

RESEARCH

Open Access



Energy-efficient power allocation for massive MIMO-enabled multi-way AF relay networks with channel aging

Fangqing Tan¹, Hongbin Chen^{1*} , Feng Zhao² and Xiaohuan Li¹

Abstract

In this paper, we consider a massive MIMO-enabled multi-way amplify-and-forward relay network with channel aging, where multiple users mutually exchange information via an intermediate relay equipped with massive antennas. For this system, we propose an energy-efficient power allocation scheme for the optimization of energy efficiency (EE). Specifically, we firstly derive accurate closed-form expressions of the system sum rate with aged channel state information (CSI) and predicted CSI. Secondly, based on the derived analytical results, a unified power allocation optimization problem with aged/predicted CSI is formulated for maximizing the system EE. To solve this challenging problem, the successive convex approximation technique is invoked to transform the original optimization problem into a tractable concave fractional programming problem. Then, Dinkelbach's algorithm and Lagrangian dual method are adopted to find the optimal solution. In addition, to strike a balance between the computational complexity and the optimality, the EE maximization problem using the equal power allocation scheme is solved by extreme value theorem, leading to a closed-form optimal solution. Numerical results demonstrate the accuracy of our analytical results and the effectiveness of the proposed algorithms. Moreover, the impact of several important system parameters on the system performance achieved by the proposed algorithms is also illustrated.

Keywords: Massive MIMO, Multi-way relaying, Channel aging, Energy efficiency, Power allocation

1 Introduction

With the unprecedented growth of mobile data traffic volumes, the carbon emission of information and communication technologies is becoming an increasingly serious problem. Internationally, several academic and industrial research projects have been dedicated to maximize the overall network capacity and improve the energy efficiency (EE) of wireless communication systems (i.e., minimizing the amount of energy required to transmit data) [1–4]. More recently, as one of the major candidate technologies for fifth-generation (5G) wireless systems, massive multiple-input multiple-output (MIMO) has received tremendous attention from both academic and industry in wireless fields [5, 6].

More specifically, it was shown in [7] that massive MIMO systems (equipped with a very large number of

antennas) are capable of achieving three orders of magnitude EE gains compared with single-antenna systems. The energy-efficient design of massive MIMO systems has emerged as a new research trend for 5G wireless communications [8, 9]. For example, in [10], the EE was analyzed in massive MIMO systems, under the effect of a general transceiver hardware impairments. In [11], the EE was maximized in a multi-user massive MIMO system, and the optimal system parameters (includes the number of base station (BS)'s antennas and users). In [12], the BS density, transmit power levels and number of antennas are optimized for maximizing EE in massive MIMO-enabled heterogeneous networks.

As another promising approach, multi-way relay networks (MWRNs) have recently received plenty of research interest [13, 14]. In general, as compared to one-way relay networks (OWRNs) and two-way relay networks (TWRNs), MWRNs are capable of achieving higher capacity and spectral efficiency (SE) and thus can be employed to effectively deal with the ever increasing

*Correspondence: chbscut@guet.edu.cn

¹Key Laboratory of Cognitive Radio and Information Processing, Guilin University of Electronic Technology, Guilin 541004, China
Full list of author information is available at the end of the article

demand for higher data rate and SE in a multi-user scenario. Therefore, the integration of massive MIMO and MWRN is regarded as a promising network architecture to meet the significant demand for mobile data applications. Additionally, it was shown in [15] that using simple relay transceivers (e.g., linear zero-forcing (ZF) transceiver), a multi-way MIMO relay system is capable of significantly alleviating the interference among different data streams/user equipments (UEs). Furthermore, similar to the observations in massive MIMO-enabled OWRNs [16] and TWRNs [17, 18], it was shown in [19] that by invoking a large-scale antenna array-equipped relay and a low-complexity ZF transceiver, the SE of MWRNs is also proportional to the number of relay antennas. At present, the existing related research on massive MIMO-enabled MWRNs mainly focused on analyzing the performance limits in various specific system configurations [19–24]. For instance, in [19], the asymptotic signal-to-interference-plus-noise ratio (SINR) of massive MIMO-enabled MWRNs was studied. Later on, the authors of [20] further analyzed the asymptotic SINR and average error rate performance and obtained the optimal pilot sequence length for maximizing the SE of multi-cell massive MIMO-enabled MWRNs. Moreover, the SE and asymptotic SINR of massive MIMO-enabled MWRNs are first analyzed with maximum-ratio processing, and then the same authors derived a closed form expression of the SE of massive MIMO-enabled MWRNs with ZF processing.

Moreover, prior works [19–22] considered the effect of channel imperfection due to channel estimation (CE), but ignored another important aspect of practical channel impairments known as *channel aging*, which refers to the phenomenon affected by the relative movement of users. This scenario is of high practical value in urban environments, where users move rapidly within a geographical area. Despite its significance, very few works have investigated its impact on the performance of massive MIMO systems. For point-to-point massive MIMO system, the impact of channel aging on the SINR performance was firstly studied by assuming matched filtering [25]. The impact of channel aging was lately investigated with ZF precoders [26] and minimum mean square error (MMSE) receivers [27]. For massive MIMO relay system, the asymptotic impact of channel aging on the performance of massive MIMO-enabled MWRNs was studied in [23]. Later on, the analysis was extended to the multi-cell massive MIMO-enabled MWRNs scenario for simultaneous wireless information and power transfer [24]. To the best of the authors's knowledge, there is a paucity of contributions on energy-efficient transmission strategies of massive MIMO-enabled MWRNs, considering the effect of channel aging. It is challenging to extend the existing energy-efficient designs conceived for single-hop

massive MIMO systems [10–12] to massive MIMO-enabled relay systems. Due to this fact, compared to single-hop transmission schemes, both signal processing schemes and the performance analysis of massive MIMO-enabled relay systems are fundamentally dependent on the more complex two-hop channels. Therefore, it is important to design energy-efficient transmission strategy for massive MIMO-enabled relay systems. Furthermore, the consideration of channel aging is of paramount importance because it can provide the robustness against the practical setting of user mobility that results to delayed and degraded channel state information (CSI).

Motivated by the above discussions, in this paper, we investigate low-complexity energy-efficient power allocation strategies for a massive MIMO-enabled MWRN with channel aging.¹ We assume that the CSI is estimated relying on the MMSE criterion, and the relay employs the low-complexity linear ZF transceivers. The main contributions of this paper are summarized as follows.

- We respectively derive closed-form expression of the achievable sum rate (SR) for aged and predicted CSI, which enables us to efficiently evaluate the system performance, thus facilitating the energy-efficient power allocation strategies.
- Based on the derived closed-form expressions, we formulate a unified optimization problem that optimizes power allocation of all UEs for maximizing the system EE, subject to limited transmit power, and minimum quality-of-service (QoS) constraints. Because of the intractable non-convexity of the formulated optimization problem, the successive convex approximation (SCA) technique is involved to transform the non-convex problem into a concave fractional programming (CFP) problem, which is then efficiently solved by Dinkelbach's algorithm and Lagrangian dual method.
- Furthermore, to strike a balance between the computational complexity and the optimality, a closed-form power control algorithm is provided under the assumption of equal power allocation (EPA) among multiple UEs, without requiring complicated iterative algorithms.
- By simulation, the impact of the maximal transmit power, of the QoS constraint, and of the transmit power of each pilot symbol on the optimum EE is quantified. Moreover, our numerical results show that the EPA scheme-based power optimization strategies strike an attractive tradeoff between the achievable EE performance and the computational complexity imposed.

The remainder of this paper is organized as follows. The system model is described in Section 3. In Section 4, the closed-form expressions for SR are derived under aged

and predicted CSI scenarios. In Section 5, we present the energy-efficient power allocation optimization problem under different performance criteria and constraints. The power allocation strategies are provided in Section 6, and the simulation results are given in Section 7. Finally, the conclusions are made in Section 8.

Notations: We use uppercase and lowercase boldface letters for denoting matrices and vectors, respectively. $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote the conjugate, transpose, and conjugate transpose, respectively. $\|\cdot\|$, $\text{tr}\{\cdot\}$, $\mathbb{E}\{\cdot\}$, $\text{Cov}(\cdot)$, and $\text{Var}\{\cdot\}$ stand for the Euclidean norm, the trace of matrices, the expectation, covariation, and variance operators, respectively. The $\text{diag}\{\mathbf{x}\}$ denotes a diagonal matrix with the vector \mathbf{x} being its diagonal entries, and the operators $\text{mod}_N(x)$ denote the modulo N of x . $[\mathbf{A}]_{i,j}$ represents the entry at the i -th row and the j -th column of a matrix \mathbf{A} . Finally, $\mathcal{CN}(\mathbf{0}, \Theta)$ denotes the circularly symmetric complex Gaussian distribution with zero mean and the covariance matrix Θ .

2 Method

This paper studies the energy-efficient power allocation problem of massive MIMO-enabled multi-way relay systems, under channel aging. The performance of the proposed framework was in depth examined through a series of simulation experiments including different system parameters, whereas the superiority of the proposed approach was clearly demonstrated by comparing it with other research works in the literature. Specifically, it has been shown that the different implementations of the proposed algorithm succeed in providing considerably higher EE in all different system settings while at the same time maintaining QoS at high levels. Moreover, the impact of normalized Doppler shifts $f_D T_S$ (i.e., channel aging) on the system achievable rate is also illustrated. The simulation code was written in MATLAB.

3 System model and transmission scheme

As shown in Fig. 1, we consider a massive MIMO-enabled AF MWRN with non-pairwise ZF transmission,² where K spatially distributed single-antenna UEs (UE_k , $k \in \{1, \dots, K\}$) exchange their information-bearing signals in K time slots among one another via a shared relay (R) equipped with M antennas.³ Without loss of significant generality, we assume that the number of relay antennas is greater than the number of UEs served at the same time-frequency resources (i.e., $M > K$). The system operates over a bandwidth of B Hz and the channels are static within the time-frequency coherence blocks composed of $T = B_C T_C$ data symbols, where B_C and T_C are the coherence bandwidth and coherence time, respectively. It is assumed that the channel coefficients do not change within one-symbol duration, but vary slowly from symbol to symbol. We assume that the relay operates on the

half-duplex TDD mode. Each coherence interval is divided into three time phases, i.e., the CE phase, the multiple-access and broadcast phases. The multiple-access phase consists of only one time slot, whereas the broadcast phase contains $K - 1$ time slots.

3.1 Data transmission

In the multiple-access phase, all K UEs simultaneously transmit their signals $\mathbf{x}_U[n]$ to the relay. These signals can be expressed as $\mathbf{x}_U[n] = \mathbf{P}_u^{1/2} \mathbf{s}[n] \in \mathbb{C}^{K \times 1}$, where $\mathbf{s}[n] = [s_1[n], \dots, s_K[n]]^T$ is the information-bearing symbol vector with $\mathbb{E}[\mathbf{s}[n] \mathbf{s}^H[n]] = \mathbf{I}_K$ and $\mathbf{P}_u = \text{diag}\{p_1, \dots, p_k, \dots, p_K\}$, p_k is the transmit power of the k th UE. The received signal $\mathbf{y}_R \in \mathbb{C}^{M \times 1}$ at the relay is given by

$$\mathbf{y}_R[n] = \mathbf{G}[n] \mathbf{x}_U[n] + \mathbf{n}_R[n], \quad (1)$$

where $\mathbf{G}[n] \in \mathbb{C}^{M \times K}$ represents the channel matrix from K UEs to the relay and $\mathbf{n}_R[n]$ denotes the additive white Gaussian noise (AWGN) that obeys $\mathcal{CN}(\mathbf{0}, \sigma_r^2 \mathbf{I}_M)$ at the relay.

To be specific, the channel matrix $\mathbf{G}[n]$ can be expressed as

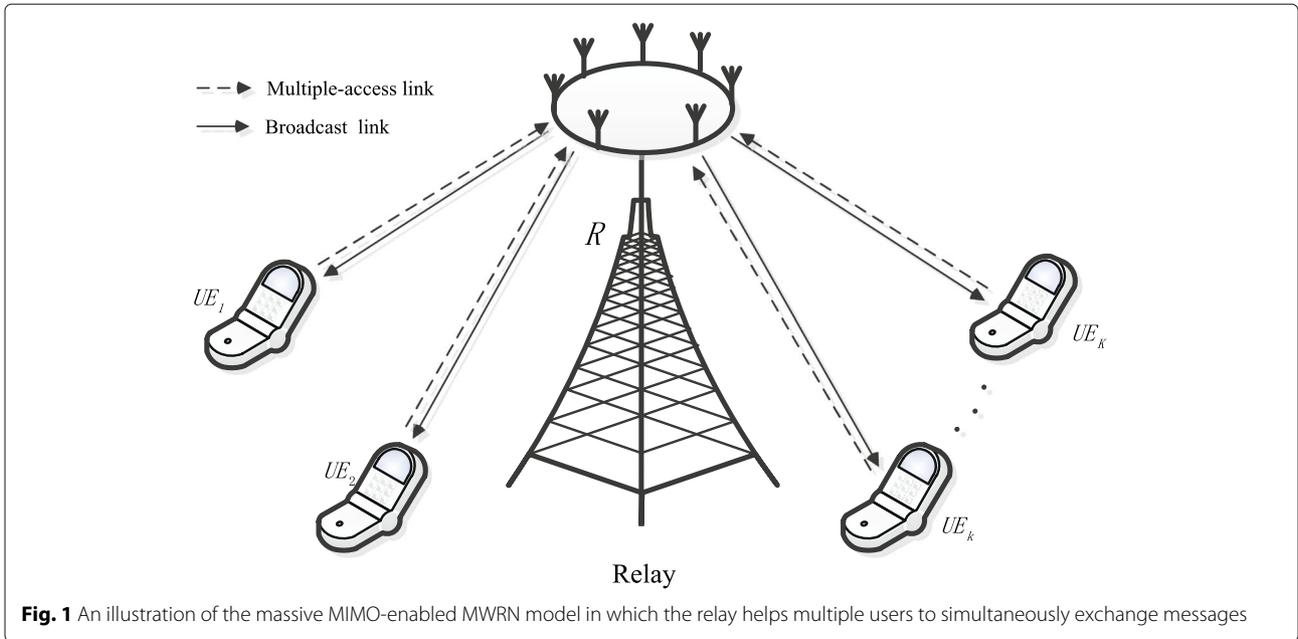
$$\mathbf{G}[n] = \mathbf{H}[n] \mathbf{D}^{1/2}, \quad (2)$$

where $\mathbf{H}[n] \in \mathbb{C}^{M \times K}$ is the small-scale fading (SSF) channel matrix and their entries obey independent identically distributed (i.i.d.) Gaussian distribution as $\mathcal{CN}(0, 1)$. \mathbf{D} is a $K \times K$ diagonal matrix with $[\mathbf{D}]_{k,k} = \beta_k$, which models the large-scale fading (LSF) capturing both path-loss and shadowing fading effects. Moreover, β_k is assumed to remain constant for all n and is assumed to be known a priori as it changes very slowly compared with SSF channel coefficients.

