SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

June 2023, Vol. 66 162303:1–162303:18
https://doi.org/10.1007/s11432-022-3616-9

Performance analysis of full-duplex densely distributed MIMO with wireless backhaul

Ziqian WAN[1], Qijun PAN[1], Jiamin LI[1,2*], Pengcheng ZHU[1,2] & Dongming WANG[1,2]

[1]*National Mobile Communications Research Laboratory, Southeast University, Nanjing* 210096*, China;*
[2]*Purple Mountain Laboratories, Nanjing* 211111*, China*

Received 28 June 2022/Revised 6 September 2022/Accepted 11 November 2022/Published online 18 May 2023

Abstract In this paper, the spectral efficiency of a full-duplex (FD) densely distributed multi-input multi-output (MIMO) system with wireless backhaul is considered. In-band full-duplex (IBFD) backhauling, in which the backhaul and access transmissions take place on the same spectrum, is exploited for wireless backhauling to enable an efficient spectrum reuse. However, the severe cross-tier interference and cross-link interference reduce the gains of IBFD backhauling. To evaluate the achievable spectral efficiency with imperfect channel state information (CSI), we propose a two-phase channel estimation scheme to estimate the CSI for two wireless links, and the scheme estimates an effective interference CSI between access points (APs) based on beamforming training to perform interference cancelation at APs. Given the estimated CSI, the closed-form expressions of the uplink and downlink achievable rates with maximum ratio transmission beamforming and maximum ratio combining receivers, respectively, are derived with Gamma approximation. Numerical results verify the accuracy of the derived closed-form expressions and the effectiveness of the two-phase channel estimation scheme for interference cancelation. Moreover, compared with half-duplex densely distributed MIMO systems, FD systems with interference cancelation have a better performance.

Keywords full-duplex, wireless backhaul, performance analysis, beamforming training, interference cancelation

Citation Wan Z Q, Pan Q J, Li J M, et al. Performance analysis of full-duplex densely distributed MIMO with wireless backhaul. Sci China Inf Sci, 2023, 66(6): 162303, https://doi.org/10.1007/s11432-022-3616-9

1 Introduction

The explosive demand for high-data-rate services poses great challenges to wireless communication systems. Densely distributed multi-input multi-output (MIMO) networks have been gaining attention and are considered a critical technology to meet the increasing demand for wireless data [1, 2]. In a densely distributed MIMO system, high-density and low-cost access points (APs) are arbitrarily developed to serve several terminals. The cooperation between APs potentially achieves a good system performance due to the increasing macro diversity and proximity gains [3–6]. The traditional distributed MIMO system largely relies on high-speed backhaul links, which exploit wired connectivity (e.g., optical fiber and digital subscriber line connections). However, the massive implementation of wired backhaul connections seems infeasible for a large number of APs in distributed massive MIMO systems. The reason is the prohibitively high cost, possibly hard-to-reach locations of APs (e.g., APs developed on roof-tops, and street fixtures), and less flexibility for system configurations [7]. In this study, wireless backhauls are taken into account to achieve a flexible development of high-density APs. The concept of wireless backhaul is associated with relay systems, in which a source signal can indirectly transfer a date to the destination via relays, which act as nodes [8]. Although wired backhaul can achieve high reliability and required data rates, a wireless backhaul system is still advantageous in terms of flexibility and scalability because APs can be easily added or relocated within the cell without the need to set up costly wired backhaul connections [9].

* Corresponding author (email: jiaminli@seu.edu.cn)

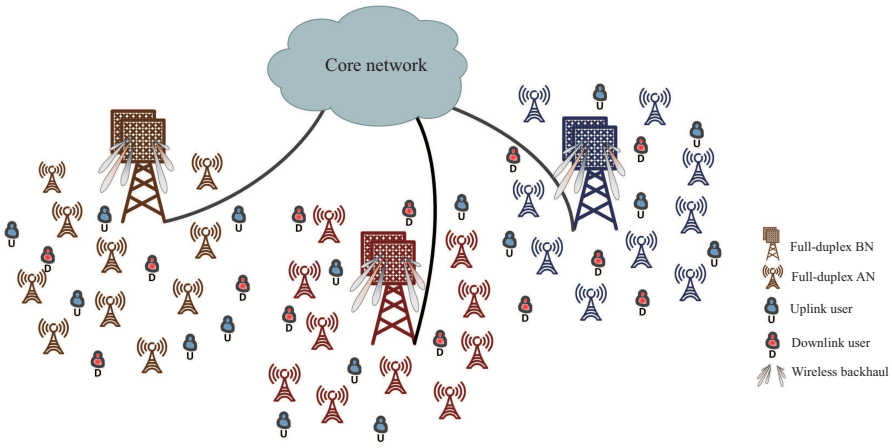

Figure 1 (Color online) System model.

To achieve efficient spectral reuse, full-duplex (FD) transmission has been considered a viable solution for 5G networks. FD networks can potentially double the spectral efficiency in comparison with half-duplex (HD) networks. To further enable an efficient spectrum reuse in wireless backhaul systems, in-band full-duplex (IBFD) backhauling can be exploited. In IBFD backhauling, the simultaneous transmission of access and backhaul information can be implemented within the same frequency band. Specifically, APs can provide simultaneous uplink (UL) and downlink (DL) transmissions for users within the same resource. At the same time, APs can transmit/receive data in that frequency band to/from macro base stations [10, 11]. However, the gains of IBFD communication are subjected to additional interference sources.

Self-interference (SI) and cross-link interference (CLI) are two major interferences in IBFD systems. SI is generated by a transmitter to its own collocated receiver [12], and CLI exists because UL communication suffers from inter-user interference, which is caused by APs serving DL users with the same time/frequency resource. Fortunately, a significant progress has been made to mitigate SI and CLI with a combination of analog and digital cancelation techniques. Based on the experimental results in [13], SI can be suppressed and reduced close to the level of noise in low-power, by using devices with recently developed digital baseband and antenna technologies. In [14], a CLI cancelation scheme is proposed and verified in a massive MIMO system. A CLI cancelation method based on beamforming training is introduced in [1], and the method is feasible to the system we investigated here. Despite the negative impacts of SI and CLI, recent advances have shown that IBFD backhauling is feasible, alleviates the need to allocate an extra spectrum for backhauling, and facilitates hardware implementation [15].

Combined with the attractive advantages of IBFD backhauling and densely distributed massive MIMO, we consider an FD densely distributed MIMO system with wireless backhaul in this study. In such a system, multiple FD-mode APs are divided into access nodes (ANs) and backhaul nodes (BNs). ANs operate like decode-and-forward relays for users and forward UL traffic to BNs over wireless backhauls, whereas BNs mainly perform baseband signal processing. Specifically, when ANs receive UL data/DL data, they decode and precode the data and then forward to BN/DL users. An implementation is shown in Figure 1. Compared with the traditional MIMO system which is still based on cellular architecture, the dense distributed development of ANs can achieve an ideal physical isolation performance in the propagation domain and control the performance loss caused by SI. Moreover, ANs can serve users collaboratively, which will efficiently improve their quality of service. Similar to a cell-free massive MIMO system, which has attracted considerable attention in recent years [16], there are no clear boundaries in the architecture we investigated, and such an architecture is highly similar to a scalable cell-free massive MIMO system. Nonetheless, there are also some differences between the two systems. First, by introducing the processing unit BN, the system investigated can execute part of the data processing work in BN, which is different from a cell-free system whose all data are transferred to the CPU for processing. Second, we consider the wireless backhaul technology, which can achieve an infinite expansion of the radio access network and comprehensive coverage of the system. In addition, there are less restrictions on ANs' configuration. That is to say, compared with the cell-free massive MIMO system, the proposed architecture is more flexible and scalable.

