n}}\sum_{m=0}^{M}\beta_{2m+1}\mathbb{E}\left\{|s_{l,n}|^{2m+2}\right\} \\
&= \sum_{m=0}^{M}\beta_{2m+1}(m+1)!P_{l,n}^{m}.
\end{aligned}
\tag{35}
$$

where (a) is by definition $\mathbb{E}(|\mathbf{s}_{l,n}|^2) = P_{l,n}$. The distortion are uncorrelated with the linear signal, i.e., $\mathbb{E}\{s_{l,n}^* e_{l,n}\} = 0$. ∎

REFERENCES

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

[2] T. T. Vu et al., "Energy-efficient massive MIMO for federated learning: Transmission designs and resource allocations," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 2329–2346, 2022.

[3] D. Wang, C. Zhang, Y. Du, J. Zhao, M. Jiang, and X. You, "Implementation of a cloud-based cell-free distributed massive MIMO system," *IEEE Commun. Mag.*, vol. 58, no. 8, pp. 61–67, Aug. 2020.

[4] X. Yu, W. Xu, S.-H. Leung, Q. Shi, and J. Chu, "Power allocation for energy efficient optimization of distributed MIMO system with beamforming," *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 8966–8981, Sep. 2019.

[5] N. Souto and J. C. Silva, "Joint beamforming algorithm for multi-stream MIMO systems assisted by multiple reconfigurable intelligent surfaces," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 1317–1333, 2023.

[6] S. Gao, X. Cheng, and L. Yang, "Spatial multiplexing with limited RF chains: Generalized beamspace modulation (GBM) for mmwave massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 9, pp. 2029–2039, Sep. 2019.

[7] D. Rönnow and P. Händel, "Nonlinear distortion noise and linear attenuation in MIMO systems—Theory and application to multi-band transmitters," *IEEE Trans. Signal Process*, vol. 67, no. 20, pp. 5203–5212, Oct. 2019.

[8] B. Liu, F. Rottenberg, and S. Pollin, "Power allocation for distributed massive LoS MIMO with nonlinear power amplifiers," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, 2022, pp. 1–5.

[9] B. Liu, Q. Wang, and S. Pollin, "Enhancing indoor-to-outdoor mmWave communication with transparent amplifying intelligent surface," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2023, pp. 6467–6473.

[10] Y. Xu, J. Pang, X. Wang, and A. Zhu, "Enhancing bandwidth and back-off range of Doherty power amplifier with modified load modulation network," *IEEE Trans. Microw. Theory Tech.*, vol. 69, no. 4, pp. 2291–2303, Apr. 2021.

[11] H. Yan, S. Ramesh, T. Gallagher, C. Ling, and D. Cabric, "Performance, power, and area design trade-offs in millimeter-wave transmitter beamforming architectures," *IEEE Circuits Syst. Mag.*, vol. 19, no. 2, pp. 33–58, May 2019.

[12] W. Chen, G. Lv, X. Liu, D. Wang, and F. M. Ghannouchi, "Doherty pas for 5G massive MIMO: Energy-efficient integrated DPA MMICs for sub-6-GHZ and mm-wave 5G massive MIMO systems," *IEEE Microw. Mag.*, vol. 21, no. 5, pp. 78–93, May 2020.

[13] T. Van Chien, E. Björnson, and E. G. Larsson, "Joint power allocation and load balancing optimization for energy-efficient cell-free massive MIMO networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6798–6812, Oct. 2020.

[14] N. Akbar, E. Björnson, N. Yang, and E. G. Larsson, "Max-min power control in downlink massive MIMO with distributed antenna arrays," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 740–751, Feb. 2021.

[15] G. Dong, H. Zhang, S. Jin, and D. Yuan, "Energy-efficiency-oriented joint user association and power allocation in distributed massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 5794–5808, Jun. 2019.

[16] J. Myung, K. Kim, and Y.-J. Ko, "Transmit power control game with a sum power constraint in cell-free massive MIMO downlink," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 632–635, Mar. 2022.

[17] X. Yu, T. Teng, X. Dang, S.-H. Leung, and F. Xu, "Joint power allocation and beamforming for energy-efficient design in multiuser distributed MIMO systems," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4128–4143, Jun. 2021.

[18] Y. Zou et al., "Impact of power amplifier nonlinearities in multi-user massive MIMO downlink," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, 2015, pp. 1–7.