In the broadcast phase, the relay simply performs transceive processing, which firstly detects the signals received and transmitted to all UEs in $K - 1$ subsequent time slots. Here, we consider an intermediate j th ($j \in \{1, \dots, K - 1\}$) time slot of the broadcast phase for the sake of exposition. In the context, the relay transmitted signal in the j th time slot of the broadcast phase is given by

$$\mathbf{x}_R^{(j)}[n] = \vartheta_j \mathbf{F}_j[n] \mathbf{y}_R[n]. \quad (3)$$

where $\mathbf{F}_j[n] = \mathbf{W}_2[n] \boldsymbol{\pi}_j \mathbf{W}_1[n]$ is the combined beamforming matrix at the relay, $\mathbf{W}_1[n]$ is a ZF detection matrix, and $\mathbf{W}_2[n]$ is a ZF precoding matrix. Moreover, $\boldsymbol{\pi}_j$ is the permutation matrix employed at the relay in the j th time slot of the broadcast phase, which is designed to ensure that the k th ($k \in \{1, \dots, K\}$) UE, receives the signal from the k' th UE, with $k' = \text{mod}_K(k + j)$. Specifically,



π_j is constructed as $\pi_j = (\pi_o)^j$, and the $K \times K$ primary permutation matrix, π_o , can be written as

$$\pi_o = [\mathbf{e}_K, \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{K-2}, \mathbf{e}_{K-1}], \tag{4}$$

where \mathbf{e}_k denotes a column vector of length K with 1 in the k th position and 0 in every other position. ϑ_j is the amplification factor designed to constrain the long-term relay transmit power p_r , which is given as follows.

$$\vartheta_j = \sqrt{\frac{p_r}{\mathbb{E}[\|\mathbf{F}_j[n] \mathbf{y}_R[n]\|^2]}}. \tag{5}$$

Then, the received signal vector at K UEs in the j th time slot of the broadcast phase can be written as follows.

$$\mathbf{y}_u^{(j)}[n] = \mathbf{G}^T[n] \mathbf{x}_R^{(j)}[n] + \mathbf{n}_u[n], \tag{6}$$

where $\mathbf{n}_u[n]$ is the AWGN vector satisfying $\mathbf{n}_u[n] \sim \mathcal{CN}(\mathbf{0}, \sigma_u^2 \mathbf{I}_K)$. The aforementioned broadcast phase continues until the completion of all $K - 1$ relay transmissions in $K - 1$ consecutive time slots.

Substituting (1) and (3) into (6), the received signal at the k th UE in the j th time slot of the broadcast phase is expressed as

$$\begin{aligned} y_{u,k}^{(j)}[n] = & \underbrace{\vartheta_j \sqrt{p_{k'}} \mathbf{g}_k^T[n] \mathbf{F}_j[n] \mathbf{g}_{k'}[n] s_{k'}[n]}_{\text{desired signal}} \\ & + \underbrace{\vartheta_j \sqrt{p_k} \mathbf{g}_k^T[n] \mathbf{F}_j[n] \mathbf{g}_k[n] s_k[n]}_{\text{self-interference}} \\ & + \underbrace{\vartheta_j \sum_{i \neq k, k'}^K \sqrt{p_i} \mathbf{g}_k^T[n] \mathbf{F}_j[n] \mathbf{g}_i[n] s_i[n]}_{\text{inter-user interference}} \\ & + \underbrace{\vartheta_j \mathbf{g}_k^T[n] \mathbf{F}_j[n] \mathbf{n}_R[n]}_{\text{noise from relay}} + \underbrace{n_{u,k}[n]}_{\text{noise at user}}, \end{aligned} \tag{7}$$

where $\mathbf{g}_k[n]$ is the k -th column of $\mathbf{G}[n]$ and $n_{u,k}[n]$ is the k -th element of $\mathbf{n}_u[n]$.

3.2 Channel estimation

The relay estimates the channel coefficients by transmitting orthogonal pilot sequences. All K UEs simultaneously transmit their pilot sequences of τ_r ($\tau_r \geq K$) symbols to the relay. The received pilot matrix at the relay is given by

$$\mathbf{X}[n] = \sqrt{\tau_r p_p} \mathbf{G}[n] \Phi + \mathbf{Z}[n], \tag{8}$$

where p_p is the transmit power of each pilot symbol and $\mathbf{Z}[n] \in \mathbb{C}^{M \times \tau_r}$ is a noise matrix whose elements are i.i.d $\mathcal{CN}(0, \sigma_r^2)$. $\Phi \in \mathbb{C}^{K \times \tau_s}$ is the pilot sequence matrix transmitted from K UEs, satisfying $\Phi \Phi^H = \mathbf{I}_K$. Correlation of the received signal $\mathbf{X}[n]$ with $\frac{1}{\sqrt{\tau_r p_p}} \Phi^H$ obtains

$$\tilde{\mathbf{X}}[n] = \frac{1}{\sqrt{\tau_r p_p}} \mathbf{X}[n] \Phi^H. \tag{9}$$

Therefore, the noisy observation of the channel vector from k th UE to the relay is expressed as

$$\tilde{\mathbf{x}}_k[n] = \mathbf{g}_k[n] + \frac{1}{\sqrt{\tau_r p_p}} \tilde{\mathbf{z}}_k[n]. \quad (10)$$

where $\tilde{\mathbf{z}}_k[n]$ are the k th columns of the matrix $\tilde{\mathbf{Z}}[n] = \mathbf{Z}[n] \Phi^H$. Since $\Phi \Phi^H = \mathbf{I}_K$, $\tilde{\mathbf{z}}_k[n] \sim \mathcal{CN}(\mathbf{0}, \sigma_r^2 \mathbf{I}_M)$.

Exploiting the MMSE criterion [28], the estimate of $\mathbf{g}_k[n]$, $\hat{\mathbf{g}}_k[n]$ is distributed as

$$\hat{\mathbf{g}}_k[n] \sim \mathcal{CN}(\mathbf{0}, \hat{\beta}_k \mathbf{I}_M), \quad (11)$$

where $\hat{\beta}_k = \frac{\tau_r p_p \beta_k^2}{\sigma_r^2 + \tau_r p_p \beta_k}$.

Due to the orthogonality property of MMSE estimation, $\mathbf{g}_k[n]$ can be decomposed into

$$\mathbf{g}_k[n] = \hat{\mathbf{g}}_k[n] + \tilde{\mathbf{g}}_k[n], \quad (12)$$

where $\tilde{\mathbf{g}}_k[n] \sim \mathcal{CN}(\mathbf{0}, (\beta_k - \hat{\beta}_k) \mathbf{I}_M)$ is the CE error and is uncorrelated with $\hat{\mathbf{g}}_k[n]$.

3.3 Channel aging

To analyze the impact of channel aging, in our analysis, we adopt an autoregressive model of order 1 for approximating the temporally correlated fading channel coefficient. As such, the channel vectors for the k -th UE at time $n+1$ can be expressed as [25]

$$\mathbf{g}_k[n+1] = \alpha \mathbf{g}_k[n] + \boldsymbol{\epsilon}_k[n+1], \quad (13)$$

where $\boldsymbol{\epsilon}_k[n+1] \sim \mathcal{CN}(\mathbf{0}, (1 - \alpha^2) \beta_k \mathbf{I})$ is a temporally uncorrelated complex Gaussian random process. We denote $\alpha = J_0(2\pi f_D T_S)$ as a temporal correlation parameter, where $J_0(\cdot)$ is the zero-order first-kind Bessel function. T_S is the channel sampling duration and $f_D = \frac{v f_c}{c}$ is the maximum Doppler frequency shift, where v , f_c , and c are the UEs' velocity, carrier frequency, and the speed of light, respectively. Without loss of generality, we assume that all users move with the same velocity. As a result, the time variation does not depend on the user index. While this seems not realistic, we stay very near to the practical case by considering the worst-case scenario where we set all users with the velocity corresponding to the most varying user.

To this end, a model accounting for the combined effects of the CE error and channel aging effect can be expressed as

$$\begin{aligned} \mathbf{g}_k[n+1] &= \alpha \hat{\mathbf{g}}_k[n] + \underbrace{\alpha \tilde{\mathbf{g}}_k[n]}_{\boldsymbol{\xi}_{a,k}[n+1]} + \boldsymbol{\epsilon}_k[n+1] \\ &= \tilde{\mathbf{g}}_{a,k}[n+1] + \boldsymbol{\xi}_{a,k}[n+1], \end{aligned} \quad (14)$$

where $\boldsymbol{\xi}_{a,k}[n+1] \sim \mathcal{CN}(\mathbf{0}, \tilde{\beta}_{a,k} \mathbf{I}_M)$ is mutually independent of $\tilde{\mathbf{g}}_{a,k}[n+1] \sim \mathcal{CN}(\mathbf{0}, \tilde{\beta}_{a,k} \mathbf{I}_M)$ with $\tilde{\beta}_{a,k} = \alpha^2 \hat{\beta}_k$, $\tilde{\beta}_{a,k} = \beta_k - \alpha^2 \hat{\beta}_k$. Obviously, the combined error $\boldsymbol{\xi}_{a,k}[n+1]$ consists of both the CE error and aged CSI effects.

3.4 Channel prediction

Channel prediction is an important approach to alleviate the channel aging effect [29–31]. In this subsection, we focus on predicting $\mathbf{g}_k[n+1]$ based on the current and previous received training signals. The detailed procedure for predicting $\mathbf{g}_k[n+1]$ is given as follows.

We adopt a Wiener predictor. Then, $\mathbf{g}_k[n+1]$ is predicted according to $\bar{\mathbf{x}}_k[n]$, where

$$\bar{\mathbf{x}}_k[n] = [\tilde{\mathbf{x}}_k^T[n], \tilde{\mathbf{x}}_k^T[n-1], \dots, \tilde{\mathbf{x}}_k^T[n-p]]^T, \quad (15)$$

with p being the predictor order. The predicted CSI is provided as follows.

$$\bar{\mathbf{g}}_{p,k}[n+1] = \mathbf{V}_k \bar{\mathbf{x}}_k[n], \quad (16)$$

where the optimal p -th linear Wiener predictor is given as [31]

$$\mathbf{V}_k = \alpha \beta_k [\boldsymbol{\delta}(p, \alpha) \otimes \mathbf{I}_M] \mathbf{T}_k^{-1}(p, \alpha), \quad (17)$$

Specifically, we have $\boldsymbol{\delta}(p, \alpha) = [1, \alpha, \dots, \alpha^p]$ and

$$\mathbf{T}_k(p, \alpha) = \beta_k [\boldsymbol{\Delta}(p, \alpha) \otimes \mathbf{I}_M] + \frac{1}{\tau_r p_p} \mathbf{I}_{M(p+1)} \quad (18)$$

with

$$\boldsymbol{\Delta}(p, \alpha) \triangleq \begin{bmatrix} 1 & \alpha & \dots & \alpha^p \\ \alpha & 1 & \dots & \alpha^{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^p & \alpha^{p-1} & \dots & 1 \end{bmatrix}. \quad (19)$$

According to [30], the covariance matrix of $\bar{\mathbf{g}}_{p,k}[n+1]$ is given by $\alpha^2 \boldsymbol{\Theta}_k(p, \alpha)$, where

$$\boldsymbol{\Theta}_k(p, \alpha) \triangleq \beta_k^2 [\boldsymbol{\delta}(p, \alpha) \otimes \mathbf{I}_M] \mathbf{T}_k^{-1}(p, \alpha) [\boldsymbol{\delta}(p, \alpha) \otimes \mathbf{I}_M]. \quad (20)$$

Thus, the real channel can be decomposed as [31]

$$\mathbf{g}_k[n+1] = \bar{\mathbf{g}}_{p,k}[n+1] + \boldsymbol{\xi}_{p,k}[n+1], \quad (21)$$

where $\boldsymbol{\xi}_{p,k}[n+1]$ is the channel prediction error vector with covariance matrix $\beta_k \mathbf{I}_M - \alpha^2 \boldsymbol{\Theta}_k(p, \alpha)$, which is independent of $\bar{\mathbf{g}}_{p,k}[n+1]$. According to [29], it can be obtained that

$$\begin{aligned} \bar{\mathbf{g}}_{p,k}[n+1] &\sim \mathcal{CN}(\mathbf{0}, \alpha^2 \boldsymbol{\Theta}_k(p, \alpha)), \\ \boldsymbol{\xi}_{p,k}[n+1] &\sim \mathcal{CN}(\mathbf{0}, \beta_k \mathbf{I}_M - \alpha^2 \boldsymbol{\Theta}_k(p, \alpha)). \end{aligned} \quad (22)$$

According to the result in ([31], Lemma 2), $\boldsymbol{\Theta}_k(p, \alpha)$ is a scaled identity matrix of size $M \times M$, which can be straightforwardly shown as follows.

$$\begin{aligned} \bar{\mathbf{g}}_{p,k}[n+1] &\sim \mathcal{CN}(\mathbf{0}, \bar{\beta}_{p,k} \mathbf{I}_M), \\ \boldsymbol{\xi}_{p,k}[n+1] &\sim \mathcal{CN}(\mathbf{0}, \tilde{\beta}_{p,k} \mathbf{I}_M), \end{aligned} \quad (23)$$

with $\bar{\beta}_{p,k} = \frac{1}{M} \text{tr}(\alpha^2 \boldsymbol{\Theta}_k(p, \alpha))$, $\tilde{\beta}_{p,k} = \frac{1}{M} \text{tr}(\beta_k \mathbf{I}_M - \alpha^2 \boldsymbol{\Theta}_k(p, \alpha))$.

4 Performance analysis for achievable sum rate

In this section, we consider two different scenarios, i.e., aged and predicted CSI. We first provide a unified achievable SR expression for two scenarios. Next, we derive closed-form expressions for achievable SR under aged and predicted CSI scenarios, which are desirable for the subsequent energy-efficient optimization problem formulation.

We assume that the temporal correlation parameter α and the LSF channel matrix \mathbf{D} are known a priori at the relay. Hence, we can have the following CSI

$$\bar{\mathbf{g}}_k[n+1] = \begin{cases} \bar{\mathbf{g}}_{a,k}[n+1], & \text{for aged CSI} \\ \bar{\mathbf{g}}_{p,k}[n+1], & \text{for predicted CSI} \end{cases} \quad (24)$$

Then, the ZF-receive matrix $\mathbf{W}_1[n+1]$ and ZF-transmit matrix $\mathbf{W}_2[n+1]$ are respectively expressed as

$$\mathbf{W}_1[n+1] = (\bar{\mathbf{G}}^H[n+1] \bar{\mathbf{G}}[n+1])^{-1} \bar{\mathbf{G}}^H[n+1], \quad (25)$$

$$\mathbf{W}_2[n+1] = \bar{\mathbf{G}}^*[n+1] (\bar{\mathbf{G}}^T[n+1] \bar{\mathbf{G}}^*[n+1])^{-1},$$

where $\bar{\mathbf{G}}[n+1] \triangleq [\bar{\mathbf{g}}_1[n+1], \dots, \bar{\mathbf{g}}_K[n+1]]$.

After the imperfect self-interference cancellation (SIC), the received signal at the k th UE of the j th time slot of the broadcast phase can be rewritten as

$$\begin{aligned} y_{u,k}^{(j)}[n+1] &= \underbrace{\vartheta_j \sqrt{p_k} \mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1] s_{k'}[n+1]}_{\text{desired signal}} \\ &+ \underbrace{\vartheta_j \sqrt{p_k} \lambda_k s_k[n+1]}_{\text{residual self-interference}} \\ &+ \underbrace{\vartheta_j \sum_{i \neq k, k'}^K \sqrt{p_i} \mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_i[n+1] s_i[n+1]}_{\text{inter-user interference}} \\ &+ \underbrace{\vartheta_j \mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{n}_R[n+1]}_{\text{noise at the relay}} + \underbrace{n_{u,k}[n+1]}_{\text{noise at the user}}, \end{aligned} \quad (26)$$

where $\lambda_k = \mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_k[n+1] - \bar{\mathbf{g}}_k^T[n+1] \mathbf{F}_j[n+1] \bar{\mathbf{g}}_k[n+1]$ is the SIC coefficient for the k th UE.