A large body of research has been conducted on the performance of distributed MIMO wireless backhaul networks with FD technologies. In [17], the authors derived the exact and asymptotic expressions for area spectral efficiencies in cellular networks with the FD technique. To further improve the performance of self-backhauling FD radio access systems, the transmit power optimization and feasibility were studied in [18]. Meanwhile, Ref. [19] studied the DL performance of FD self-backhauling small cells and derived analytical expressions for the coverage and average DL rate of a two-tier heterogeneous network. However, these studies did not consider the effect of imperfect channel state information (CSI), which can reduce the spectral efficiency in massive MIMO systems. Hence, they assumed that backhaul and access links have perfect CSI and did not investigate the way to estimate the CSI of two successive wireless links, which is a critical problem in wireless backhaul systems. Moreover, the recent investigation did not illustrate how to cancel the CLI interference between ANs but only based on the assumption that the interference could be eliminated.

To take full advantage of the benefits provided by wireless backhaul massive MIMO systems and FD technologies, accurate CSI is necessary at core networks and terminals. The reason is twofold. First, the CSI of access and backhaul links is a prerequisite for the core network to design beamforming and receivers [20]. Because there are two wireless links in networks, the imperfect CSI of one wireless link will influence the signal quality of the other one, which increases the difficulty in channel estimation. Second, the inter-ANs interference caused by the simultaneous occurrence of DL and UL transmission is the major factor limiting the UL capacity without considering SI. Therefore, interference cancelation is crucial for UL spectral efficiency in FD massive MIMO systems. To regenerate and eliminate inter-AN interference, the estimation of the interference channel matrix between ANs is necessary.

The estimation of two successive wireless links is a crucial problem in wireless backhaul systems. To estimate the CSI of the access link, the training pilot from users will be transmitted to the core network through two wireless links, so the inaccuracy CSI of the backhaul link will influence the estimation accuracy of CSI for the access link. In [21], the authors studied the achievable sum rate of FD wireless backhaul links with imperfect CSI under the assumption that ANs can perform channel estimation independently, as a result in pitting a high request to ANs. To estimate the CSI between ANs, the simple method is that each AN sends a pilot sequence to other ANs. Estimating the CSI needed is inefficient in massive MIMO systems due to the large number of antennas equipped by ANs. In [22], to reduce the DL pilot training overhead, the beamforming training method was proposed, and the benefits of beamforming training in distributed massive MIMO were investigated in [23]. Ref. [1] proposed a method of requiring the effective CSI between ANs in FD densely distributed massive MIMO systems for interference cancelation and verified the benefits of beamforming training in FD distributed massive MIMO systems. To the best of our knowledge, how to acquire CSI between ANs for interference cancelation has not been considered in wireless backhaul systems.

In this study, we used a two-phase channel estimation scheme to successively estimate the CSI of two wireless links and acquire effective CSI to cancel inter-AN interference in networks. The scheme can acquire CSI through UL training without requiring extra functions of AN. Moreover, beamforming training is improved to estimate the effective interference CSI, defined as the inner products of beamforming and channel vectors, between ANs. Further details will be given in Section 3. Under the assumption that the backhaul and access links are unrelative, the pilot consumption can be reduced. Moreover, the overhead of beamforming training can be significantly reduced because the dimension of effective CSI is proportional to the number of DL users. Specifically, the main contributions of our paper can be summarized as follows.

First, we model the UL and DL transmission signals of FD densely distributed massive MIMO system with wireless backhaul. A two-tier network structure is considered in which multiple BNs are deployed to provide simultaneous backhaul to multiple ANs. Each AN configures in the IBFD backhaul mode. The IBFD backhaul mode leads to numerous interferences, i.e., inter-tier interference, SI, and CLI.

Second, we propose a two-phase channel estimation method to acquire the CSI of two wireless links (backhaul and access links) successively and effective CSI between ANs. During the first phase, the core network estimates the CSI between BNs and ANs, and then estimates the CSI between all ANs and terminals based on the estimated CSI of backhaul links. During the second phase, the improved beamforming training is adopted, which enables the core network to estimate the effective CSI between ANs. Based on the estimated CSI, the core network can perform interference cancelation during the data transmission phase.

Third, with the estimated CSI, we analyze the UL and DL achievable rates and provide closed-form

expressions for the rates with the maximum-ratio combining (MRC) receivers and maximum-ratio transmission (MRT) beamforming. These analytical expressions enable the analysis of system spectral efficiency with the two-phase channel estimation scheme. Also, the benefits of the improved beamforming training is shown by the derived expressions.

Forth, simulations are performed to verify the closed-form expressions, and based on the analytical expressions, insightful conclusions are drawn according to the comparison of the spectral efficiency with and without the two-phase channel estimation method and the comparison of FD and HD densely distributed massive MIMO systems. The impact of beamforming training on the spectral efficiency is also shown.

The rest of this paper is organized as follows. In Section 2, we describe the system model, including channel model and signal transmission models. In Section 3, we describe the two-phase channel estimation scheme. In Section 4, the closed-form expressions of the achievable rates of access and backhaul links are introduced. In Section 5, the simulation results and discussions are presented. Lastly, in Section 6, our contributions are summarized and the conclusion is drawn.

Notation. Vectors and matrices are denoted by boldface lower-case and upper-case letters, respectively. $\boldsymbol{I}_N$ denotes an N-dimensional identity matrix. $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ represent the transpose operator and conjugate transpose operator, respectively. $|\cdot|$ and $\|\cdot\|$ represent the absolute value of a scalar and spectral norm of a matrix, respectively. $\mathbb{E}[\cdot]$, $\mathbb{E}[\cdot|\cdot]$, and $\mathrm{Cov}(\cdot,\cdot)$ denote the expectation, conditional expectation, and covariance operator, respectively. $\mathrm{diag}\{\cdot\}$ denotes a square diagonal matrix with the elements of the given vector on the main diagonal. $\otimes$ is the Kronecker product. A circularly symmetric complex Gaussian random variable x with the mean zero and variance σ^2 is defined as $x \sim \mathcal{CN}(0, \sigma^2)$. $\Gamma(k, \theta)$ denotes the Gamma distribution with parameters k and θ, and Nakagami(α, β) represents the Nakagami distribution with parameters α and β.

2 System model

In Subsection 2.1, we describe the system configuration of FD11t densely distributed MIMO system and the mathematical description of the channel model is presented in Subsection 2.2. The signal models for DL and UL transmissions are given in Subsections 2.3 and 2.4.

2.1 System configuration

We consider an FD densely distributed MIMO system with wireless backhaul network that is composed of a core network, M ANs, and L BNs. ANs jointly serve K terminals. The l-th BN offers the backhaul services to M_l ANs and each AN only receives/transfers effective information from/to one BN, which means $M = \sum_{l=1}^{L} M_l$. This configuration reduces the backhaul distance between ANs and BNs, and expands the coverage of the system. In such system, ANs transfer data for BNs and terminals, BNs perform the signal processing from ANs and core networks. Assume that users work in HD mode and there are K_{U} UL users and K_{D} DL users in the area, where $K = K_{\mathrm{U}} + K_{\mathrm{D}}$. Assuming that both ANs and BNs work in FD mode, ANs are configured with $M_{\mathrm{t}}^{\mathrm{A}}$ transmitting antennas and $M_{\mathrm{r}}^{\mathrm{A}}$ receiving antennas, BNs are configured with $M_{\mathrm{t}}^{\mathrm{B}}$ transmitting antennas and $M_{\mathrm{r}}^{\mathrm{B}}$ receiving antennas. All users are configured with a single antenna. An implementation of the system is shown in Figure 1.