[19] B. Liu, A. P. Guevara, L. Van der Perre, and S. Pollin, "Nonlinear distortion in distributed Massive MIMO systems: An indoor chan-nel measurement analysis," in *Proc. IEEE Glob. Commun. Conf. Workshops (GLOBECOM Workshops)*, 2021, pp. 1–6.

[20] S. R. Aghdam, S. Jacobsson, U. Gustavsson, C. Durisi, C. Studer, and T. Eriksson, "Distortion-aware linear precoding for massive MIMO downlink systems with nonlinear power amplifiers," 2020, *arXiv:2012.13337*.

[21] F. Rottenberg, G. Callebaut, and L. Van der Perre, "The Z3RO family of precoders cancelling nonlinear power amplification distortion in large array systems," *IEEE Trans. Wireless Commun.*, vol. 22, no. 3, pp. 2036–2047, Mar. 2023.

[22] E. G. Larsson and L. Van der Perre, "Out-of-band radiation from antenna arrays clarified," *IEEE Commun. Lett.*, vol. 7, no. 4, pp. 610–613, Aug. 2018.

[23] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[24] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4247–4261, Jul. 2020.

[25] C. Mollén, U. Gustavsson, T. Eriksson, and E. G. Larsson, "Spatial characteristics of distortion radiated from antenna arrays with transceiver nonlinearities," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6663–6679, Oct. 2018.

[26] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[27] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418–4431, Jul. 2012.

[28] B. Liu, Q. Wang, and S. Pollin, "TAIS: Transparent amplifying intelligent surface for indoor-to-outdoor mmWave communications," *IEEE Trans. Commun.*, vol. 72, no. 2, pp. 1223–1238, Feb. 2024.

[29] N. N. Moghadam, G. Fodor, M. Bengtsson, and D. J. Love, "On the energy efficiency of MIMO hybrid beamforming for millimeter-wave systems with nonlinear power amplifiers," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7208–7221, Nov. 2018.

[30] E. Björnson, L. Sanguinetti, and J. Hoydis, "Hardware distortion correlation has negligible impact on ul massive MIMO spectral efficiency," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1085–1098, Feb. 2019.

[31] D. Dardari, V. Tralli, and A. Vaccari, "A theoretical characterization of nonlinear distortion effects in OFDM systems," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1755–1764, Oct. 2000.

[32] R. Corvaja and A. G. Armada, "Phase noise degradation in massive MIMO downlink with zero-forcing and maximum ratio transmission precoding," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8052–8059, Oct. 2016.

[33] P. J. Smith and M. Shafi, "On a Gaussian approximation to the capacity of wireless MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, vol. 1, 2002, pp. 406–410.

[34] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Trans. Signal Process*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.

[35] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[36] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 506–522, Feb. 2005.

[37] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput analysis of massive MIMO uplink with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 4038–4051, Jun. 2017.

[38] S. Wei, D. L. Goeckel, and P. A. Kelly, "Convergence of the complex envelope of bandlimited OFDM signals," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4893–4904, Oct. 2010.

[39] D.-M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavcic, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3498–3508, Aug. 2006.

[40] D. G. Owen, "A punitive damages overview: Functions, problems and reform," *Vill. L. Rev.*, vol. 39, no. 2, p. 363, 1994.

[41] L. Wang et al., "Energy efficient train-ground mmWave mobile relay system for high speed railways," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 1, pp. 16–28, Mar. 2023.

[42] M. Kobayashi and G. Caire, "An iterative water-filling algorithm for maximum weighted sum-rate of Gaussian MIMO-BC," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1640–1646, Aug. 2006.

BIN LIU (Member, IEEE) is currently pursuing the Ph.D. degree in electrical engineering with KU Leuven, Belgium. From March to October 2022, he was a Visiting Scholar with the Delft University of Technology, The Netherlands. His research interests encompass signal processing in massive MIMO, reconfigurable intelligent surfaces, and mmWave communications. He was awarded the Best Paper at the IEEE ICC 2023.

SOFIE POLLIN (Senior Member, IEEE) is a Professor with KU Leuven focusing on wireless communication systems. Before that, she worked with imec and the University of California at Berkeley, and she is currently still a Principal Member of technical staff with imec. Her research centers around wireless networks that require networks that are ever more dense, heterogeneous, battery powered, and spectrum constrained. Her research interests are cell-free networks, integrated communication and sensing, and nonterrestrial networks.