From (26), the ergodic achievable rate of the k th UE in the j th time slot of the broadcast phase can be expressed as

$$\mathcal{R}_k^{(j)} = \mathbb{E} \left[\log_2 \left(1 + \gamma_k^{(j)} \right) \right] \quad (27)$$

where

$$\gamma_k^{(j)} = \frac{\text{DS}_k^{(j)}}{\text{RSI}_k^{(j)} + \text{IUI}_k^{(j)} + \text{NR}_k^{(j)} + \text{NU}_k}, \quad (28)$$

with $\text{DS}_k^{(j)} \triangleq \vartheta_j^2 p_{k'} |\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1]|^2$, $\text{RSI}_k^{(j)} \triangleq \vartheta_j^2 p_k |\lambda_k|^2$, $\text{NU}_k \triangleq |n_{u,k}[n+1]|^2$, $\text{IUI}_k^{(j)} \triangleq \vartheta_j^2 p_i |\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_i[n+1]|$ and $\text{NR}_k^{(j)} \triangleq \vartheta_j^2 |\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{n}_R[n+1]|$.

However, further derivation of (27) is difficult because of the intractability to carry out the ensemble average analytically. Instead, we adopt another technique to derive a worst-case lower bound of achievable rate. According to [32], we can rewrite $y_{u,k}^{(j)}[n+1]$ as

$$y_{u,k}^{(j)}[n+1] = \vartheta_j \sqrt{p_{k'}} \mathbb{E} \left[\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1] \right] s_{k'}[n+1] + \tilde{n}_k[n+1] \quad (29)$$

with

$$\begin{aligned} \tilde{n}_k[n+1] &= \vartheta_j \sqrt{p_{k'}} \left(\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1] \right. \\ &\quad \left. - \mathbb{E} \left[\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1] \right] \right) s_{k'}[n+1] \\ &+ \vartheta_j \sqrt{p_k} \lambda_k s_k[n+1] \\ &+ \vartheta_j \sum_{i \neq k, k'}^K \sqrt{p_i} \mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_i[n+1] s_i[n+1] \\ &+ \vartheta_j \mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{n}_R[n+1] + n_{u,k}[n+1]. \end{aligned} \quad (30)$$

In (29), the first part $\vartheta_j \sqrt{p_{k'}} \mathbb{E} [\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1]] s_{k'}[n+1]$ is considered as “desired signal,” and the second term $\tilde{n}_k[n+1]$ is considered as “effective noise,” uncorrelated with the first term. Therefore, by approximating the effective noise as independent Gaussian noise of the same variance [32], we can obtain the statistical CSI based achievable rate of the k th UE in the j th time slot of the broadcast phase as

$$\widehat{\mathcal{R}}_k^{(j)} = \log_2 \left(1 + \widehat{\gamma}_k^{(j)} \right), \quad (31)$$

with

$$\widehat{\gamma}_k^{(j)} = \frac{\vartheta_j^2 p_{k'} |\mathbb{E} [\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1]]|^2}{\vartheta_j^2 p_{k'} \text{Var} [\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1]] + \vartheta_j^2 \text{SI}_k^{(j)} + \vartheta_j^2 \text{UI}_k^{(j)} + \vartheta_j^2 \text{NR}_k^{(j)} + \text{NU}_k}, \quad (32)$$

where $\text{SI}_k^{(j)}$, $\text{UI}_k^{(j)}$, $\text{NR}_k^{(j)}$, and NU_k denote the residual self-interference after SIC, the inter-user interference, the amplified noise from the relay, and the noise at k th UE, respectively, i.e.,

$$\text{SI}_k^{(j)} \triangleq p_k \mathbb{E} [|\lambda_k|^2], \quad (33)$$

$$\text{UI}_k^{(j)} \triangleq \sum_{i \neq k, k'}^K p_i \mathbb{E} \left[\left| \mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_i[n+1] \right|^2 \right],$$

$$\text{NR}_k^{(j)} \triangleq \mathbb{E} \left[\left| \mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{n}_R[n+1] \right|^2 \right],$$

$$\text{NU}_k \triangleq \mathbb{E} [|n_{u,k}[n+1]|^2].$$

Remark 1 The above worst-case lower bound of achievable rate in (31) is obtained by assuming that UE_k uses only statistical information of the channel gains (i.e.,

$\mathbb{E}[\mathbf{g}_k^T[n+1]\mathbf{F}_j[n+1]\mathbf{g}_{k'}[n+1]]$ to decode the signal transmitted by UE $_{k'}$. By contrast, the ergodic rate in (27) is obtained by a sophisticated receiver, i.e., UE $_{k'}$ knows perfectly $\mathbf{g}_k^T[n+1]\mathbf{F}_j[n+1]\mathbf{g}_{k'}[n+1]$. In Section 8, it is demonstrated via simulations that the performance gap between the achievable SRs given by (31) and (27) is rather small in massive MIMO-enabled MWRNs. It is clear that (31) is a very useful metric for obtaining the achievable rate in practical applications where CSI is not available.

Accordingly, the statistical CSI-based achievable SR of the considered system with aged and predicted CSI is uniformly given as

$$\widehat{\mathcal{R}} = \sum_{k=1}^K \sum_{j=1}^{K-1} \widehat{\mathcal{R}}_k^{(j)}, \quad (34)$$

In the following theorems, two accurate closed-form expressions for the worst-case lower bound of achievable SR are derived under aged and predicted CSI scenarios.

Theorem 1 For ZF transceivers, with aged CSI, the worst-case lower bound of achievable SR in the considered massive MIMO-enabled MWRNs is given by

$$\widehat{\mathcal{R}}_a = \sum_{k=1}^K \sum_{j=1}^{K-1} \widehat{\mathcal{R}}_{ak}^{(j)}, \quad (35)$$

where the closed-form formula of $\widehat{\mathcal{R}}_{ak}^{(j)}$ is defined as

$$\widehat{\mathcal{R}}_{ak}^{(j)} = \log_2 \left(1 + \frac{p_{k'} p_r}{p_r (\zeta_{a1,k}^{(j)} + \zeta_{a2,k}^{(j)} + \zeta_{a3,k}^{(j)}) + \zeta_{a4,k}^{(j)} \zeta_a^{(j)}} \right), \quad (36)$$

in which

$$\begin{aligned} \zeta_{a1,k}^{(j)} &= p_{k'} \left(\mu_1 \tilde{\beta}_{a,k'} \tilde{\beta}_{a,k'}^{-1} + \mu_1 \tilde{\beta}_{a,k} \tilde{\beta}_{a,k}^{-1} + \mu_2 \tilde{\beta}_{a,k} \tilde{\beta}_{a,k'} \mathcal{A}_{aj} \right), \\ \zeta_{a2,k}^{(j)} &= p_k \left(\mu_1 \tilde{\beta}_{a,k} \tilde{\beta}_{a,k'}^{-1} + \mu_1 \tilde{\beta}_{a,k} \tilde{\beta}_{a,k''}^{-1} + \mu_2 \tilde{\beta}_{a,k}^2 \mathcal{A}_{aj} \right), \\ \zeta_{a3,k}^{(j)} &= \sum_{i \neq k, k'}^K p_i \left(\mu_1 \tilde{\beta}_{a,i} \tilde{\beta}_{a,k'}^{-1} + \mu_1 \tilde{\beta}_{a,i} \tilde{\beta}_{a,i''}^{-1} + \mu_2 \tilde{\beta}_{a,i} \tilde{\beta}_{a,k} \mathcal{A}_{aj} \right), \\ \zeta_{a4,k}^{(j)} &= \mu_1 \sigma_r^2 \tilde{\beta}_{a,k'}^{-1} + \sigma_r^2 \tilde{\beta}_{a,k} \mathcal{A}_{aj} + \sigma_u^2, \\ \zeta_a^{(j)} &= \sum_{k=1}^K p_k \left(\mu_1 \tilde{\beta}_{a,k''}^{-1} + \mu_2 \tilde{\beta}_{a,k} \mathcal{A}_{aj} \right) + \mu_2 \sigma_r^2 \mathcal{A}_{aj}, \end{aligned} \quad (37)$$

where $k'' = \text{mod}_K(K + k - j)$, $i'' = \text{mod}_K(K + i - j)$, $\mu_1 = \frac{1}{M-K-1}$, $\mu_2 = \frac{(2+(M-K)(M-K-3))}{(M-K)(M-K-1)^2(M-K-3)}$ and $\mathcal{A}_{aj} = \sum_{k=1}^K \tilde{\beta}_{a,k}^{-1} \tilde{\beta}_{a,k''}$.

Proof Please see Appendix 1. \square

For predicted CSI, a closed-form expression for the statistical CSI-based achievable SR is derived as follows.

Theorem 2 For ZF transceivers, with predicted CSI, the worst-case lower bound of achievable SR in the considered massive MIMO-enabled MWRNs is given by

$$\widehat{\mathcal{R}}_p = \sum_{k=1}^K \sum_{j=1}^{K-1} \widehat{\mathcal{R}}_{pk}^{(j)}, \quad (38)$$

where $\widehat{\mathcal{R}}_{pk}^{(j)}$ is derived as

$$\widehat{\mathcal{R}}_{pk}^{(j)} = \log_2 \left(1 + \frac{p_k p_r}{p_r (\zeta_{p1,k}^{(j)} + \zeta_{p2,k}^{(j)} + \zeta_{p3,k}^{(j)}) + \zeta_{p4,k}^{(j)} \zeta_p^{(j)}} \right), \quad (39)$$

in which

$$\begin{aligned} \zeta_{p1,k}^{(j)} &= p_{k'} \left(\mu_1 \tilde{\beta}_{p,k'} \tilde{\beta}_{p,k'}^{-1} + \mu_1 \tilde{\beta}_{p,k} \tilde{\beta}_{p,k}^{-1} + \mu_2 \tilde{\beta}_{p,k} \tilde{\beta}_{p,k'} \mathcal{A}_{pj} \right), \\ \zeta_{p2,k}^{(j)} &= p_k \left(\mu_1 \tilde{\beta}_{p,k} \tilde{\beta}_{p,k'}^{-1} + \mu_1 \tilde{\beta}_{p,k} \tilde{\beta}_{p,k''}^{-1} + \mu_2 \tilde{\beta}_{p,k}^2 \mathcal{A}_{pj} \right), \\ \zeta_{p3,k}^{(j)} &= \sum_{i \neq k, k'}^K p_i \left(\mu_1 \tilde{\beta}_{p,i} \tilde{\beta}_{p,k'}^{-1} + \mu_1 \tilde{\beta}_{p,i} \tilde{\beta}_{p,i''}^{-1} + \mu_2 \tilde{\beta}_{p,i} \tilde{\beta}_{p,k} \mathcal{A}_{pj} \right), \\ \zeta_{p4,k}^{(j)} &= \mu_1 \sigma_r^2 \tilde{\beta}_{p,k'}^{-1} + \sigma_r^2 \tilde{\beta}_{p,k} \mathcal{A}_{pj} + \sigma_u^2, \\ \zeta_p^{(j)} &= \sum_{k=1}^K p_k \left(\mu_1 \tilde{\beta}_{p,k''}^{-1} + \mu_2 \tilde{\beta}_{p,k} \mathcal{A}_{pj} \right) + \mu_2 \sigma_r^2 \mathcal{A}_{pj}. \end{aligned} \quad (40)$$

where $\mathcal{A}_{pj} = \sum_{k=1}^K \tilde{\beta}_{p,k}^{-1} \tilde{\beta}_{p,k''}$.

Proof Since the proof follows similar lines as the proof of Theorem 1, it is omitted. \square

Remark 2 Through Theorems 1 and 2, two simple closed-form expression of the SR of the considered system have been derived. The advantage of these expressions are that it only depends on the LSF channel coefficients and the configurable system parameters. Thus, complicated calculations involving large-dimensional matrix variables that represent the SSF channel coefficients are avoided. In this way, the computational complexity which relates to the SSF-based signal processing is greatly reduced. It is underlined that these closed-form expressions establish an explicit functional relationship between the SR, the transmit powers of UEs, thus facilitating the introduction of the following novel energy-efficient resource allocation methodology.

5 EE optimization problem formulation

In the EE optimization, we employ a realistic power consumption model similar to those used in [11, 33]. The

total power consumption of the considered system can be quantified as

$$P_{\text{tot}} = P_{\text{PA}} + P_{\text{C}}, \quad (41)$$

where P_{PA} is the power consumed by power amplifiers (PAs) given as

$$P_{\text{PA}} = \frac{(1 - \frac{\tau_r}{T}) \sum_{k=1}^K p_k}{\eta_{\text{PA,U}}} + \frac{\frac{\tau_r}{T} K p_{\text{p}}}{\eta_{\text{PA,U}}} + \frac{(1 - \frac{\tau_r}{T})(K-1)p_r}{\eta_{\text{PA,R}}}, \quad (42)$$

in which $\eta_{\text{PA,U}} \in (0, 1)$ and $\eta_{\text{PA,R}} \in (0, 1)$ are the efficiency of PAs at the UEs and at the relay, respectively. P_{C} denotes the total circuit power consumption, to be more specific, we have

$$P_{\text{C}} = K P_{\text{CU}} + P_{\text{CE}} + P_{\text{LP}} + M(P_{\text{RR}} + P_{\text{cR}}) + P_{\text{f}}, \quad (43)$$

in which $P_{\text{CE}} = M \frac{\log_2(\tau_r) R_{\text{flops}}}{\eta_{\text{C}}}$ is the power consumed for CE at the relay, R_{flops} is the floating-point operations per second (flops) per antenna for each user, and η_{C} is the power efficiency of computing measured in *flops/W*.

$P_{\text{LP}} = 2M \frac{(K+K^2)R_{\text{flops}}}{\eta_{\text{C}}}$ is the power consumed for the ZF-receive detector and ZF-transmit precoder. P_{RR} is the other baseband processing power (such as ADC/DAC, modulation/demodulation) at each antenna, P_{cR} and P_{CU} are the power consumed at circuit components of each antenna at the relay and each UE, respectively, and P_{f} is the fixed power consumption at the relay.

The power consumption in (41) can be rewritten as

$$P_{\text{tot}} \triangleq v_1 \sum_{k=1}^K p_k + P_{\text{fixed}}, \quad (44)$$

where $v_1 \triangleq \frac{(1 - \frac{\tau_r}{T})}{\eta_{\text{PA,U}}}$, $P_{\text{fixed}} = \frac{\frac{\tau_r}{T} K p_{\text{p}}}{\eta_{\text{PA,U}}} + \frac{(1 - \frac{\tau_r}{T})(K-1)p_r}{\eta_{\text{PA,R}}} + K P_{\text{CU}} + P_{\text{CE}} + P_{\text{LP}} + M(P_{\text{RR}} + P_{\text{cR}}) + P_{\text{f}}$.

Given the values of the other system parameters, the EE η_{EE} [bits/Joule] under aged and predicted CSI scenarios is unifiedly defined as

$$\eta_{\text{EE}} = \frac{B(1 - \frac{\tau_r}{T}) \frac{K-1}{K} \sum_{k=1}^K \sum_{j=1}^{K-1} \widehat{\mathcal{R}}_k^{(j)}}{P_{\text{tot}}}, \quad (45)$$

where

$$\widehat{\mathcal{R}}_k^{(j)} = \log_2 \left(1 + \frac{p_r p_{k'}}{\sum_{i=1}^K p_i \rho_{k,i}^{(j)} + \mu_k^j} \right), \quad (46)$$

with $\rho_{k,i}^{(j)}$, $\mu_k^{(j)}$ are constant value (independent of transmit powers), which are different for aged CSI and predicted CSI. More precisely,

- For aged CSI, $\rho_{k,i}^{(j)} = p_r \left(\mu_1 \tilde{\beta}_{a,i} \tilde{\beta}_{a,k'}^{-1} + \mu_1 \tilde{\beta}_{a,k} \tilde{\beta}_{a,i'}^{-1} + \mu_2 \tilde{\beta}_{a,i} \tilde{\beta}_{a,k} \mathcal{A}_{aj} \right) + \sigma_u^2 \left(\mu_1 \tilde{\beta}_{a,i'}^{-1} + \mu_2 \tilde{\beta}_{a,i} \mathcal{A}_{aj} \right)$,
 $\mu_k^{(j)} = p_r \left(\mu_1 \sigma_r^2 \tilde{\beta}_{a,k'}^{-1} + \mu_2 \sigma_r^2 \tilde{\beta}_{a,k} \mathcal{A}_{aj} \right) + \mu_2 \sigma_u^2 \sigma_r^2 \mathcal{A}_{aj}$.