2.2 Channel model

The channel vectors between users and ANs as well as ANs and BNs are assumed to be frequency-flat fading. In DL transmission, the l-th BN transmits the data to all ANs it serves through backhaul link and then ANs jointly transfer the signal to the corresponding k-th DL user. The channel vector between the l-th AN and k-th DL user is modeled as

$$\boldsymbol{g}_{a_l,d_k} = \left[\sqrt{\lambda_{a_l,1,d_k}}\boldsymbol{h}_{a_l,1,d_k}^{\mathrm{T}}, \ldots, \sqrt{\lambda_{a_l,M_l,d_k}}\boldsymbol{h}_{a_l,M_l,d_k}^{\mathrm{T}}\right]^{\mathrm{T}}, \tag{1}$$

where $\lambda_{a_l,m_l,d_k} = d_{a_l,m_l,d_k}^{-\alpha}$ represents the large-scale fading between the m_l-th AN served by l-th BN and DL user k. d_{a_l,m_l,d_k} represents the corresponding distance, α is the path loss exponent, and $\boldsymbol{h}_{a_l,m_l,d_k} \in \mathbb{C}^{M_{\mathrm{t}}^{\mathrm{A}} \times 1} \sim \mathcal{CN}(0, \boldsymbol{I}_{M_{\mathrm{t}}^{\mathrm{A}}})$ is the small-scale fast fading.

Similarly, in DL backhaul transmission, the channel vector between l-th BN and ANs it serves $\boldsymbol{g}_{b_l,a_l} \in \mathbb{C}^{M_{\mathrm{t}}^{\mathrm{B}} \times M_l M_{\mathrm{t}}^{\mathrm{A}}}$ is modeled as

$$\boldsymbol{g}_{b_l,a_l} = \left[\sqrt{\lambda_{b_l,a_l,1}}\boldsymbol{H}_{b_l,a_l,2}^{\mathrm{T}}, \ldots, \sqrt{\lambda_{b_l,a_l,M_l}}\boldsymbol{H}_{b_l,a_l,M_l}^{\mathrm{T}}\right]^{\mathrm{T}}, \tag{2}$$

where $\boldsymbol{H}_{b_l,a_l,m_l} \in \mathbb{C}^{M_{\mathrm{t}}^{\mathrm{B}} \times M_{\mathrm{t}}^{\mathrm{A}}}$ is also the small-scale fast fading. To facilitate the performance analysis, divide the channel vector according to the antenna of AN, which is

$$\boldsymbol{g}_{a_l,b_l,m_l} = \begin{bmatrix} g_{11}^{m_l} & g_{12}^{m_l} & \cdots & g_{1M_{\mathrm{t}}^{\mathrm{A}}}^{m_l} \\ \vdots & \vdots & & \vdots \\ g_{M_{\mathrm{t}}^{\mathrm{B}}1}^{m_l} & \cdots & \cdots & g_{M_{\mathrm{r}}^{\mathrm{B}}M_{\mathrm{t}}^{\mathrm{A}}}^{m_l} \end{bmatrix}. \tag{3}$$

Since large-scale fading changes slowly and is considered constant over long periods of time, we assume that the large-scale fading is known at ANs and BNs [24].

2.3 Downlink data transmission

During DL data transmission, all ANs transmit data to DL users while receiving messages from the BNs they connected with in backhaul simultaneously using the same frequency band.

(1) Access link. The received signal at k-th DL user can be expressed as

$$y_{d_k} = D_{a,d_k} + \sum_{j=1,j\neq k}^{K_{\mathrm{D}}} \mathcal{I}_{a,d_j} + \sum_{i=1}^{L} \mathcal{I}_{a_i,b_i} + \sum_{i=1}^{L} \mathcal{I}_{b_i,a_i} + \sum_{i=1}^{K_{\mathrm{U}}} \mathcal{I}_{u_i,d_k} + n_{d_k}, \tag{4}$$

where $D_{a,d_k} = \sqrt{p_{a,d_k}}\boldsymbol{g}_{a,d_k}^{\mathrm{H}}\boldsymbol{w}_{a,d_k}x_{a,d_k}$ represents the desired signal for DL user, $\sum_{j=1,j\neq k}^{K_{\mathrm{D}}}\mathcal{I}_{a,d_j} = \sqrt{p_{a,d_j}}\boldsymbol{g}_{a,d_j}^{\mathrm{H}}\boldsymbol{w}_{a,d_j}x_{a,d_j}$ is the co-tier interference transmitted by all ANs in the same access link but intended for other users. $\sum_{i=1}^{L}\mathcal{I}_{a_i,b_i} = \sum_{i=1}^{L}\sqrt{p_{a_i,b_i}}\boldsymbol{g}_{a_i,d_k}^{\mathrm{H}}\boldsymbol{w}_{a_i,b_i}\boldsymbol{x}_{a_i,b_i}$ and $\sum_{i=1}^{L}\mathcal{I}_{b_i,a_i} = \sum_{i=1}^{L}\sqrt{p_{b_i,a_i}}\boldsymbol{g}_{b_i,d_k}^{\mathrm{H}}\boldsymbol{w}_{b_i,a_i}\boldsymbol{x}_{b_i,a_i}$ are cross-tier interference transmitted by ANs and BNs over backhual link, respectively. $\sum_{i=1}^{K_{\mathrm{U}}}\mathcal{I}_{u_i,d_k} = \sum_{i=1}^{K_{\mathrm{U}}}\sqrt{p_{u_i,d_k}}\boldsymbol{g}_{u_i,d_k}^{\mathrm{H}}x_{u_i}$ represents the interference transmitted by UL users and the last term is the noise.

The symbols $\boldsymbol{g}_{a,d_k}^{\mathrm{H}}$, $\boldsymbol{g}_{a_l,d_k}^{\mathrm{H}}$, $\boldsymbol{g}_{b_l,d_k}^{\mathrm{H}}$ and $\boldsymbol{g}_{u_i,d_k}^{\mathrm{H}}$ represent the channel vector between all ANs and all DL users, ANs connected with l-th BN and all DL users, BN and DL users, the channel between UL user i and DL user k, respectively. $\boldsymbol{w}_{a,d_k}$ is the precode vector for DL users, adopting MRT beamforming scheme, for instance, $\boldsymbol{w}_{a,d_k} = \frac{\hat{\boldsymbol{g}}_{a,d_k}}{\|\hat{\boldsymbol{g}}_{a,d_k}\|}$. p_{a,d_k} is the transmission power from AN to k-th DL user. x_{a,d_k} ($\boldsymbol{x}_{b_i,a_i}$) represents the data (vector) transmitted by transmitter with $\mathbb{E}\{x_{a,d_k}x_{a,d_k}^{\mathrm{H}}\} = 1$ or $\mathbb{E}\{\boldsymbol{x}_{b_i,a_i}\boldsymbol{x}_{b_i,a_i}^{\mathrm{H}}\} = \boldsymbol{I}_{M_{\mathrm{t}}^{\mathrm{A}}}$ and $\mathbb{E}\left\{\boldsymbol{x}_k\boldsymbol{x}_{k'}^{\mathrm{H}}\right\} = 0$. $n_{\mathrm{dl}} \sim \mathcal{CN}(0,\sigma_{d_k}^2)$ is the complex additive white Gaussian noise (AWGN).