- For predicted CSI, $\rho_{k,i}^{(j)} = p_r \left(\mu_1 \tilde{\beta}_{p,i} \tilde{\beta}_{p,k'}^{-1} + \mu_1 \tilde{\beta}_{p,k} \tilde{\beta}_{p,i'}^{-1} + \mu_2 \tilde{\beta}_{p,i} \tilde{\beta}_{p,k} \mathcal{A}_{pj} \right) + \sigma_u^2 \left(\mu_1 \tilde{\beta}_{p,i'}^{-1} + \mu_2 \tilde{\beta}_{p,i} \mathcal{A}_{pj} \right)$,
 $\mu_k^{(j)} = p_r \left(\mu_1 \sigma_r^2 \tilde{\beta}_{p,k'}^{-1} + \mu_2 \sigma_r^2 \tilde{\beta}_{p,k} \mathcal{A}_{pj} \right) + \mu_2 \sigma_u^2 \sigma_r^2 \mathcal{A}_{pj}$.

Remark 3 In (45), the pre-log factor $(1 - \frac{\tau_r}{T})$ is due to the fact that during each coherence interval of T symbols, we spend τ_r symbols for pilot-based CE. Moreover, the numerator $K-1$ of the pre-log factor $\frac{K-1}{K}$ is due to the fact that any user node receives signals from other $K-1$ user nodes, while the denominator K follows by the single time slot in the multiple-access phase and $K-1$ time slots used for full-data exchange in the broadcast phase.

It is seen from (45) that the EE η_{EE} is a function of the transmit powers of K UEs, $\{p_k\}_{k=1}^K$. How to wisely allocate the transmit power among the K UEs is crucial for achieving the optimum EE in the context of green communications. Hence, the energy-efficient power allocation is formulated as the following optimization problem:

$$\begin{aligned} & \max_{\{p_k\}_{k=1}^K} \eta_{\text{EE}} \\ & \text{s. t. } \begin{cases} \text{C1: } p_k \geq 0, \forall k \\ \text{C2: } p_k \leq p_{\text{max}}, \forall k \\ \text{C3: } \widehat{\mathcal{R}}_k^{(j)} \geq R_0, \forall k, j \end{cases} \end{aligned} \quad (47)$$

where the objective function η_{EE} is defined by (45), and p_{max} is the maximum transmit power of each UE. The constraints C1 and C2 are the boundary values for the transmit powers of K UEs. The constraint C3 guarantees the transmission link quality by satisfying the minimum QoS requirement R_0 for each UE at each time slot of the broadcast phase. Here, R_0 denotes the required achievable rate for all UEs.

6 Energy-efficient power allocation algorithm

6.1 Optimal power allocation (OPA) scheme

It is easy to observe that (47) is not a CFP optimization problem, because the numerator of the objection function η_{EE} and the QoS constraints C3 are non-convex with respect to $\{p_k\}_{k=1}^K$. Therefore, (47) cannot be directly solved by classic fractional programming tools. To overcome this difficulty, we employ the SCA technique proposed in [34–36] to sequentially approximate $\mathcal{R}_k^{(j)}$ by using the following inequality:

$$\log_2(1 + z_{k,j}) \geq a_{k,j} \log_2 z_{k,j} + b_{k,j}. \quad (48)$$

The above inequation is tight at a particular value $z_{k,j} = \bar{z}_{k,j}$ when the approximation constants $a_{k,j}$ and $b_{k,j}$ are chosen as

$$a_{k,j} = \frac{\bar{z}_{k,j}}{1 + \bar{z}_{k,j}}, b_{k,j} = \log_2(1 + \bar{z}_{k,j}) - \frac{\bar{z}_{k,j}}{1 + \bar{z}_{k,j}} \log_2 \bar{z}_{k,j}. \quad (49)$$

Motivated by the above convexity approximation, we employ the inequality (48) to approximate $\widehat{\mathcal{R}}_k^{(j)}$, where $z_{k,j}$ corresponds to $\widehat{\gamma}_{k,j} = \frac{p_i p_{k'}}{\sum_{i=1}^K p_i \rho_{k,i}^{(j)} + \mu_k}$. Then, the variable change $p_k = 2^{q_k}$, for $\forall k$ was used. Finally, we arrive at the following approximated optimization problem

$$\begin{aligned} \max_{\{q_k\}_{k=1}^K} \widetilde{\eta}_{EE}(\mathbf{Q}) &= \frac{B \left(1 - \frac{\tau_r}{T}\right) \frac{K-1}{K} \sum_{k=1}^K \sum_{j=1}^{K-1} \widetilde{\mathcal{R}}_k^{(j)}(\mathbf{Q})}{P_{\text{tot}}(\mathbf{Q})} \\ \text{s. t. } &\begin{cases} \text{C1}' : 2^{q_k} \geq 0, \forall k \\ \text{C2}' : 2^{q_k} \leq p_{\text{max}}, \forall k \\ \text{C3}' : \widetilde{\mathcal{R}}_k^{(j)}(\mathbf{Q}) \geq R_0, \forall k, j \end{cases} \end{aligned} \quad (50)$$

where $\mathbf{Q} = \text{diag}\{q_1 \cdots, q_k \cdots, q_K\}$, $P_{\text{tot}}(\mathbf{Q}) = v_1 \sum_{k=1}^K 2^{q_k} + P_{\text{fixed}}$ and

$$\begin{aligned} \widetilde{\mathcal{R}}_k^{(j)}(\mathbf{Q}) &= b_{k,j} + a_{k,j} \log_2(p_r) \\ &\quad + a_{k,j} q_{k'} - a_{k,j} \log_2 \left(\sum_{i=1}^K 2^{q_i} \rho_{k,i}^{(j)} + \mu_k \right), \end{aligned} \quad (51)$$

with $a_{k,j} = \frac{\bar{\gamma}_k^{(j)}}{1 + \bar{\gamma}_k^{(j)}}$ and $b_{k,j} = \log_2 \left(1 + \bar{\gamma}_k^{(j)} \right) - \frac{\bar{\gamma}_k^{(j)}}{1 + \bar{\gamma}_k^{(j)}}$. $\log_2 \bar{\gamma}_k^{(j)}$ being the approximation constants computed as (49), where $\bar{\gamma}_k^{(j)} = \frac{p_r 2^{q_{k'}}}{\sum_{i=1}^K 2^{q_i} \rho_{k,i}^{(j)} + \mu_k}$. For any fixed $a_{k,j}$ and $b_{k,j}$, it can be easily verified that (51) is convex with respect to $\{q_k\}_{k=1}^K$.⁴ Therefore, the optimization problem (50) is a CFP problem with a quasi-concave objective function $\widetilde{\eta}_{EE}(\mathbf{Q})$ ⁵ and convex constraints, which can be transformed into a convex optimization in a subtractive by the Dinkelbach's method as follows [37].

$$\begin{aligned} \max_{\mathbf{Q}} \mathcal{F}(\mathbf{Q}, \lambda) \\ \text{s. t. C1}' - \text{C3}', \end{aligned} \quad (52)$$

where

$$\mathcal{F}(\mathbf{Q}, \lambda) = B \left(1 - \frac{\tau_r}{T}\right) \frac{K-1}{K} \sum_{k=1}^K \sum_{j=1}^{K-1} \widetilde{\mathcal{R}}_k^{(j)}(\mathbf{Q}) - \lambda P_{\text{tot}}(\mathbf{Q}). \quad (53)$$

Here, λ is a non-negative parameter, it can be noted that when $\lambda \rightarrow 0$, it implies that the energy-efficient problem (52) is degenerated to an optimization problem for the SE maximization. The optimal factor λ^* (i.e., the optimal objective function value of (50)) works as the optimal EE for the system. For fixed parameters $a_{k,j}$, $b_{k,j}$, and λ , the optimization problem (52) is a convex optimization problem, which can be efficiently solved using standard convex

optimization tools, e.g., CVX [38]. Next, we derive an iterative algorithm for solving this optimization by applying the Lagrangian dual method.

Thus, the dual problem associated with the primal problem (52) can be written as

$$\begin{aligned} \mathcal{L}(\mathbf{Q}, \lambda, \boldsymbol{\mu}, \boldsymbol{\psi}) &= B \left(1 - \frac{\tau_r}{T}\right) \frac{K-1}{K} \sum_{k=1}^K \sum_{j=1}^{K-1} \widetilde{\mathcal{R}}_k^{(j)}(\mathbf{Q}) \\ &\quad - \lambda \left(v_1 \sum_{k=1}^K 2^{q_k} + P_{\text{fixed}} \right) \\ &\quad - \sum_{k=1}^K \mu_k (2^{q_k} - p_{\text{max}}) \\ &\quad + \sum_{k=1}^K \sum_{j=1}^{K-1} \psi_{k,j} \left(\widetilde{\mathcal{R}}_k^{(j)} - R_0 \right) \end{aligned} \quad (54)$$

where $\boldsymbol{\mu} = \{\mu_k\}$, $\forall k$ are the Lagrangian multipliers associated with the transmit power constraints C2'. while $\boldsymbol{\psi} = \{\psi_{k,j}\}$, $\forall k, j$ are the Lagrangian multipliers for QoS constraints C3'.

In the following, we solve the dual problem (54) using Lagrangian dual approach, which alternates between a *subproblem (inner problem)*, updating the power allocation variables \mathbf{Q} by fixing the Lagrangian multipliers $\boldsymbol{\mu}$, $\boldsymbol{\psi}$, and a *master problem (outer problem)*, updating the Lagrangian multipliers $\boldsymbol{\mu}$, $\boldsymbol{\psi}$ for the obtained solution of the inner problem \mathbf{Q}^* . The Lagrangian dual approach is outlined as follows.

The optimization problem (54) is in a standard concave form, which can be efficiently solved by using standard optimization techniques and KKT conditions [38]. Thus, to obtain the optimal power allocation for users, we take the partial derivative of (54) with q_k , $k = 1, \dots, K$, and equate the results to zero, thus the power allocation at the $(m+1)$ th iteration is updated as follows.

$$q_k(m+1) = \log_2 \left[\frac{\sum_{k=1}^K \sum_{j=1}^{K-1} \left(B \left(1 - \frac{\tau_r}{T}\right) \frac{K-1}{K} + \psi_{k,j} \right) \frac{\rho_{k,k}^{(j) a_{k,j}}}{\sum_{i=1}^K 2^{q_i(m)} \rho_{k,i}^{(j)} + \mu_k}}{(\mu_k + \lambda v_1) \ln 2} \right]^+ \quad (55)$$

where $[x]^+ = \max\{0, x\}$.

Since the dual problem in (54) is differentiable, the gradient method may be readily used for updating the Lagrangian dual variables μ_k and $\psi_{k,j}$, $\forall k, j$ as follows [39].

$$\mu_k(m+1) = [\mu_k(m) - \epsilon_1(m) (p_{\text{max}} - 2^{q_k})]^+, \forall k, \quad (56)$$

$$\psi_{k,j}(m+1) = [\psi_{k,j}(m) - \epsilon_2(m) (\widetilde{\mathcal{R}}_k^{(j)} - R_0)]^+, \forall k, j, \quad (57)$$

where $\epsilon_1(m)$ and $\epsilon_2(m)$ are the step sizes used for moving in the direction of the negative gradient for the Lagrangian multipliers μ_k and $\psi_{k,j}$, respectively. The

updated Lagrange multipliers are used for updating the power allocation policy. We repeat this process until convergence. The detailed iterative procedure is summarized in Algorithm 1.

Algorithm 1 Iteration algorithm for optimizing the transmit powers of the users

1. Set the maximum number of iterations $I_{\max 1}$;
2. Initialize the iteration counter $l = 0$ and penalty factor $\lambda(l) = 0.01$;
3. **repeat** (Corresponding to the outer loop)
 - (a) Set the maximum number of iteration $l_{\max 2}$;
 - (b) Initialize the iteration counter $m = 0$ and the step $\epsilon_1(m)$, $\epsilon_2(m)$, $(a_{k,j}, b_{k,j})$ by $(1, 0)$, $\mu_k(m)$, $\psi_{k,j}(m)$ and $q_k(m)$, $\forall k, j$;
- repeat** (Corresponding to the inner loop) **repeat** (Solving optimization problem (54))
 - (a) Update $q_k(m)$ using (55);
 - (b) Update $\mu_k(m)$ and $\psi_{k,j}(m)$ using (56) and (57), respectively;
- Until** convergence to the optimal solution q_k^* , $\forall k$
4. Update the coefficients $(a_{k,j}, b_{k,j})$
5. Set $q_k(m+1) \leftarrow q_k^*$, $\forall k$
6. $m \leftarrow m+1$
- until** convergence or $m > I_{\max 2}$
7. Update $\lambda(l+1) = \frac{B(1-\frac{\tau}{T}) \sum_{k=1}^{K-1} \sum_{j=1}^{K-1} \tilde{\mathcal{R}}_k^{(j)}(\mathbf{Q}^*)}{P_{\text{tot}}(\mathbf{Q}^*)}$
8. $l \leftarrow l+1$
9. **Until** convergence or $l > I_{\max 1}$
10. Obtain the optimal power allocation $p_k^* = 2^{q_k^*}$, $\forall k$.

To get a better insight into the computational complexity of our proposed algorithm, we perform an exhaustive complexity analysis. First, it is assumed that the network factor λ converges in W iterations. The optimization problem (52) consists of $K \times (K-1)$ sub-problems due to K UEs operating on $K-1$ effective time slots. Besides, the computational complexity resulted by these constraints C1' - C3' is $\mathcal{O}(V^3 + 2)$, where V denotes each UE's power level. Furthermore, the computational complexity of updating Lagrangian dual variables is given as $\mathcal{O}(K^\varpi)$ (for example $\varpi = 2$ if the ellipsoid method is used [40]). Let us suppose if the dual objective function (54) converges in \mathcal{G} iterations, then the total complexity for the proposed OPA scheme becomes $\mathcal{O}(2W\mathcal{G}(K-1)(K)^{\varpi+2}(V^3+2))$.

6.2 Equal power allocation (EPA) scheme

To strike a balance between the computational complexity and the optimality, we propose another lower-complexity

power allocation scheme in this paper, i.e., an EPA scheme among all UEs. The same-level transmit powers of K UEs is set as $p_k = p_u$, for $\forall k$; then, the optimization problem (47) under the EPA scheme is simplified as

$$\begin{aligned} \max_{p_u} \bar{\eta}_{EE}(p_u) \quad (58) \\ \text{s. t. } \begin{cases} \text{C4 : } p_u \geq 0, \\ \text{C5 : } p_u \leq p_{\max}, \\ \text{C6 : } \mathcal{R}_k^{(j)}(p_u) \geq R_0, \forall k, j, \end{cases} \end{aligned}$$

where

$$\bar{\eta}_{EE}(p_u) = \frac{B(1-\frac{\tau}{T}) \sum_{k=1}^{K-1} \sum_{j=1}^{K-1} \bar{\mathcal{R}}_k^{(j)}(p_u)}{u_1 K p_u + P_{\text{fixed}}}, \quad (59)$$

$$\bar{\mathcal{R}}_k^{(j)}(p_u) = \log_2 \left(1 + \frac{p_u p_u}{p_u \sum_{i=1}^K \rho_{k,i}^{(j)} + \mu_k^{(j)}} \right). \quad (60)$$

To solve the above optimization problem (58), we firstly find the feasible region of p_u and then find the global extrema values. The detailed steps are shown as follows.