(2) Backhaul link. The received signal at AN served by l-th BN can be expressed as

$$y_{b_l,a_l} = \mathcal{D}_{b_l,a_l} + \sum_{i=1,i\neq l}^{L}\mathcal{I}_{b_i,a_i} + \sum_{k=1}^{K_{\mathrm{D}}}\mathcal{I}_{a,d_k}^{b} + \sum_{i=1}^{L}\mathcal{I}_{a_i,b_i} + \sum_{k=1}^{K_{\mathrm{U}}}\mathcal{I}_{u_k}^{b} + n_{b_l}, \tag{5}$$

where $\mathcal{D}_{b_l,a_l} = \sqrt{p_{b_l,a_l}}\boldsymbol{v}_{b_l,a_l}\boldsymbol{g}_{b_l,a_l}^{\mathrm{H}}\boldsymbol{w}_{b_l,a_l}\boldsymbol{x}_{b_l,a_l}$ represents the desired signal AN receives, $\sum_{i=1,i\neq l}^{L}\mathcal{I}_{b_i,a_i} = \sum_{i=1,i\neq l}^{L}\sqrt{p_{b_i,a_i}}\boldsymbol{v}_{b_l,a_l}\boldsymbol{g}_{b_i,a_l}^{\mathrm{H}}\boldsymbol{w}_{b_i,a_i}\boldsymbol{x}_{b_i,a_i}$ represents the co-tier interference caused by the signals from BNs not connected with this AN. $\sum_{k=1}^{K_{\mathrm{D}}}\mathcal{I}_{a,d_k}^{b} = \sum_{k=1}^{K_{\mathrm{D}}}\sqrt{p_{a,d_k}}\boldsymbol{v}_{b_l,a_l}\boldsymbol{g}_{a,a_l}^{\mathrm{H}}\boldsymbol{w}_{a,d_k}x_{a,d_k}$ is the cross-tier interference caused by ANs over the access link. $\sum_{i=1}^{L}\mathcal{I}_{a_i,b_i} = \sum_{i=1}^{L}\sqrt{p_{a_i,b_i}}\boldsymbol{v}_{b_l,a_l}\boldsymbol{g}_{a_i,a_l}^{\mathrm{H}}\boldsymbol{w}_{a_i,b_i}\boldsymbol{x}_{a_i,b_i}$ is the interference transmitted by other ANs for UL backhaul transmission. $\sum_{k=1}^{K_{\mathrm{U}}}\mathcal{I}_{u_k}^{b} = \sum_{k=1}^{K_{\mathrm{U}}}\sqrt{p_{u_k,a}}\boldsymbol{v}_{b_l,a_l}\boldsymbol{g}_{u_k,a_l}^{\mathrm{H}}x_{u_k,a}$ represents the interference caused by UL users and the last part is the noise. Note that the SI and CLI are all contained in (5). During spectral efficiency analysis, the corresponding interference cancelation scheme is used to eliminate the impact of SI and CLI [25, 26].

Specifically, similar to access link signal models, $\boldsymbol{g}_{b_i,a_l}^{\mathrm{H}}$, $\boldsymbol{g}_{a_i,a_l}^{\mathrm{H}}$ and $\boldsymbol{g}_{u_k,a_l}^{\mathrm{H}}$ represent the channel vector between i-th BN and ANs connected with l-th BN, the channel vector between ANs connected with i-th BN and ANs connected with l-th BN, and the channel vector between AN connected with l-th BN and UL user k, respectively. $\boldsymbol{w}_{b_l,a_l}$ represents the precode vector, $\boldsymbol{v}_{b_l,a_l}$ is the receiver vector, other parameters are defined the same as access link.

2.4 Uplink data transmission

During UL transmission, ANs receive data from UL users in access link. Simultaneously, ANs transmit the received signal to BN they connected with using the same frequency band.

(1) Access link. The received signal at ANs can be expressed as

$$y_{u_k} = \mathcal{D}_{u_k} + \sum_{j=1,j\neq k}^{K_{\mathrm{U}}} \mathcal{I}_{u_j,a} + \sum_{j=1}^{K_{\mathrm{D}}} \mathcal{I}_{a,d_j} + \sum_{l=1}^{L} \mathcal{I}_{a_l,b_l} + \sum_{l=1}^{L} \mathcal{I}_{b_l,a_l} + n_{u_k}, \tag{6}$$

where $\mathcal{D}_{u_k} = \sqrt{p_{u_k,a}}\boldsymbol{v}_{u_k,a}\boldsymbol{g}_{u_k,a}^{\mathrm{H}} x_{u_k,a}$ represents the desired signal from k-th UL user. $\sum_{j=1,j\neq k}^{K_{\mathrm{U}}} \mathcal{I}_{u_j,a} = \sum_{j=1,j\neq k}^{K_{\mathrm{U}}} \sqrt{p_{u_j,a}}\boldsymbol{v}_{u_k,a}\boldsymbol{g}_{u_j,a}^{\mathrm{H}} x_{u_j,a}$ is the co-tier interference caused by other UL users in the same access link. $\sum_{j=1}^{K_{\mathrm{D}}} \mathcal{I}_{a,d_j} = \sum_{j=1}^{K_{\mathrm{D}}} \sqrt{p_{a,d_j}}\boldsymbol{v}_{u_k,a}\boldsymbol{g}_{a,a}^{\mathrm{H}}\boldsymbol{w}_{a,d_j} x_{a,d_j}$ is the interference when ANs transmit DL data to downlink users. $\sum_{l=1}^{L} \mathcal{I}_{a_l,b_l} = \sum_{l=1}^{L} \sqrt{p_{a_l,b_l}}\boldsymbol{v}_{u_k,a}\boldsymbol{g}_{a_l,a}^{\mathrm{H}}\boldsymbol{w}_{a_l,b_l}\boldsymbol{x}_{a_l,b_l}$ and $\sum_{l=1}^{L} \mathcal{I}_{b_l,a_l} = \sum_{l=1}^{L} \sqrt{p_{b_l,a_l}}\boldsymbol{v}_{u_k,a}\boldsymbol{g}_{b_l,a}^{\mathrm{H}}\boldsymbol{w}_{b_l,a_l}\boldsymbol{x}_{b_l,a_l}$ are cross-tier interference caused by ANs and BNs over backhaul link similar to (4) and the last term is the noise.

The symbols $\boldsymbol{g}_{a,a}^{\mathrm{H}}$, $\boldsymbol{g}_{a_l,a}^{\mathrm{H}}$ and $\boldsymbol{g}_{b_l,a}^{\mathrm{H}}$ represent the channel matrix between ANs, the channel vector between l-th region and all ANs in the area, and the channel vector between l-th BN and ANs, respectively. $\boldsymbol{v}_{u_k,a}$ is the receiver vector at ANs and BNs. In this study, MRC receivers are adopted, where $\boldsymbol{v}_{u_k,a} = \frac{\hat{\boldsymbol{g}}_{u_k,a}}{\|\hat{\boldsymbol{g}}_{u_k,a}\|}$, other parameters are defined the same as DL access link.

Comparing (5) and (6), the received signal at ANs is the same, while the desired signal and interference are different when analyzing different wireless links. In Section 4, the achievable rates of both AN and UL users are given.

(2) Backhaul link. The received signal at l-th BN can be expressed as

$$y_{a_l,b_l} = \mathcal{D}_{a_l,b_l} + \sum_{i=1,i\neq l}^{L} \mathcal{I}_{a_i,b_i} + \sum_{k=1}^{K_{\mathrm{D}}} \mathcal{I}_{a,d_k} + \sum_{i=1}^{L} \mathcal{I}_{a_i,b_i} + \sum_{k=1}^{K_{\mathrm{U}}} \mathcal{I}_{u_k} + n_{b_l}, \tag{7}$$

where $\mathcal{D}_{a_l,b_l} = \sqrt{p_{a_l,b_l}}\boldsymbol{v}_{a_l,b_l}\boldsymbol{g}_{a_l,b_l}^{\mathrm{H}}\boldsymbol{w}_{a_l,b_l}\boldsymbol{x}_{a_l,b_l}$ represents the desired signal from ANs l-th BN served, $\sum_{i=1,i\neq l}^{L} \mathcal{I}_{a_i,b_i} = \sum_{i\neq l} \sqrt{p_{a_i,b_i}}\boldsymbol{v}_{a_l,b_l}\boldsymbol{g}_{a_i,b_l}^{\mathrm{H}}\boldsymbol{w}_{a_i,b_i}\boldsymbol{x}_{a_i,b_i}$ is the interference caused by ANs l-th BN not connected with. $\sum_{k=1}^{K_{\mathrm{D}}} \mathcal{I}_{a,d_k} = \sum_{k=1}^{K_{\mathrm{D}}} \sqrt{p_{a,d_k}}\boldsymbol{v}_{a_l,b_l}\boldsymbol{g}_{a,b_l}^{\mathrm{H}}\boldsymbol{w}_{a,d_k} x_{a,d_k}$ is the cross-tier interference caused by ANs when transmitting signals to DL users. $\sum_{i=1}^{L} \mathcal{I}_{a_i,b_i} = \sum_{i=1}^{L} \sqrt{p_{b_i,a_i}}\boldsymbol{v}_{a_l,b_l}\boldsymbol{g}_{b_i,b_l}^{\mathrm{H}}\boldsymbol{w}_{b_i,a_i}\boldsymbol{x}_{b_i,a_i}$ is the co-tier interference over backhaul link, $\sum_{k=1}^{K_{\mathrm{U}}} \mathcal{I}_{u_k} = \sum_{k=1}^{K_{\mathrm{U}}} \sqrt{p_{u_k,a}}\boldsymbol{v}_{a_l,b_l}\boldsymbol{g}_{u_k,b_l}^{\mathrm{H}} x_{u_k,a}$ is the interference transmitted by UL users and the last part is the noise. Specifically, similar to access link signal models, $\boldsymbol{w}$ represents the beamforming vector, $\boldsymbol{v}_{a_l,b_l}$ is the receiver vector, $\boldsymbol{g}_{u_k,b_l}^{\mathrm{H}}$ is the channel vector between UL user k and l-th BN, other parameters are defined as the access link.

Remark 1. The signal transmission model we discussed is universal. Each AN and BN works in FD mode and performs UL and DL transmission simultaneously. In practice, most APs only perform UL receiving or DL transmitting in a time slot. Therefore, the signal model we considered is the most complex condition in FD systems, which is meaningful and universal to the performance analysis of multi-tier systems with the FD technique.

As discussed above, there are two wireless links in UL/DL transmission. In order to get CSI for both backhaul and access link and also to acquire effective CSI between ANs for interference cancelation, we propose a two-phase channel estimation scheme. In Section 3, the two-phase channel estimation scheme is introduced.

3 Channel estimation

In this section, a two-phase channel estimation scheme is proposed. Assume that the system adopts time division duplexing (TDD) mode. In the first phase, the first τ_{p} symbols are used for uplink training. Specifically, the core network estimates the CSI of backhaul link, and then based on the CSI of backhaul link, the access link is estimated. In the second phase, the following τ_1 symbols are used for DL channel estimation where the improved beamforming training scheme is adopted to help ANs to estimate the effective CSI between ANs. Finally, the remaining symbols are used for UL and DL data transmission. The uplink training and downlink channel estimation are presented in Subsections 3.1 and 3.2.

3.1 Uplink training

During the uplink training phase, all ANs send pilot sequences of length $\tau_{\mathrm{p}}^{\mathrm{B}}$ to BN they connect with. BNs transfer the received pilot sequences to the core network for backhaul channel estimation. Simultaneously, all terminals transmit pilot sequences of length $\tau_{\mathrm{p}}^{\mathrm{A}}$ to ANs and then transfer to BNs via wireless backhaul links for estimation. Noting that when estimating the CSI of access link, the CSI for backhaul link has been known at the core network; therefore, the quality of the CSI for access link depends on the quality of the CSI for backhaul link. The minimum mean-square error (MMSE) channel estimation is used in uplink training and downlink channel estimation. To simplify the analysis, the impact of pilot contamination is not considered. Due to the relatively fixed position and long coherence time between ANs and BNs, the frequent channel estimation is not required [24].

(1) Backhaul link. Let $\varphi_{k,n}\in\mathbb{C}^{\tau_{\mathrm{p}}^{\mathrm{B}}\times 1}$ represent the pilot sequence sent from k-th AN's n-th antenna to BN with $\|\varphi_{k,n}\|^2=1,\forall k=1,\ldots,M_l,n=1,\ldots,M_{\mathrm{t}}^{\mathrm{A}}$. The length of pilot sequence satisfies $\tau_{\mathrm{p}}^{\mathrm{B}}\geqslant M_l M_{\mathrm{t}}^{\mathrm{A}}$. The pilots received by BN are

$$\boldsymbol{Y}_{\mathrm{p}}=\sqrt{\tau_{\mathrm{p}}^{\mathrm{B}}p_{\mathrm{p}}}\sum_{k=1}^{M_l}\sum_{n=1}^{M_{\mathrm{t}}^{\mathrm{A}}}g_{k,n}\varphi_{k,n}^{\mathrm{H}}+\boldsymbol{N}_{\mathrm{p}},\tag{8}$$

where p_{p} is the transmit power of pilot sequence, $g_{k,n}$ is the channel between n-th antenna on k-th AN to BN, which is the $(k-1)M_{\mathrm{t}}^{\mathrm{A}}+n$-th column of $\boldsymbol{g}_{a_l,b_l}$. $\boldsymbol{N}_{\mathrm{p}}\in{}^{M_{\mathrm{r}}^{\mathrm{B}}\times\tau_{\mathrm{p}}^{\mathrm{B}}}$ is the complex AWGN vector with i.i.d. $\mathcal{CN}\left(0,\sigma_p^2\right)$ elements.

In order to estimate the channel response for each antenna of each AN, the received pilot sequence $\boldsymbol{Y}_{\mathrm{p}}$ is first projected onto $\varphi_{k,n}$,

$$\breve{y}_{\mathrm{p},k,n}\triangleq\boldsymbol{Y}_{\mathrm{p}}\varphi_{k,n}=\sqrt{\tau_{\mathrm{p}}^{\mathrm{B}}p_{\mathrm{p}}}g_{k,n}+n_{\mathrm{p},k,n}.\tag{9}$$

The MMSE estimate of the channel response $g_{k,n}$ is

$$\hat{g}_{k,n}=\mathbb{E}\left\{g_{k,n}\breve{y}_{\mathrm{p},k,n}\right\}\left(\mathbb{E}\left\{\breve{y}_{\mathrm{p},k,n}\breve{y}_{\mathrm{p},k,n}^{\mathrm{H}}\right\}\right)^{-1}\breve{y}_{\mathrm{p},k,n}=\frac{\sqrt{\tau_{\mathrm{p}}^{\mathrm{B}}p_{\mathrm{p}}}\lambda_k}{\tau_{\mathrm{p}}^{\mathrm{B}}p_{\mathrm{p}}\lambda_k+1}\breve{y}_{\mathrm{p},k,n}.\tag{10}$$

Let $\tilde{g}_{k,n}$ be the channel estimation error, i.e., $\tilde{g}_{k,n}=g_{k,n}-\hat{g}_{k,n}$. Due to the MMSE estimation property, $\hat{g}_{k,n}$ and $\tilde{g}_{k,n}$ are independent. Furthermore, we have that the elements of $\hat{g}_{k,n}$ are i.i.d. $\mathcal{CN}\left(0,\sigma_{\mathrm{p}}^2\right)$, where $\beta_k^2\triangleq\frac{\tau_{\mathrm{p}}^{\mathrm{B}}p_{\mathrm{p}}\lambda_k^2}{\tau_{\mathrm{p}}^{\mathrm{B}}p_{\mathrm{p}}\lambda_k+\sigma^2}$, and the elements of $\tilde{g}_{k,n}$ are i.i.d. $\mathcal{CN}\left(0,\lambda_k-\beta_k^2\right)$. σ^2 is the variance of estimation noise. To sum up, the estimated channel between ANs and BN is $\hat{\boldsymbol{g}}_{a_l,b_l}=[\beta_{a_l,b_l,1}\hat{\boldsymbol{H}}_{a_l,b_l,2}^{\mathrm{T}},\ldots,\beta_{a_l,b_l,M_l}\hat{\boldsymbol{H}}_{a_l,b_l,M_l}^{\mathrm{T}}]^{\mathrm{T}}$.