Firstly, we solve the Eq. $\bar{\mathcal{R}}_k^{(j)}(p_u) = R_0$ and get the solution $p_{u,k,j}^*$ for $\forall k, j$. It can be easily determined that $\bar{\mathcal{R}}_k^{(j)}(p_u)$ is a monotonically increasing function for p_u . Hence, the QoS constraints of (58) can be reset as $p_u \geq p_{u,k,j}^*$ for $\forall k, j$, i.e.,

$$p_u \geq p_{u,\max}^*, \quad (61)$$

where $p_{u,\max}^* = \max \{p_{u,1,1}^*, \dots, p_{u,k,j}^*, \dots, p_{u,K,K-1}^*\}$. Considering both C5 and (61), the feasible region of p_u for (58) becomes $[p_{u,\max}^*, p_{\max}]$. If $p_{u,\max}^* > p_{\max}$, the optimization problem becomes infeasible. Namely, there is no solution of p_u satisfying the QoS constraints, so the algorithm should adjust p_{\max} . If $p_{u,\max}^* < p_{\max}$, (58) is feasible on $[p_{u,\max}^*, p_{\max}]$.

Once feasible, we can find the global maximum of $\bar{\eta}_{EE}(p_u)$ in $[p_{u,\max}^*, p_{\max}]$. To be more specific, it can be readily proved that $\bar{\eta}_{EE}(p_u)$ is quasi-concave in p_u and therefore has a unique stationary point \bar{p}_u , which coincides with its global maximizer and can be found from the first-order derivative (i.e., $\frac{\partial \bar{\eta}_{EE}(\bar{p}_u)}{\partial \bar{p}_u} = 0$). Then, since $\bar{\eta}_{EE}(p_u)$ is strictly increasing for $p_u \leq \bar{p}_u$ and strictly decreasing for $p_u > \bar{p}_u$. Therefore, the solution of (58), p_u^* is obtained as follows.

$$p_u^* = \begin{cases} p_{u,\max}^*, & \text{if } \bar{p}_u \leq p_{u,\max}^*, \\ \bar{p}_u, & \text{if } p_{u,\max}^* < \bar{p}_u < p_{\max}, \\ p_{\max}, & \text{if } p_{\max} \leq \bar{p}_u. \end{cases} \quad (62)$$

Remark 4 Compared to the optimization problem (47), where K variables $\{p_k\}_{k=1}^K$ are optimized, the new formulation (58) only uses p_u as the optimization variable. Hence, the computational complexity of the EPA scheme in (58) is significantly lower than that of the OPA scheme solving (47). Moreover, the OPA scheme based (47) is solved by the iterative Dinkelbach's algorithm and Lagrangian dual method, which requires a complexity of $\mathcal{O}(2WG(K-1)(K)^{\varpi+2}(V^3+2))$, the EPA scheme based (58) can obtain a closed-form optimal solution by comparing extreme values and boundary values of the optimization problem (58), without iteration. In contrast to OPA scheme, the computational complexity of the EPA scheme can be negligible.

7 Numerical results

In this section, we evaluate the EE performance of the considered massive MIMO-enabled MWRN that uses the proposed energy-efficient power allocation strategies, and demonstrate the accuracy of our analytical results as well as the impacts of several relevant parameters on the optimum EE via numerical simulations. Several key simulation parameters are set as Table 1 [11, 33]. Assume that the relay coverage area is modeled as a disc and the relay is located at the geometric center of the disc. Furthermore, all UEs are assumed to be randomly and uniformly distributed in the circular cell with a radius R , we assume that no UE is closer to the relay than R_{\min} , and the log-normal shadowing $\xi_k \sim \ln \mathcal{N}(0, \sigma_k^2)$.

7.1 Accuracy of analytical results

In this subsection, we evaluate the accuracy of analytical results given in (35) with aged CSI, as well as in (38) with predicted CSI for different $f_D T_S$ and p . We use normalized Doppler shifts $f_D T_S$ to characterize channel aging. Larger

normalized Doppler shifts correspond to large CSI delays (i.e., the more serious channel aging effect). We choose $\sigma_r^2 = \sigma_u^2 = 1$ and $\tau_r = K$. For the clarity of analysis, we assume that the EPA scheme used at UEs is considered, i.e., $p_k = p_u$. All the simulated values are obtained by averaging over 10^6 independent Monte Carlo channel realizations.

Figure 2 shows the system's achievable SR versus the number of antennas at the relay M for different normalized Doppler shifts $f_D T_S$. It can be clearly seen from Fig. 2 that the relative performance gaps between the analytical results (35) (marked as Analytical) and the simulated values (27) (marked as Simulated) are very small, which demonstrates analytical results' accuracy. In addition, we can see a intuitive result that channel aging degrades the system's achievable SR. Again, it is noted that increasing the number of relay antennas M improves the system's achievable SR, as expected. This observation also implies that, when $f_D T_S$ is relatively large, the contribution of the increasing of M diminishes quickly.

We now investigate the benefits of channel prediction on the achievable SR in Fig. 3. As can be readily, our analytical results (38) are in perfect agreement with the simulated curves (27), demonstrating the accuracy of analytical results. In addition, it is noted that, as the normalized Doppler shift $f_D T_S$ becomes large, the achievable SR loss increases significantly. Apparently, when the channel prediction order grows large, the achievable SR gain improves considerably. We also observe that, when the channel aging effect is less severe (i.e., $f_D T_S$ is small), channel prediction becomes more important. Finally, it can be observed that, the predicted CSI case achieves a higher SR than the current CSI (no channel aging) case when $f_D T_S$ is small, while its performance degrades substantially when $f_D T_S$ is large and becomes worse than that with the current CSI case.

7.2 Optimality of the proposed optimization strategy

In Fig. 4, we show the convergence behavior of the proposed power allocation strategies (including both the OPA and the EPA schemes) under different channel prediction orders p . It can be observed that the EEs of the OPA scheme (by solving (47)) are monotonically increased with the iteration number, then converge to the optimal EE value after only a few iterations. In addition, in order to further demonstrate the effectiveness of the proposed schemes, a performance comparison is given with other algorithm (i.e., Charnes-Cooper transformation (CCT)-based method) for power allocation in [41]. From Fig. 4, we can observe that the OPA scheme with the CCT-based method is slightly superior to the proposed OPA scheme with lower iterations, but the CCT-based method involves perspective transformations, which increases the computational complexity.

Table 1 Simulation parameters

Parameter	Value
Reference distance: R_{\min}	30 m
Variance of log-normal shadowing fading: σ_k^2	8 dB
Transmission bandwidth: B	20 MHz
PA efficiency at the relay: η_r	0.39
Channel coherence bandwidth: B_c	180 kHz
PA efficiency at UEs: η_u	0.3
Channel coherence time: T_c	10 ms
Circuit power consumption at the relay: P_{cR}	1 W
Fixed power consumption: P_f	20 W
Circuit power consumption at UEs: P_{cU}	0.1 W
Power efficiency of float-point operation: $\frac{P_{flops}}{\eta_C}$	1800
baseband processing power consumption at the relay: P_{RR}	29.6 W

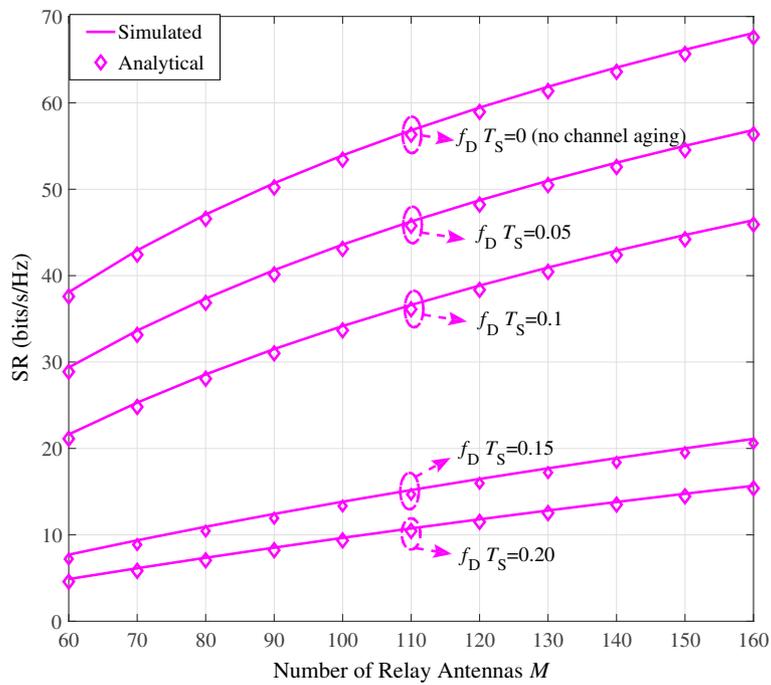


Fig. 2 The system’s achievable SR versus the number of relay antennas M for different normalized Doppler shifts $f_D T_S$. $K = 20$, $p_u = 20$ dBm, $p_r = 40$ dBm, and $p_p = 40$ dBm

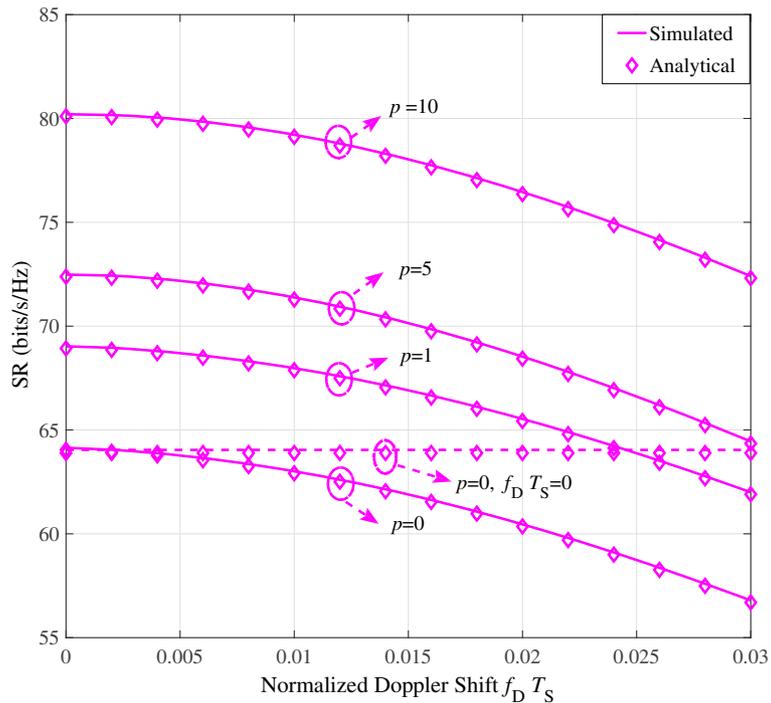


Fig. 3 The system’s achievable SR versus the normalized Doppler shifts $f_D T_S$ for different prediction order p . $K = 20$, $M = 128$, $p_u = 20$ dBm, $p_r = 40$ dBm, and $p_p = 40$ dBm

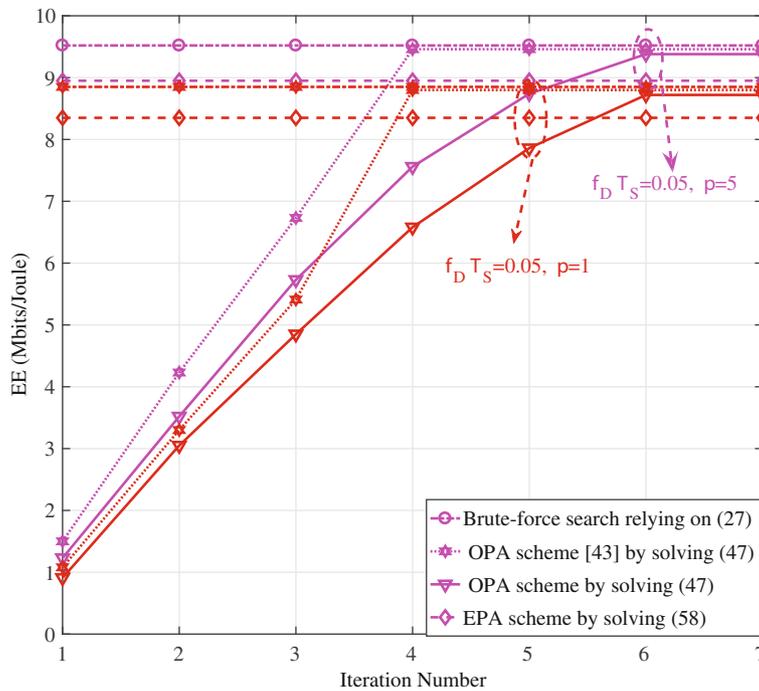


Fig. 4 Convergence and optimality of the proposed power allocation strategies, under different prediction orders p . $p_p = 40$ dBm, $M = 128$, $K = 20$, $R_0 = 4$ bit/s/Hz, $p_{max} = 24$ dBm and $p_r = 40$ dBm

At the same time, the lower-complexity EPA schemes for solving (58) achieve near-optimum performances, dispensing with any iteration. Moreover, it can be observed that the obtained EE performances of the EPA schemes are slightly worse those of the OPA schemes. Finally, in order to valid the accuracy of the derived lower bound and the optimality of the proposed power allocation strategies, in Fig. 4, we provide a performance benchmark that correspond to solving the problem (47) via the high-complexity brute-force searching relying on the ergodic achievable SR in (27). From Fig. 4, we can see that the EEs of the proposed methods are slightly inferior to the benchmarks, and this is mainly because the proposed schemes are sub-optimal methods which involve iterations and convex approximation. When using OPA scheme, K variables $\{p_k\}_{k=1}^K$ must be optimized. By contrast, with EPA scheme, we only need to optimize the single variable p_u . Furthermore, the OPA scheme obtains the optimal power allocation solution in virtue of the complicated iterative Dinkelbach’s algorithm and Lagrangian dual method. The EPA scheme can obtain a closed-form optimal solution by only comparing extreme values and boundary values of the optimization problem (58), without iteration. Hence, compared with the the OPA scheme, the computational complexity of the EPA scheme is significantly reduced. Therefore, the EPA scheme is a good choice in terms of the tradeoff between the achievable EE performance and the computational complexity. Finally, numerical results

also reveal that higher prediction order can obtain the improvement of EE performance.

Figure 5 illustrates the optimum EE achieved by the proposed power allocation strategies versus the transmit power constraint p_{max} . It is observed that the OPA scheme slightly outperforms the EPA scheme in terms of the optimum EE achieved. Furthermore, we can see that, when $p_{max} \leq 26$ dBm, the optimum EEs achieved by these proposed schemes can be substantially improved as p_{max} increases. This observation suggests that at this region [10, 26] dBm, increasing the available power budget is an energy-efficient choice. However, when $p_{max} \geq 26$ dBm, the optimum EEs of the proposed power allocation scheme converge to a certain stable level. This important observation suggests that, when p_{max} is large enough, the increasing of transmit power may not be a good choice from the perspective of EE. Finally, it is observed that the smaller $f_S T_D$ achieves a higher EE for either power allocation strategy. This is rather expected, since the smaller $f_S T_D$ means the less serious channel aging effect, and the EE loss becomes smaller accordingly.

Figure 6 illustrates the impact of the QoS threshold R_0 on the optimum EEs achieved by the proposed power allocation schemes. It can be readily noted that when $R_0 \leq 4$ bit/s/Hz, each optimum EE remains unchanged. This happens because when R_0 takes small values, it is easy to satisfy the link’s QoS requirement. This observation suggests that, at the low QoS requirement region

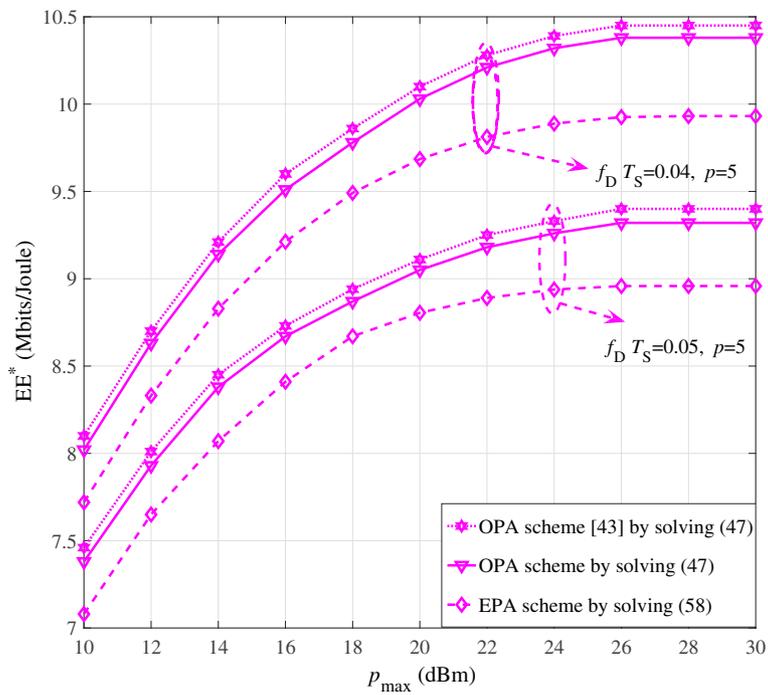


Fig. 5 The optimum EE versus the transmit power constraint, p_{\max} , under different normalized Doppler shifts $f_D T_S$. $K = 20$, $M = 128$, $p_r = 40$ dBm, $R_0 = 4$ bit/s/Hz and $p_p = 40$ dBm

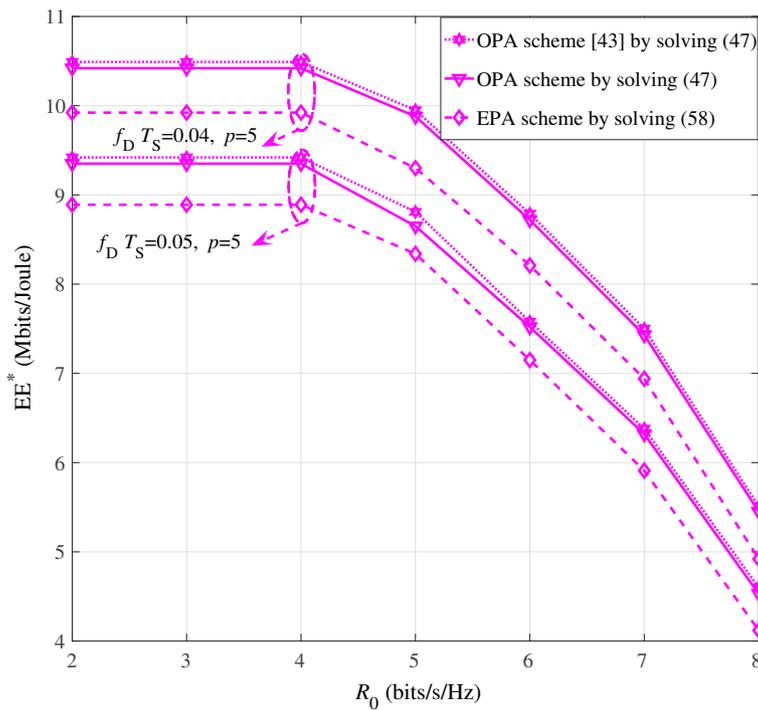
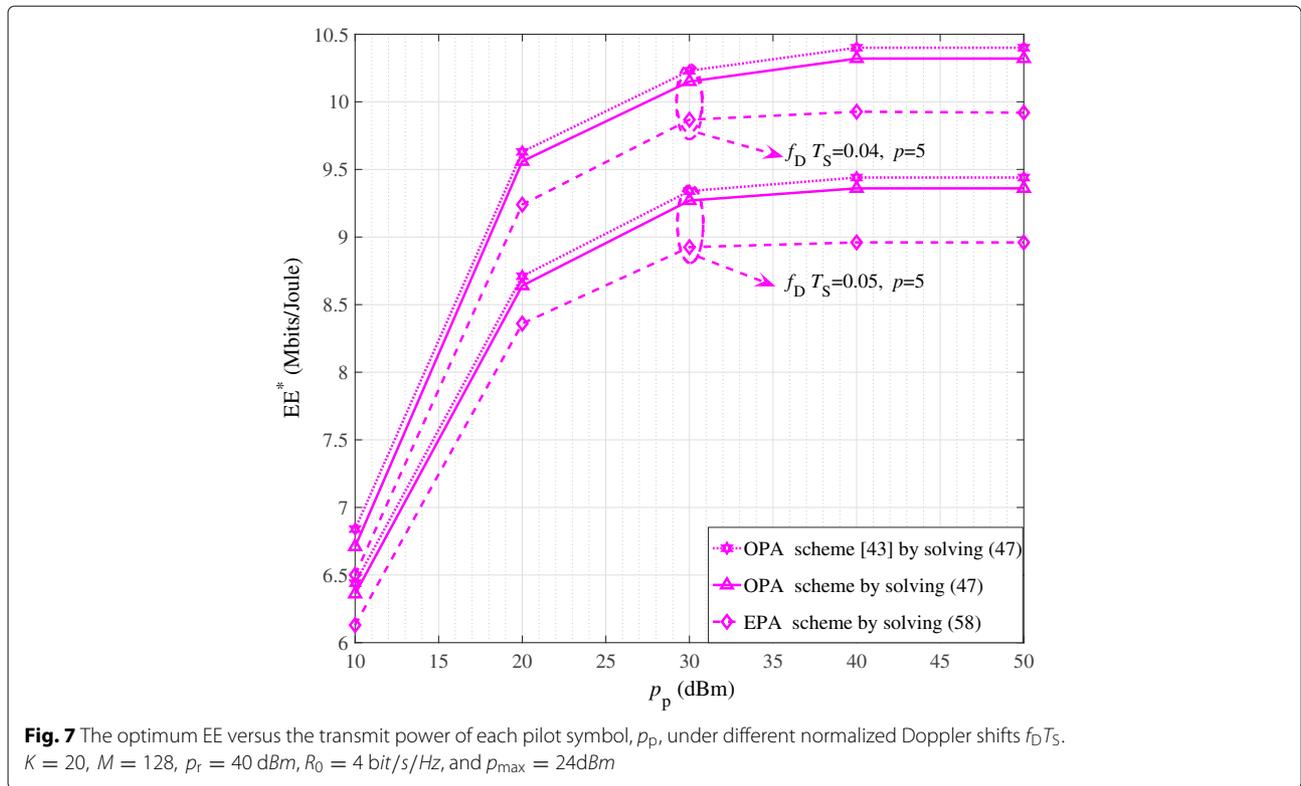


Fig. 6 The optimum EE versus the QoS constraint, R_0 , under different normalized Doppler shifts $f_D T_S$. $K = 20$, $M = 128$, $p_r = 40$ dBm, $p_{\max} = 24$ dBm and $p_p = 40$ dBm



$R_0 \leq 4$ bit/s/Hz, we can make the best use of all the available power to achieve the maximum EE, without having to waste more power on unfavorable links. Meanwhile, when $R_0 \geq 4$ bit/s/Hz, the optimum EEs decrease as R_0 increases. This is due to the fact that when R_0 increases, an excess fraction of power has to be allocated to compensate for disadvantageous links, which results in a degradation of optimum EE. In other words, a higher minimum rate R_0 is satisfied at the expense of a reduction of the optimum EE.

In Fig. 7, we show the impact of the transmit power of each pilot symbol p_p on the optimum EEs achieved by the proposed power allocation schemes. From these results and as it was expected, it can readily be observed that the the optimum EEs of all schemes increases with increasing p_p . Moreover, as p_p grows large, the growth of the achievable EE gradually slows down and saturates to the value that relies on perfect CSI estimation. This implies that although the system with high transmit power of each pilot symbol (i.e., $p_p = 50$ dBm) is capable of improving the CE accuracy, then achieves a better EE performance, the extremely high CE accuracy is not a wise choice at the cost of consuming more power.

8 Conclusions

In this paper, we have provided the performance analysis of the system’s achievable SR and proposed low-complexity power allocation strategies for maximizing the EE of a massive MIMO-enabled MWRN with channel

aging. Specifically, we derived closed-form expressions for the system’s achievable SR with/without channel prediction. Based on the derived analytical results, a unified power allocation optimization problem is established, under the transmit power and QoS constraints. Owing to the non-convexity of the objective function and QoS constraints, the original non-convex problem is sequentially approximated as a solvable CFP problem with the aid of the SCA technique, which can be efficiently solved by the Dinkelbach’s algorithm and Lagrangian dual method. Moreover, we have proposed a closed-form power control algorithm for the lower-complexity EPA scheme. The impacts of normalized Doppler shifts $f_D T_S$, channel prediction order, and other relevant system parameters on the SR and EE performance are investigated via numerical simulations, which have verified the accuracy of our analytical results, and confirmed the effectiveness of the proposed power allocation schemes.

Endnotes

¹Although we respectively studied the resource allocation for EE maximization in the massive MIMO-enabled OWRNs and MWRNs in [42, 43], the works in [42, 43] only considered the channel estimation (CE) error and ignored the effect of channel aging. Contrast to the transmission schemes proposed in [42, 43], the performance analysis and power allocation algorithms in this

paper have stronger robustness over the practical communication scenario.

² This setup is general enough to model a variety of communication scenarios. Certain practical applications such as multimedia teleconferencing via a satellite or mutual data exchange between sensor nodes and the data fusion center in wireless sensor networks require mutual data exchange among more than just two terminals.

³ In this paper, for simplicity and tractability, we assumed that the direct link between any two UEs is ignored due to severe path loss. This assumption has been widely made in multi-way relay systems [13, 14], and it is easily extended to complex system models with a direct communication link between two UEs.

⁴ The log-sum-exp is convex [38].

⁵ The numerator is concave, and the denominator is convex [38].

Appendix 1 Proof of Theorem 1

With aged CSI, $\tilde{\mathbf{g}}_k[n+1] \sim \mathcal{CN}(\mathbf{0}, \tilde{\beta}_{a,k} \mathbf{I}_K)$ is independent of $\xi_{a,k}[n+1] \sim \mathcal{CN}(\mathbf{0}, \tilde{\beta}_{a,k} \mathbf{I}_K)$.

1. Derive ϑ_j : According to (3) and (25), we have

$$\frac{p_r}{\vartheta_j^2} = \mathbb{E} \left[\left\| \mathbf{F}_j[n+1] \mathbf{y}_R[n+1] \right\|^2 \right] = \Psi_1^{(j)} + \Psi_2^{(j)} + \Psi_3^{(j)}, \quad (63)$$

where

$$\Psi_1^{(j)} = \mathbb{E} \left[\left\| \mathbf{F}_j[n+1] \bar{\mathbf{G}}[n+1] \mathbf{x}_U[n+1] \right\|^2 \right] \quad (64)$$

$$\Psi_2^{(j)} = \mathbb{E} \left[\left\| \mathbf{F}_j[n+1] \tilde{\mathbf{G}}[n+1] \mathbf{x}_U[n+1] \right\|^2 \right]$$

$$\Psi_3^{(j)} = \sigma_r^2 \mathbb{E} \left[\left\| \mathbf{F}_j[n+1] \right\|^2 \right].$$

with $\tilde{\mathbf{G}}[n+1] = [\xi_{a,1}[n+1], \dots, \xi_{a,K}[n+1]]$.

- Compute $\Psi_1^{(j)}$: According to the definition of $\mathbf{F}_j[n+1]$, we have

$$\begin{aligned} \Psi_1^{(j)} &\stackrel{(a)}{=} \mathbb{E} \left[\text{Tr} \left\{ \mathbf{P} \boldsymbol{\pi}_j^H \mathbf{W}_2^H[n+1] \mathbf{W}_2[n+1] \boldsymbol{\pi}_j \right\} \right] \\ &= \sum_{k=1}^K p_k \mathbb{E} \left[\left[\boldsymbol{\pi}_j^H (\bar{\mathbf{G}}^H[n+1] \bar{\mathbf{G}}[n+1])^{-1} \boldsymbol{\pi}_j \right]_{k,k} \right] \\ &\stackrel{(b)}{=} \mu_1 \sum_{k=1}^K p_k \tilde{\beta}_{a,k}^{-1}, \end{aligned} \quad (65)$$

where $\mu_1 = \frac{1}{M-K-1}$ and $k'' = \text{mod}_K(K+k-j)$. (a) results from the property $\text{Tr}\{\mathbf{A}\mathbf{B}\} = \text{Tr}\{\mathbf{B}\mathbf{A}\}$. As to the detailed derivation of (b), we use the identity as follows [44]: $\boldsymbol{\Omega} \triangleq (\bar{\mathbf{G}}^H[n+1] \bar{\mathbf{G}}[n+1])^{-1}$ is an inverted Wishart matrix, i.e., $\boldsymbol{\Omega} \sim \mathcal{W}_K^{-1}(M+K+1, \hat{\mathbf{D}}_a^{-1})$ with $\hat{\mathbf{D}}_a^{-1} = \text{diag} \{ \tilde{\beta}_{a,1}^{-1}, \dots, \tilde{\beta}_{a,K}^{-1} \}$. Hence, we have [45]

$$\mathbb{E}[\boldsymbol{\Omega}] = \mathbb{E} \left[(\bar{\mathbf{G}}^H[n+1] \bar{\mathbf{G}}[n+1])^{-1} \right] = \frac{\hat{\mathbf{D}}_a^{-1}}{M-K-1}. \quad (66)$$

- Compute $\Psi_2^{(j)}$: Since $\bar{\mathbf{G}}[n+1]$ and $\tilde{\mathbf{G}}[n+1]$ are independent, we obtain

$$\begin{aligned} \Psi_2^{(j)} &\stackrel{(a)}{=} \mathbb{E} \left[\text{Tr} \left\{ \mathbf{P} \tilde{\mathbf{G}}^H[n+1] \mathbf{F}_j^H[n+1] \mathbf{F}_j[n+1] \tilde{\mathbf{G}}[n+1] \right\} \right] \\ &= \sum_{k=1}^K p_k \tilde{\beta}_{a,k} \mathbb{E} \left[\text{Tr} \left\{ \mathbf{F}_j^H[n+1] \mathbf{F}_j[n+1] \right\} \right], \end{aligned} \quad (67)$$

in which

$$\begin{aligned} \mathbb{E} \left[\text{Tr} \left\{ \mathbf{F}_j^H[n+1] \mathbf{F}_j[n+1] \right\} \right] &= \mathbb{E} \left[\text{Tr} \left\{ \boldsymbol{\pi}_j^H \boldsymbol{\Omega} \boldsymbol{\pi}_j \right\} \right] \\ &= \sum_{k=1}^K \mathbb{E} \left[[\boldsymbol{\Omega}]_{k'',k''} [\boldsymbol{\Omega}]_{k,k} \right] \\ &\quad + \sum_{m=1, m \neq k}^K \mathbb{E} \left[[\boldsymbol{\Omega}]_{k'',m''} [\boldsymbol{\Omega}]_{m,k} \right] \\ &\stackrel{(c)}{=} \mu_2 \mathcal{A}_{aj} \end{aligned} \quad (68)$$

where $m'' = \text{mod}_K(K+m-j)$, $\mu_2 = \frac{(2+(M-K)(M-K-3))}{(M-K)(M-K-1)^2(M-K-3)}$ and $\mathcal{A}_{aj} = \sum_{k=1}^K \tilde{\beta}_{a,k}^{-1} \tilde{\beta}_{a,k''}^{-1}$. The detailed derivation of (c) is given as follows [45]:

$$\begin{aligned} \mathbb{E} \left[[\boldsymbol{\Omega}]_{k,k} [\boldsymbol{\Omega}]_{k'',k''} \right] &= \text{Cov} \left([\boldsymbol{\Omega}]_{k,k}, [\boldsymbol{\Omega}]_{k'',k''} \right) \\ &\quad + \mathbb{E} \left[[\boldsymbol{\Omega}]_{k,k} \right] \mathbb{E} \left[[\boldsymbol{\Omega}]_{k'',k''} \right] \\ &= \frac{2\tilde{\beta}_{a,k}^{-1} \tilde{\beta}_{a,k''}^{-1} + (M-K)(M-K-3)\tilde{\beta}_{a,k}^{-1} \tilde{\beta}_{a,k''}^{-1}}{(M-K)(M-K-1)^2(M-K-3)}. \end{aligned} \quad (69)$$

and for $m \neq k$

$$\mathbb{E} \left[[\boldsymbol{\Omega}]_{k'',m''} [\boldsymbol{\Omega}]_{m,k} \right] = 0. \quad (70)$$