(2) Access link. For the access link training phase, ANs receive pilot sequences from all users and then transfer the sequences to BNs through wireless backhaul and finally the CSI of access link is estimated in the core network. As discussed above, the channel estimation of access link is based on the CSI of backhaul link. In this study, assuming that the channel estimation quality of backhaul link is not ideal and the estimation error of backhaul link is represented as $\tilde{\boldsymbol{g}}_{\mathrm{B}}$, $\tilde{\boldsymbol{g}}_{\mathrm{B}}$ satisfies i.i.d. $\mathcal{CN}\left(0,\lambda_k-\beta_k^2\right)$.

The estimated channel between k-th UL user and ANs connected with l-th BN $\hat{\boldsymbol{g}}_{a_l,u_k}$ and the estimated channel between k-th DL user and ANs connected with l-th BN $\hat{\boldsymbol{g}}_{a_l,d_k}$ can be respectively expressed as

$$\hat{\boldsymbol{g}}_{a_l,u_k}=\left[\beta_{a_l,1,u_k}\hat{h}_{a_l,1,u_k}^{\mathrm{T}},\ldots,\beta_{a_l,M_l,u_k}\hat{h}_{a_l,M_l,u_k}^{\mathrm{T}}\right]^{\mathrm{T}},\tag{11}$$

$$\hat{\boldsymbol{g}}_{a_l,d_k}=\left[\beta_{a_l,1,d_k}\hat{h}_{a_l,1,d_k}^{\mathrm{T}},\ldots,\beta_{a_l,M_l,d_k}\hat{h}_{a_l,M_l,d_k}^{\mathrm{T}}\right]^{\mathrm{T}},\tag{12}$$

where $\beta_{a_l,m,u_k/d_k}^2=\frac{\tau_{\mathrm{p}}^{\mathrm{A}}p_{\mathrm{p}}\lambda_{a_l,1,u_k/d_k}^2}{\tau_{\mathrm{p}}^{\mathrm{A}}p_{\mathrm{p}}\lambda_{a_l,1,u_k/d_k}+\sigma^2}$, $\beta_{a_l,m,u_k/d_k}$ is the equivalent large-scale fading, p_{p} is the UL pilot transmitted power, σ^2 is the variance of AWGN, $\hat{h}_{a_l,m,u_k}\triangleq[\hat{h}_{a_l,1,u_k}^{\mathrm{T}},\ldots,\hat{h}_{a_l,M_l,u_k}^{\mathrm{T}}]^{\mathrm{T}}\sim\mathcal{CN}(0,\boldsymbol{I}_{M_l M_{\mathrm{r}}^{\mathrm{A}}})$ is the equivalent small-scale fast fading component of the estimated channel.

Using the MMSE estimation property, combined with the estimation error of wireless backhaul link, access link can be decomposed into

$$\boldsymbol{g}_{a_l,u_k/d_k}=\hat{\boldsymbol{g}}_{a_l,u_k/d_k}+\tilde{\boldsymbol{g}}_{a_l,u_k/d_k}+\tilde{\boldsymbol{g}}_{\mathrm{B}},\tag{13}$$

where $\tilde{\boldsymbol{g}}_{a_l,u_k/d_k}$ is the channel estimation error, $\tilde{\boldsymbol{g}}_{a_l,u_k/d_k}$ and $\hat{\boldsymbol{g}}_{a_l,u_k/d_k}$ are independent.

According to Lemmas A1 and A2 in Appendix A, we have

$$\hat{g}^{\mathrm{H}}_{a_l,u_k/d_k}\hat{g}_{a_l,u_k/d_k} \sim \Gamma\left(\hat{k}_{a_l,u_k/d_k}, \hat{\theta}_{a_l,u_k/d_k}\right), \tag{14a}$$

$$\tilde{g}^{\mathrm{H}}_{a_l,u_k/d_k}\tilde{g}_{a_l,u_k/d_k} \sim \Gamma\left(\tilde{k}_{a_l,u_k/d_k}, \tilde{\theta}_{a_l,u_k/d_k}\right), \tag{14b}$$

where $\hat{k}_{a_l,d_k} = \frac{M^{\mathrm{A}}_{\mathrm{r}}\left(\sum_{m=1}^{M_l}\beta^2_{a_l,m,d_k}\right)^2}{\sum_{m=1}^{M_l}\beta^4_{a_l,m,d_k}}$, $\hat{\theta}_{a_l,d_k} = \frac{\sum_{m=1}^{M_l}\beta^4_{a_l,m,d_k}}{\sum_{m=1}^{M_l}\beta^2_{a_l,m,d_k}}$, and in (14b), $\tilde{k}_{a_l,d_k} = \frac{M^{\mathrm{A}}_{\mathrm{r}}\left(\sum_{m=1}^{M_l}\eta^2_{a_l,m,d_k}\right)^2}{\sum_{m=1}^{M_l}\eta^4_{a_l,m,d_k}}$, $\tilde{\theta}_{a_l,d_k} = \frac{\sum_{m=1}^{M_l}\eta^4_{a_l,m,d_k}}{\sum_{m=1}^{M_l}\eta^2_{a_l,m,d_k}}$, $\eta^2_{a_l,m,d_k} = \lambda_{a_l,m,d_k} - \beta^2_{a_l,m,d_k}$. The definition of the channel vectors between ANs and uplink users is similar.

Remark 2. The performance of access link estimation is affected by backhaul link estimation. The estimation error $\tilde{\boldsymbol{g}}_{\mathrm{B}}$ is introduced to show the degree of influence, and the estimation error of access link is recomposed using (13). We assume that the backhaul link and access link are unrelated. The location of BNs and ANs is relatively stationary and the backhaul link CSI keeps unchanged for a long period. If the core network can obtain perfect CSI for backhaul link, $\tilde{\boldsymbol{g}}_{\mathrm{B}}$ equals 0.

3.2 Downlink channel estimation

During the downlink pilot transmission phase, we adopt beamforming training to obtain the estimation of the effective interference CSI $\boldsymbol{f}_{\mathrm{d}_k} \triangleq \boldsymbol{g}_{a,a}\boldsymbol{w}_{a,d_k}$ instead of the channel matrix $\boldsymbol{g}^{\mathrm{H}}_{a,a}$ between ANs, which greatly reduces the training overhead. In this study, we consider MRT beamforming, mathematically, $\boldsymbol{w}_{a,d_k} = \frac{\hat{\boldsymbol{g}}_{a,d_k}}{\|\hat{\boldsymbol{g}}_{a,d_k}\|}$. Specifically, all ANs broadcast the pilot signals which have been multiplied with beamforming to ANs including itself, and then all ANs estimate $\boldsymbol{f}_{\mathrm{d}_k}$ jointly. The details are as follows.