According to (71) and (68), we get

$$\Psi_2^{(j)} = \mu_2 \mathcal{A}_{aj} \sum_{k=1}^K p_k \tilde{\beta}_{a,k}. \quad (71)$$

- Compute $\Psi_3^{(j)}$: Similarly, according to (68), we can obtain

$$\begin{aligned} \Psi_3^{(j)} &= \sigma_r^2 \mathbb{E} \left[\text{Tr} \left\{ \mathbf{F}_j[n+1] \mathbf{F}_j^H[n+1] \right\} \right] \\ &= \mu_2 \sigma_r^2 \mathcal{A}_{aj}. \end{aligned} \quad (72)$$

Substituting (65), (71), and (72) into (8), we have

$$\begin{aligned} \vartheta_j^2 &= \frac{p_r}{\Psi_1^{(j)} + \Psi_2^{(j)} + \Psi_3^{(j)}} \\ &= \frac{p_r}{\sum_{k=1}^K p_k \left(\mu_1 \tilde{\beta}_{a,k}^{-1} + \mu_2 \tilde{\beta}_{a,k} \mathcal{A}_{aj} \right) + \mu_2 \sigma_r^2 \mathcal{A}_{aj}}. \end{aligned} \quad (73)$$

- Derive $\widehat{\mathcal{R}}_k^{(j)}$: From (31), we need to compute

$$\mathbb{E} \left[\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1] \right], \text{SI}_k^{(j)}, \text{UI}_k^{(j)}, \text{NR}_k^{(j)}, \text{NU}_k, \text{ and } \text{Var} \left[\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1] \right].$$

- Compute $\mathbb{E} \left[\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1] \right]$: We have

$$\mathbb{E} \left[\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1] \right] = \varphi_1 + \varphi_2 + \varphi_3 + \varphi_4, \quad (74)$$

with

$$\begin{aligned} \varphi_1 &= \mathbb{E} \left[\bar{\mathbf{g}}_k^T[n+1] \mathbf{F}_j[n+1] \bar{\mathbf{g}}_{k'}[n+1] \right] = 1, \\ \varphi_2 &= \mathbb{E} \left[\bar{\mathbf{g}}_k^T[n+1] \mathbf{F}_j[n+1] \boldsymbol{\xi}_{a,k'}[n+1] \right] \stackrel{(d)}{=} 0, \\ \varphi_3 &= \mathbb{E} \left[\boldsymbol{\xi}_{a,k}^T[n+1] \mathbf{F}_j[n+1] \bar{\mathbf{g}}_{k'}[n+1] \right] \stackrel{(d)}{=} 0, \\ \varphi_4 &= \mathbb{E} \left[\boldsymbol{\xi}_{a,k}^T[n+1] \mathbf{F}_j[n+1] \boldsymbol{\xi}_{a,k'}[n+1] \right] \stackrel{(d)}{=} 0, \end{aligned} \quad (75)$$

where (d) results from the independence between $\bar{\mathbf{g}}_i[n+1]$ and $\boldsymbol{\xi}_{a,j}[n+1]$ for $\forall i, j$. Hence, we have

$$\mathbb{E} \left[\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1] \right] = 1. \quad (76)$$

- Compute $\text{Var} \left[\mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1] \right]$: According to the definition of variance, we have

$$\text{Var} \left[|\varsigma_{k,j}|^2 \right] = \mathbb{E} \left[|\varsigma_{k,j}|^2 \right] - \left(\mathbb{E}[\varsigma_{k,j}] \right)^2, \quad (77)$$

where $\varsigma_{k,j} = \mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_{k'}[n+1]$, $\mathbb{E} \left[|\varsigma_{k,j}|^2 \right]$ can be decomposed into the following parts:

$$\mathbb{E} \left[|\varsigma_{k,j}|^2 \right] = \omega_1 + \omega_2 + \omega_3 + \omega_4, \quad (78)$$

with

$$\begin{aligned} \omega_1 &= \mathbb{E} \left[\left| \bar{\mathbf{g}}_k^T[n+1] \mathbf{F}_j[n+1] \bar{\mathbf{g}}_{k'}[n+1] \right|^2 \right] \\ &= \mathbb{E} \left[\mathbf{e}_k^T \boldsymbol{\pi}_j \mathbf{e}_{k'} \mathbf{e}_{k'}^H \boldsymbol{\pi}_j^H \mathbf{e}_k \right] = 1, \\ \omega_2 &= \mathbb{E} \left[\left| \bar{\mathbf{g}}_k^T[n+1] \mathbf{F}_j[n+1] \boldsymbol{\xi}_{a,k'}[n+1] \right|^2 \right] \\ &= \tilde{\beta}_{a,k'} \mathbb{E} \left[\mathbf{e}_k^T \boldsymbol{\pi}_j \mathbf{W}_1[n+1] \mathbf{W}_1^H[n+1] \boldsymbol{\pi}_j^H \mathbf{e}_k \right] \\ &= \tilde{\beta}_{a,k'} \mathbb{E} \left[[\boldsymbol{\Omega}]_{k',k'} \right] \stackrel{(b)}{=} \mu_1 \tilde{\beta}_{a,k'} \tilde{\beta}_{a,k'}^{-1}, \\ \omega_3 &= \mathbb{E} \left[\left| \boldsymbol{\xi}_{a,k}^T[n+1] \mathbf{F}_j[n+1] \bar{\mathbf{g}}_{k'}[n+1] \right|^2 \right] \\ &= \tilde{\beta}_{a,k} \mathbb{E} \left[\mathbf{e}_{k'}^T \boldsymbol{\pi}_j^H \mathbf{W}_2^H[n+1] \mathbf{W}_2[n+1] \boldsymbol{\pi}_j \mathbf{e}_{k'} \right] \\ &= \tilde{\beta}_{a,k} \mathbb{E} \left[[\boldsymbol{\Omega}]_{k,k} \right] \stackrel{(b)}{=} \mu_1 \tilde{\beta}_{a,k} \tilde{\beta}_{a,k}^{-1}, \\ \omega_4 &= \mathbb{E} \left[\left| \boldsymbol{\xi}_{a,k}^T[n+1] \mathbf{F}_j[n+1] \boldsymbol{\xi}_{a,k'}[n+1] \right|^2 \right] \\ &= \tilde{\beta}_{a,k} \tilde{\beta}_{a,k'} \mathbb{E} \left[\text{Tr} \left\{ \mathbf{F}_j^H[n+1] \mathbf{F}_j[n+1] \right\} \right] \\ &= \tilde{\beta}_{a,k} \tilde{\beta}_{a,k'} \sum_{k=1}^K \mathbb{E} \left[[\boldsymbol{\Omega}]_{k,k} [\boldsymbol{\Omega}]_{k',k'} \right] \\ &\stackrel{(c)}{=} \mu_2 \tilde{\beta}_{a,k} \tilde{\beta}_{a,k'} \mathcal{A}_{aj}. \end{aligned} \quad (79)$$

According to (77)–(79), we obtain

$$\begin{aligned} \text{Var}(\varsigma_{k,j}) &= \mu_1 \tilde{\beta}_{a,k'} \tilde{\beta}_{a,k'}^{-1} + \mu_1 \tilde{\beta}_{a,k} \tilde{\beta}_{a,k}^{-1} \\ &\quad + \mu_2 \tilde{\beta}_{a,k} \tilde{\beta}_{a,k'} \mathcal{A}_{aj}. \end{aligned} \quad (80)$$

- Compute $\text{SI}_k^{(j)}$: We have

$$\begin{aligned} \mathbb{E} \left[|\lambda_k|^2 \right] &\stackrel{(e)}{=} \mathbb{E} \left[\left| \mathbf{g}_k^T[n+1] \mathbf{F}_j[n+1] \mathbf{g}_k[n+1] \right|^2 \right] \\ &= \omega_5 + \omega_6 + \omega_7 + \omega_8, \end{aligned} \quad (81)$$

where (e) results from the property $\bar{\mathbf{g}}_k^T[n+1] \mathbf{F}_j[n+1] \bar{\mathbf{g}}_k[n+1] = \mathbf{e}_k^T \boldsymbol{\pi}_j \mathbf{e}_k = 0$, and

$$\begin{aligned}
 \omega_5 &= \mathbb{E} \left[\left| \bar{\mathbf{g}}_k^T [n+1] \mathbf{F}_j [n+1] \bar{\mathbf{g}}_k [n+1] \right|^2 \right] \\
 &= \mathbb{E} \left[\mathbf{e}_k^T \boldsymbol{\pi}_j \mathbf{e}_k \mathbf{e}_k^T \boldsymbol{\pi}_j^H \mathbf{e}_k \right] = 0, \\
 \omega_6 &= \mathbb{E} \left[\left| \bar{\mathbf{g}}_k^T [n+1] \mathbf{F}_j [n+1] \boldsymbol{\xi}_{a,k} [n+1] \right|^2 \right] \\
 &= \tilde{\beta}_{a,k} \mathbb{E} \left[\mathbf{e}_k^T \boldsymbol{\pi}_j \mathbf{W}_1 [n+1] \mathbf{W}_1^H [n+1] \boldsymbol{\pi}_j^H \mathbf{e}_k \right] \\
 &= \tilde{\beta}_{a,k} \mathbb{E} \left[[\boldsymbol{\Omega}]_{k',k'} \right] \stackrel{(b)}{=} \mu_1 \tilde{\beta}_{a,k} \tilde{\beta}_{a,k}^{-1}, \\
 \omega_7 &= \mathbb{E} \left[\left| \boldsymbol{\xi}_{a,k}^T [n+1] \mathbf{F}_j [n+1] \bar{\mathbf{g}}_k [n+1] \right|^2 \right] \\
 &= \tilde{\beta}_{a,k} \mathbb{E} \left[\mathbf{e}_k^T \boldsymbol{\pi}_j^H \boldsymbol{\Omega} \boldsymbol{\pi}_j \mathbf{e}_k \right] \\
 &= \tilde{\beta}_{a,k} \mathbb{E} \left[[\boldsymbol{\Omega}]_{k'',k''} \right] \stackrel{(b)}{=} \mu_1 \tilde{\beta}_{a,k} \tilde{\beta}_{a,k}^{-1}, \\
 \omega_8 &= \mathbb{E} \left[\left| \boldsymbol{\xi}_{a,k}^T [n+1] \mathbf{F}_j [n+1] \boldsymbol{\xi}_{a,k} [n+1] \right|^2 \right] \\
 &= \tilde{\beta}_{a,k}^2 \mathbb{E} \left[\text{Tr} \left[\boldsymbol{\pi}_j^H \boldsymbol{\Omega} \boldsymbol{\pi}_j \boldsymbol{\Omega} \right] \right] \\
 &= \tilde{\beta}_{a,k}^2 \sum_{k=1}^K \mathbb{E} \left[[\boldsymbol{\Omega}]_{k'',k''} [\boldsymbol{\Omega}]_{k,k} \right] \stackrel{(c)}{=} \mu_2 \tilde{\beta}_{a,k}^2 \mathcal{A}_{aj}.
 \end{aligned} \tag{82}$$

Substituting (82) into (81), we obtain

$$\mathbb{E} [|\lambda_k|^2] = \mu_1 \tilde{\beta}_{a,k} \tilde{\beta}_{a,k}^{-1} + \mu_1 \tilde{\beta}_{a,k} \tilde{\beta}_{a,k}^{-1} + \mu_2 \tilde{\beta}_{a,k}^2 \mathcal{A}_{aj}. \tag{83}$$

- Compute $\text{UI}_k^{(j)}$: Following the same methodology used for computing $\mathbb{E} [|\lambda_k|^2]$, we can easily derive

$$\begin{aligned}
 \mathbb{E} \left[\left| \mathbf{g}_k^T [n+1] \mathbf{F}_j [n+1] \mathbf{g}_i [n+1] \right|^2 \right] &= \omega_9 + \omega_{10} \\
 &\quad + \omega_{11} + \omega_{12}
 \end{aligned} \tag{84}$$

with

$$\begin{aligned}
 \omega_9 &= \mathbb{E} \left[\left| \bar{\mathbf{g}}_k^T [n+1] \mathbf{F}_j [n+1] \bar{\mathbf{g}}_i [n+1] \right|^2 \right] \\
 &= \mathbf{e}_k^T \boldsymbol{\pi}_j \mathbf{e}_i \mathbf{e}_i^T \boldsymbol{\pi}_j^H \mathbf{e}_k = 0, \\
 \omega_{10} &= \mathbb{E} \left[\left| \bar{\mathbf{g}}_k^T [n+1] \mathbf{F}_j [n+1] \boldsymbol{\xi}_{a,i} [n+1] \right|^2 \right] \\
 &= \tilde{\beta}_{a,i} \mathbb{E} \left[\mathbf{e}_k^T \boldsymbol{\pi}_j \boldsymbol{\Omega} \boldsymbol{\pi}_j^H \mathbf{e}_k \right] \stackrel{(b)}{=} \mu_1 \tilde{\beta}_{a,i} \tilde{\beta}_{a,k}^{-1},
 \end{aligned} \tag{85}$$

$$\begin{aligned}
 \omega_{11} &= \mathbb{E} \left[\left| \boldsymbol{\xi}_{a,k}^T [n+1] \mathbf{F}_j [n+1] \bar{\mathbf{g}}_i [n+1] \right|^2 \right] \\
 &= \tilde{\beta}_{a,k} \mathbb{E} \left[\mathbf{e}_i^T \boldsymbol{\pi}_j^H \boldsymbol{\Omega} \boldsymbol{\pi}_j \mathbf{e}_i \right] \stackrel{(b)}{=} \mu_1 \tilde{\beta}_{a,k} \tilde{\beta}_{a,i}^{-1}, \\
 \omega_{12} &= \mathbb{E} \left[\left| \boldsymbol{\xi}_{a,k}^T [n+1] \mathbf{F}_j [n+1] \boldsymbol{\xi}_{a,i} [n+1] \right|^2 \right] \\
 &= \tilde{\beta}_{a,k} \tilde{\beta}_{a,i} \mathbb{E} \left[\text{Tr} \left\{ \mathbf{F}_j [n+1] \mathbf{F}_j^H [n+1] \right\} \right] \\
 &\stackrel{(c)}{=} \mu_2 \tilde{\beta}_{a,k} \tilde{\beta}_{a,i} \mathcal{A}_{aj}.
 \end{aligned} \tag{86}$$

where $i'' = \text{mod}_K (K + i - j)$.