First, ANs transfer the received pilots to the core networks through wireless backhual link. The received signals at the CPU can be given by

$$\boldsymbol{Y}^{\mathrm{p}} = \boldsymbol{g}_{a,a}\boldsymbol{w}\boldsymbol{P}\boldsymbol{\Phi} + \boldsymbol{N}_{\mathrm{up}}, \tag{15}$$

where $\boldsymbol{N}_{\mathrm{up}} \in \mathbb{C}^{MM^{\mathrm{A}}_{\mathrm{t}}\times\tau_1}$ is the complex AWGN matrix with i.i.d. $\mathcal{CN}(0,\sigma^2_{\mathrm{up}})$ elements. Note that the estimation error of backhual links is considered in $\boldsymbol{N}_{\mathrm{up}}$. $\boldsymbol{\Phi} \in \mathbb{C}^{K_{\mathrm{D}}\times\tau_1}$ is the downlink pilot matrix which is assumed as $\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{H}} = I_{K_{\mathrm{D}}}$, $\boldsymbol{P} = \mathrm{diag}\left(\sqrt{p_{\mathrm{dp},1}},\ldots,\sqrt{p_{\mathrm{dp},K_{\mathrm{D}}}}\right)$ is the pilot power matrix, and $p_{\mathrm{dp},k}$ is the transmitted power of the k-th pilot sequence.

After correlating the received pilot signal with $\boldsymbol{\Phi}^{\mathrm{H}}$, Eq. (15) can be rewritten as

$$\tilde{\boldsymbol{Y}}^{\mathrm{p}} = \left[\sqrt{p_{\mathrm{dp},1}}\boldsymbol{f}_1,\ldots,\sqrt{p_{\mathrm{dp},K_{\mathrm{D}}}}\boldsymbol{f}_{K_{\mathrm{D}}}\right] + \tilde{\boldsymbol{N}}_{\mathrm{up}},$$

where $\tilde{\boldsymbol{N}}_{\mathrm{up}} = \boldsymbol{N}_{\mathrm{up}}\boldsymbol{\Phi}^{\mathrm{H}}$.

Based on the above observation, the MMSE estimation of $\boldsymbol{f}_i$ can be written as

$$\hat{\boldsymbol{f}}_i = \mathbb{E}[\boldsymbol{f}_i] + \mathrm{cov}(\boldsymbol{f}_i,\tilde{\boldsymbol{y}}^{\mathrm{p}}_i)\mathrm{cov}(\tilde{\boldsymbol{y}}^{\mathrm{p}}_i,\tilde{\boldsymbol{y}}^{\mathrm{p}}_i)^{-1}(\tilde{\boldsymbol{y}}^{\mathrm{p}}_i - \mathbb{E}[\tilde{\boldsymbol{y}}^{\mathrm{p}}_i]), \tag{16}$$

where $\tilde{\boldsymbol{y}}^{\mathrm{p}}_i = \sqrt{p_{\mathrm{dp},i}}\boldsymbol{f}_i + \tilde{\boldsymbol{n}}_{\mathrm{up},i}$ is the i-th column of $\tilde{\boldsymbol{Y}}^{\mathrm{p}}$, and $\tilde{\boldsymbol{n}}_{\mathrm{up},i}$ is the i-th column of $\tilde{\boldsymbol{N}}_{\mathrm{up}}$.

Based on the derived progress in [1], the closed-form expressions for (16) with MRT beamforming are given in the following proposition.

Proposition 1. With MRT beamforming, the estimation of $\boldsymbol{f}_i$ can be given by

$$\hat{\boldsymbol{f}}_i = \sqrt{p_{\mathrm{dp},i}}\boldsymbol{D}(p_{\mathrm{dp},i}\boldsymbol{D} + \sigma^2_{\mathrm{up}}\boldsymbol{I}_{M^{\mathrm{A}}_{\mathrm{t}}M})^{-1}\tilde{\boldsymbol{y}}^{\mathrm{p}}_i, \tag{17}$$

which can be rewritten as

$$\hat{\boldsymbol{f}}_i = \left[\rho_{1,i}\hat{\boldsymbol{q}}^{\mathrm{T}}_{1,i},\ldots,\rho_{M_{\mathrm{ul}},i}\hat{\boldsymbol{q}}^{\mathrm{T}}_{M,i}\right]^{\mathrm{T}}, \tag{18}$$

and replace the subscript a,a as I, the estimation error $\tilde{\boldsymbol{f}}_i$ is distributed as

$$\tilde{\boldsymbol{f}}_i \sim \mathcal{CN}\left(0,\mathrm{diag}\left(\delta^2_{1,i}\boldsymbol{I}_{M^{\mathrm{A}}_{\mathrm{t}}},\ldots,\delta^2_{M,i}\boldsymbol{I}_{M^{\mathrm{A}}_{\mathrm{t}}}\right)\right), \tag{19}$$

where $\boldsymbol{D} \triangleq \mathrm{diag}(\sum_{i=1}^{M}\lambda_{\mathrm{I},1,m},\ldots,\sum_{m=1}^{M}\lambda_{\mathrm{I},M,m}) \otimes \frac{1}{M}\boldsymbol{I}_{M_{\mathrm{t}}^{\mathrm{A}}}$, $\rho_{l,i} \triangleq \frac{\sqrt{p_{\mathrm{dp},i}}\sum_{m=1}^{M}\lambda_{\mathrm{I},l,m}}{\sqrt{Mp_{\mathrm{dp},i}\sum_{m=1}^{M}\lambda_{\mathrm{I},l,m}+M^2\sigma_{\mathrm{up}}^2}}$, $\delta_{l,i}^2 \triangleq \frac{1}{M}\sum_{m=1}^{M}\lambda_{\mathrm{I},l,m}-\rho_{l,i}^2$, and $\hat{\boldsymbol{q}}_{l,i} \sim \mathcal{CN}(0,\boldsymbol{I}_{M_{\mathrm{t}}^{\mathrm{A}}})$ is the equivalent small-scale fast fading. The proof refers to [1].

In this study, ANs work in IBFD mode and the downlink channel estimation based on beamforming can be used for eliminating SI and CLI.

Note that the two-phase estimation scheme is composed of uplink training and downlink channel estimation, which combines CSI estimation of two wireless links and the effective CSI estimation of interference channel between ANs. Compared with the traditional channel estimation scheme, which only estimates the CSI of access link, the channel estimation scheme proposed is more realistic because the CSI of two wireless links is imperfect. Moreover, it is efficient and novel to combine channel estimation and interference cancelation together.

4 Spectral efficiency analysis

In this section, based on the estimated CSI, the closed-form expressions of UL and DL achievable rates are derived. In Subsection 4.1, the interference cancelation is performed and the closed-form expressions of DL ergodic achievable rates with MRT beamforming are derived. In Subsection 4.2, the closed-form expressions of UL ergodic achievable rates with MRC receivers are derived.

4.1 Downlink spectral efficiency

For DL transmissions, ANs jointly send data to DL terminals while simultaneously receiving the data from BN over the wireless backhaul link. The signal transmission model is shown in (4) and (5).