According to (84) and (85), we can obtain

$$\begin{aligned}
 \mathbb{E} \left[\left| \mathbf{g}_k^T [n+1] \mathbf{F}_j [n+1] \mathbf{g}_i [n+1] \right|^2 \right] &= \mu_1 \tilde{\beta}_{a,i} \tilde{\beta}_{a,k}^{-1} \\
 &\quad + \mu_1 \tilde{\beta}_{a,k} \tilde{\beta}_{a,i}^{-1} \\
 &\quad + \mu_2 \tilde{\beta}_{a,i} \tilde{\beta}_{a,k} \mathcal{A}_{aj}.
 \end{aligned} \tag{87}$$

Therefore, we can conclude

$$\text{UI}_k^{(j)} = \sum_{i \neq k, k'}^K p_i \mu_1 \tilde{\beta}_{a,i} \tilde{\beta}_{a,k}^{-1} + p_i \mu_1 \tilde{\beta}_{a,k} \tilde{\beta}_{a,i}^{-1} + \mu_2 \tilde{\beta}_{a,i} \tilde{\beta}_{a,k} \mathcal{A}_{aj}. \tag{88}$$

- Compute $\text{NR}_k^{(j)}$: We have

$$\mathbb{E} \left[\left| \mathbf{g}_k^T [n+1] \mathbf{F}_j [n+1] \mathbf{n}_R [n+1] \right|^2 \right] = \omega_{13} + \omega_{14}, \tag{89}$$

with

$$\begin{aligned}
 \omega_{13} &= \mathbb{E} \left[\left| \bar{\mathbf{g}}_k^T [n+1] \mathbf{F}_j [n+1] \mathbf{n}_R [n+1] \right|^2 \right] \\
 &= \sigma_r^2 \mathbb{E} \left[\mathbf{e}_k^T \boldsymbol{\pi}_j \boldsymbol{\Omega} \boldsymbol{\pi}_j^H \mathbf{e}_k \right] \stackrel{(b)}{=} \mu_1 \sigma_r^2 \tilde{\beta}_{a,k}^{-1}, \\
 \omega_{14} &= \mathbb{E} \left[\left| \boldsymbol{\xi}_{a,k}^T [n+1] \mathbf{F}_j [n+1] \mathbf{n}_R [n+1] \right|^2 \right] \\
 &= \sigma_r^2 \tilde{\beta}_{a,k} \mathbb{E} \left[\text{Tr} \left\{ \mathbf{F}_j^H [n+1] \mathbf{F}_j [n+1] \right\} \right] \\
 &\stackrel{(c)}{=} \mu_2 \sigma_r^2 \tilde{\beta}_{a,k} \mathcal{A}_{aj}.
 \end{aligned} \tag{90}$$

Substituting (90) into (89), we can obtain

$$\text{NR}_k^{(j)} = \mu_1 \sigma_r^2 \tilde{\beta}_{a,k}^{-1} + \mu_2 \sigma_r^2 \tilde{\beta}_{a,k} \mathcal{A}_{aj}. \tag{91}$$

- Compute NU_k : We have

$$\text{NU}_k \triangleq \mathbb{E} \left[|n_{u,k} [n+1]|^2 \right] = \sigma_u^2. \tag{92}$$

Substituting (73), (76), (80), (83), (88), (91), and (92) into (32), we can obtain (35) after some simple algebraic manipulations. Thus, the proof of Theorem 1 is complete.

Abbreviations

5G: Fifth-generation; AWGN: Additive white Gaussian noise; BS: Base station; CCT: Charnes-Cooper transformation; CE: Channel estimation; CFP: Concave fractional programming; CSI: Channel state information; EE: Energy efficiency; EPA: Equal power allocation; LSF: Large-scale fading; MIMO: Multiple-input multiple-output; MMSE: Minimum mean square error; MWRN: Multi-way relay network; OPA: Optimal power allocation; OWRN: One-way relay network; QoS: Quality-of-service; SCA: Successive convex approximation; SE: Spectral efficiency; SINR: Signal-to-interference-plus-noise ratio; SR: Sum rate; SSF: Small-scale fading; TWRN: Two-way relay network; UE: User equipment; ZF: Zero-forcing

Funding

This work was supported by the National Natural Science Foundation of China (61471135, 61671165); the Guangxi Natural Science Foundation (2016GXNSFGA380009); the PhD Research Startup Fund of Guilin University of Electronic Technology (UF17048Y); the Fund of Key Laboratory of Cognitive Radio and Information Processing (Guilin University of Electronic Technology), China; and the Guangxi Key Laboratory of Wireless Wideband Communication and Signal Processing (CRKL160105, CRKL170101).

Authors' contributions

FT was responsible for mathematical derivation, numerical simulation, and paper writing. HC was responsible for problem formulation and paper revision. FZ was responsible for problem and result discussion. XL was responsible for model validation and result check. All authors read and approved the final manuscript.

Authors' information

Fangqing Tan received the M.Eng. degree in communication and information systems from Chongqing University of Post and Telecommunications, China, in 2012, and the Ph.D. degree in communication and information systems from Beijing University of Posts and Telecommunications, China, in 2017. He is currently a lecturer with the School of Information and Communication, Guilin University of Electronic Technology, Guilin, China. His research interests include massive MIMO systems, cooperative communications, and energy-efficient wireless communications. Hongbin Chen received the B.Eng. degree in electronic and information engineering from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2004 and the Ph.D. degree in circuits and systems from South China University of Technology, Guangzhou, China, in 2009. From October 2006 to May 2008, he was a Research Assistant with the Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong. From March to April 2014, he was a Research Associate with the same department. From May 2015 to May 2016, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. He is currently a Professor with the School of Information and Communication, Guilin University of Electronic Technology, Guilin, China. His research interests include energy-efficient wireless communications. Feng Zhao received the Ph.D. degree in communication and information systems from Shandong University, China in 2007. Now, he is a Professor with the School of Information and Communication, Guilin University of Electronic Technology, China. His research interests include wireless communications, signal processing, and information security. Xiaohuan Li received the B.Eng. and M.Sc. degrees from Guilin University of Electronic Technology, China, in 2006 and 2009, respectively, and the Ph.D. degree from South China University of Technology, China, in 2015. He is currently an Associate Professor with the School of Information and Communication, Guilin University of Electronic Technology, China. His research interests include vehicular ad hoc networks.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Key Laboratory of Cognitive Radio and Information Processing, Guilin University of Electronic Technology, Guilin 541004, China. ²School of Electronics and Communication Engineering, Yulin Normal University, Yulin 537000, China.

Received: 9 March 2018 Accepted: 2 August 2018

Published online: 20 August 2018

References

1. I. Chih-Lin, C. Rowell, S. Han, Z. Xu, G. Li, Z. Pan, Toward green and soft: a 5G perspective. *IEEE Commun. Mag.* **52**(2), 66–73 (2014)
2. Z. Hasan, H. Boostanimehr, V. Bhargava, Green cellular networks: a survey, some research issues and challenges. *IEEE Commun. Surv. Tutorials.* **13**(4), 524–540, Fourth Quarter (2011)
3. C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P.M. Grant, H. Haas, J. Thompson, I. Ku, C.X. Wang, T.A. Le, M. Nakhai, J. Zhang, L. Hanzo, Green radio: radio techniques to enable energy-efficient wireless networks. *IEEE Commun. Mag.* **49**(6), 46–54 (2011)
4. Y. Chen, S. Zhang, S. Xu, G. Li, Fundamental trade-offs on green wireless networks. *IEEE Commun. Mag.* **49**(6), 30–37 (2011)
5. F. Rusek, D. Persson, B.K. Lau, E. Larsson, T. Marzetta, O. Edfors, F. Tufvesson, Scaling up MIMO: opportunities and challenges with very large arrays. *IEEE Signal Process. Mag.* **30**(1), 40–60 (2013)
6. E.G. Larsson, O. Edfors, F. Tufvesson, T.L. Marzetta, Massive MIMO for next generation wireless systems. *IEEE Commun. Mag.* **52**(2), 186–195 (2014)
7. H.Q. Ngo, E. Larsson, T. Marzetta, Energy and spectral efficiency of very large multiuser MIMO systems. *IEEE Trans. Commun.* **61**(4), 1436–1449 (2013)
8. L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, R. Zhang, An overview of massive MIMO: benefits and challenges. *IEEE J. Sel. Topics Signal Process.* **8**(5), 742–758 (2014)
9. S. Buzzi, I. Chih-Lin, T. E. Klein, H. V. Poor, C. Yang, A. Zappone, A survey of energy-efficient techniques for 5G networks and challenges ahead. *IEEE J. Sel. Areas in Commun.* **34**(4), 697–709 (2016)
10. E. Björnson, J. Hoydis, M. Kountouris, M. Debbah, Massive mimo systems with non-ideal hardware: energy efficiency, estimation, and capacity limits. *IEEE Trans. Inf. Theory.* **60**(11), 7112–7139 (2014)
11. E. Björnson, L. Sanguinetti, J. Hoydis, M. Debbah, Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer? *IEEE Trans. Wireless Commun.* **14**(6), 3059–3075 (2015)
12. E. Björnson, L. Sanguinetti, M. Kountouris, Deploying dense networks for maximal energy efficiency: small cells meet massive mimo. *IEEE J. Sel. Areas Commun.* **34**(4), 832–847 (2016)
13. D. Gundüz, A. Yener, A. Goldsmith, H.V. Poor, The multiway relay channel. *IEEE Trans. Inf. Theory.* **59**(1), 51–63 (2013)
14. Y. Tian, A. Yener, in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*. Degrees of freedom for the MIMO multi-way relay channel (IEEE, Istanbul, 2013), pp. 1576–1580
15. G. Amarasuriya, C. Tellambura, M. Ardakani, Multi-way MIMO amplify-and-forward relay networks with zero-forcing transmission. *IEEE Trans. Commun.* **61**(12), 4847–4863 (2013)
16. H. A. Suraweera, H. Q. Ngo, T. Q. Duong, C. Yuen, E. G. Larsson, in *Proc. IEEE Int. Conf. on Commun. (ICC)*. Multi-pair amplify-and-forward relaying with very large antenna arrays (IEEE, Budapest, 2013), pp. 1–6
17. H. Cui, L. Song, B. Jiao, Multi-pair two-way amplify-and-forward relaying with very large number of relay antennas. *IEEE Trans. Wireless Commun.* **13**(5), 2636–2645 (2014)
18. S. Jin, X. Liang, K-K. Wong, X. Gao, Q. Zhu, Ergodic rate analysis for multipair massive MIMO two-way relay networks. *IEEE Trans. Wireless Commun.* **14**(3), 1480–1491 (2015)
19. G. Amarasuriya, H. V. Poor, in *Proc. 2014 IEEE 25th Annual, Int. Symp. on Personal, Indoor, and Mobile Radio Commun. (PIMRC)*. Multi-way amplify-and-forward relay networks with massive MIMO (IEEE, Washington DC, 2014), pp. 595–600
20. Kudathanthirige D.P., G. A. A Baduge, Multicell multiway massive MIMO relay networks. *IEEE Trans. Veh. Technol.* **66**(8), 6831–6848 (2017)
21. C. D. Ho, H. Q. Ngo, M. Matthaiou, T. Q. Duong, in *Proc. IEEE Int. Conf. on Recent Advances in Signal Processing, Telecommunications Computing (SigTelCom)*. Multi-way massive MIMO relay networks with maximum-ratio processing (IEEE, Da Nang, 2017), pp. 124–128
22. C. D. Ho, H. Q. Ngo, M. Matthaiou, T. Q. Duong, On the performance of zero-forcing processing in multi-way massive MIMO relay networks. *IEEE Commun. Lett.* **21**(4), 849–852 (2017)
23. G. Amarasuriya, H. V. Poor, in *Proc. 2015 IEEE Int. Conf. on Commun. (ICC)*. Impact of channel aging in multi-way relay networks with massive MIMO, (London, UK, 2015), pp. 1951–1957

24. G. Amaraluriya, E. G. Larsson, H. V. Poor, Wireless information and power transfer in multiway massive MIMO relay networks. *IEEE Trans. Wireless Commun.* **15**(6), 3837–3855 (2016)
25. K. T. Truong, R. W. Heath, Effects of channel aging in massive MIMO systems. *J. Commun. and Netw.* **15**(4), 338–351 (2013)
26. A. K. Papazafeiropoulos, T. Ratnarajah, in *Proc. IEEE Wireless Commun. and Netw. Conf. (WCNC)*. Linear precoding for downlink massive MIMO with delayed CSIT and channel prediction (IEEE, Istanbul, 2014), pp. 809–914
27. A. K. Papazafeiropoulos, T. Ratnarajah, in *Proc. IEEE Conf. on Acoustics, Speech and Signal Process. (ICASSP)*. Uplink performance of massive MIMO subject to delayed CSIT and anticipated channel prediction (IEEE, Florence, 2014), pp. 3162–3165
28. H. Q. Ngo, H. Suraweera, M. Matthaiou, E. Larsson, Multipair full-duplex relaying with massive arrays and linear processing. *IEEE J. Sel. Areas Commun.* **32**(9), 1721–1737 (2014)
29. S. Zhou, G. B. Giannakis, How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over rayleigh MIMO channels? *IEEE Trans. Wireless Commun.* **3**(4), 1285–1294 (2004)
30. A. K. Papazafeiropoulos, T. Ratnarajah, Deterministic equivalent performance analysis of time-varying massive MIMO systems. *IEEE Trans. Wireless Commun.* **14**(10), 5795–5809 (2015)
31. C. Kong, C. Zhong, A. K. Papazafeiropoulos, M. Matthaiou, Zhang Z, Sum-rate and power scaling of massive MIMO systems with channel aging. *IEEE Trans. Commun.* **63**(12), 4879–4893 (2015)
32. J. Jose, A. Ashikhmin, T. L. Marzetta, S. Vishwanath, Pilot contamination and precoding in multi-cell TDD systems. *IEEE Trans. Wireless Commun.* **10**(8), 2640–2651 (2011)
33. W. Liu, A. Zappone, C. Yang, E. Jorswieck, in *Proc. IEEE Int. Works. on Signal Process in Advances in Wireless Commun. (SPAWC)*. Global EE optimization of massive MIMO systems (IEEE, Stockholm, 2015), pp. 221–225
34. B. R. Marks, G. P. Wright, A general inner approximation algorithm for nonconvex mathematical programs. *Oper. Res.* **26**(4), 681–683 (1978)
35. J. Papandriopoulos, J. S. Evans, Scale: a low-complexity distributed protocol for spectrum balancing in multiuser DSL networks. *IEEE Trans. Inf. Theory.* **55**(8), 3711–3724 (2009)
36. A. Zappone, E. Jorswieck, Energy efficiency in wireless networks via fractional programming theory. *Found. Trends Commun. Inf. Theory.* **11**(3–4), 185–396 (2015)
37. W. Dinkelbach, On nonlinear fractional programming. *Manag. Sci.* **13**(7), 492–498 (1967)
38. S. Boyd, L. Vandenberghe, *Convex optimization*. (Cambridge University Press, Cambridge, United Kingdom, 2004)
39. W. Yu, R. Lui, Dual methods for nonconvex spectrum optimization of multicarrier systems. *IEEE Trans. Commun.* **54**(7), 1310–1322 (2006)
40. H. Zhang, Y. Liu, M. Tao, Resource allocation with subcarrier pairing in OFDMA two-way relay networks. *IEEE Wireless Commun. Lett.* **1**(2), 61–64 (2012)
41. M. Naeem, K. Illanko, A. Karmokar, A. Anpalagan, M. Jaseemuddin, Decode and forward relaying for energy-efficient multiuser cooperative cognitive radio network with outage constraints. *IET Commun.* **8**(5), 578–586 (2014)
42. F. Tan, T. Lv, S. Yang, Power allocation optimization for energy-efficient massive MIMO aided multi-pair decode-and-forward relay systems. *IEEE Trans. Commun.* **65**(6), 2368–2381 (2017)
43. F. Tan, T. Lv, P. Huang, Global energy efficiency optimization for wireless-powered massive MIMO aided multiway af relay networks. *IEEE Trans. Signal Process.* **66**(9), 2384–2398 (2018)
44. K. V. Mardia, J. T. Kent, J. M. Bibby, *Multivariate analysis*. (Academic press, London, United Kingdom, 1979)
45. Inverse Wishart distribution (2018). Available <https://en.wikipedia.org/wiki/Inverse-Wishartdistribution>

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