According to the property of MMSE channel estimation described in Section 3, the inner product of channel estimation error and the channel estimation vector is 0 since they are independent. Therefore, the DL signal model for access link can be rewritten as

$$y_{d_k} = \sqrt{p_{a,d_k}}\hat{\boldsymbol{g}}_{a,d_k}^{\mathrm{H}}\boldsymbol{w}_{a,d_k}x_{a,d_k} + \sum_{j\neq k}^{K_{\mathrm{D}}}\sqrt{p_{a,d_j}}\boldsymbol{g}_{a,d_k}^{\mathrm{H}}\boldsymbol{w}_{a,d_j}x_{a,d_j} + \sum_{i=1}^{L}\mathcal{I}_{a_i,b_i} + \sum_{i=1}^{L}\mathcal{I}_{b_i,a_i} + \sum_{i=1}^{K_{\mathrm{U}}}\mathcal{I}_{u_i,d_k} + n_{d_k}. \tag{20}$$

The omitted parts $\sum_{i=1}^{L}\mathcal{I}_{a_i,b_i}$, $\sum_{i=1}^{L}\mathcal{I}_{b_i,a_i}$ and $\sum_{i=1}^{K_{\mathrm{U}}}\mathcal{I}_{u_i,d_k}$ are the same as (4). The downlink ergodic achievable rates of the k-th downlink user can be expressed as

$$R_{d_k} = \mathbb{E}\left[\log_2\left(1+\gamma_{d_k}^{\mathrm{mrt}}\right)\right], \tag{21}$$

where $\gamma_{d_k}^{\mathrm{mrt}}$ is SINR of the received signal. Put the expectation into $\gamma_{d_k}^{\mathrm{mrt}}$, then

$$\gamma_{d_k}^{\mathrm{mrt}} = \frac{\left|\sqrt{p_{a_l,d_k}}\hat{\boldsymbol{g}}_{a_l,d_k}^{\mathrm{H}}\boldsymbol{w}_{a_l,d_k}\right|^2}{\mathcal{V}+\mathcal{N}_{a,d_j}+\mathcal{N}_{a,b}+\mathcal{N}_{b,a}+\mathcal{N}_{u,d}+\sigma_{d_k}^2}, \tag{22}$$

where $\mathcal{V} = \mathbb{E}\{|\sqrt{p_{a_l,d_k}}(\hat{\boldsymbol{g}}_{a_l,d_k}^{\mathrm{H}}\boldsymbol{w}_{a_l,d_k} - [\hat{\boldsymbol{g}}_{a_l,d_k}^{\mathrm{H}}\boldsymbol{w}_{a_l,d_k}])|^2\}$ denotes the variance of desired signal, $\mathcal{N}_{a,d_j} = \sum_{j=1,j\neq k}^{K_{\mathrm{D}}}\mathbb{E}[|\sqrt{p_{a,d_j}}\boldsymbol{g}_{a_l,d_k}^{\mathrm{H}}\boldsymbol{w}_{a_l,d_j}|^2]$ denotes the co-tier interference caused by other DL terminals, $\mathcal{N}_{a,b} = \sum_{i=1}^{L}\mathbb{E}[|\sqrt{p_{a_i,b_i}}\boldsymbol{g}_{a_l,d_k}^{\mathrm{H}}\boldsymbol{w}_{a_i,b_i}|^2]$ and $\mathcal{N}_{b,a} = \sum_{i=1}^{L}\mathbb{E}[|\sqrt{p_{b_i,a_i}}\boldsymbol{g}_{b_l,d_k}^{\mathrm{H}}\boldsymbol{w}_{b_i,a_i}|^2]$ denote the cross-tier interference, $\mathcal{N}_{u,d} = \sum_{i=1}^{K_{\mathrm{U}}}\mathbb{E}[|\sqrt{p_{u_i,d_k}}\boldsymbol{g}_{u_i,d_k}^{\mathrm{H}}|^2]$ denotes the interference transmitted by UL terminals.

Define the average spectral efficiency of the k-th DL user as

$$\mathrm{SE}_{d_k} = \left(1-\frac{\tau_{\mathrm{p}}+\tau_1}{\tau_{\mathrm{c}}}\right)R_{d_k}, \tag{23}$$

where τ_{c} is the length of the signal sequence.

Under the MRT beamforming, we derive the closed-form expression of the DL ergodic achievable rate k-th DL user as

$$R_{d_k}^{\mathrm{mrt}} = \log_2\left(1+\gamma_{d_k}^{\mathrm{mrt}}\right), \tag{24}$$

$$\gamma_{d_k}^{\mathrm{mrt}} \simeq \frac{\xi(\hat{k}_{a,d_k})\hat{\theta}_{a,d_k}}{(\hat{k}_{a,d_k}-\xi(\hat{k}_{a,d_k}))\hat{\theta}_{a,d_k}+\frac{1}{MM_{\mathrm{t}}^{\mathrm{A}}}\tilde{k}_{a,d_k}\tilde{\theta}_{a,d_k}+\mathcal{I}_{a,d_j}+\mathcal{I}_{a,b}+\mathcal{I}_{b,a}+\mathcal{I}_{u,d}+\sigma_{d_k}^2}, \tag{25}$$

where the function in (25) is $\xi(x) \triangleq \frac{\Gamma(x+1/2)}{\Gamma(x)}$. The specific expressions of omitted parts in (25) are represented as $\mathcal{I}_{a,d_j} = p_{a,d_j}(\frac{K_{\mathrm{D}}-1}{MM_{\mathrm{t}}^{\mathrm{A}}}\hat{k}_{a,d_k}\hat{\theta}_{a,d_k}+\frac{K_{\mathrm{D}}-1}{MM_{\mathrm{t}}^{\mathrm{A}}}\tilde{k}_{a,d_k}\tilde{\theta}_{a,d_k})$, $\mathcal{I}_{a,b} = \sum_{i=1}^{L} M_i M_{\mathrm{t}}^{\mathrm{A}} p_{a_i,b_i}\lambda_{a_i,d_k}$, $\mathcal{I}_{b,a} = M_{\mathrm{t}}^{\mathrm{B}}\sum_{i=1}^{L} p_{b_i,a_i}\lambda_{b_i,d_k}$, and $\mathcal{I}_{u,d} = \sum_{i=1}^{K_{\mathrm{U}}} p_{u_i,d_k}\lambda_{u_i,d_k}$.

The proof of (24) is shown in Appendix B.

For DL transmission over wireless backhaul links, BN transmits signals to ANs it connects with. BN transmits signals with the MRT precoding scheme and ANs adopt MRC receivers, that is, $\boldsymbol{w}_{b_l,a_{l,m}} = \frac{\hat{\boldsymbol{g}}_{b_l,a_{l,m}}^{\mathrm{H}}}{\|\hat{\boldsymbol{g}}_{b_l,a_{l,m}}^{\mathrm{H}}\|}$, $\boldsymbol{v}_{b_l,a_{l,m}} = \frac{\hat{\boldsymbol{g}}_{b_l,a_{l,m}}^{\mathrm{H}}}{\|\hat{\boldsymbol{g}}_{b_l,a_{l,m}}^{\mathrm{H}}\|}$. BNs and ANs send or receive data signals with multiple antennas. According to [21], the signal sent by l-th BN and received by m-th AN it connects with can be rewritten as

$$y_{b_l,a_{l,m}} = \mathcal{D}_{b_l,a_{l,m}} + \mathcal{I}_{b_l,a_{l,m'}} + \sum_{i\neq l}^{L}\mathcal{I}_{b_i,a_i} + \sum_{k=1}^{K_{\mathrm{D}}}\mathcal{I}_{a,d_k}^{b} + \sum_{i=1}^{L}\mathcal{I}_{a_i,b_i} + \sum_{k=1}^{K_{\mathrm{U}}}\mathcal{I}_{u_k}^{b} + n_{a_l,m}, \tag{26}$$

where $\mathcal{D}_{b_l,a_{l,m}} = \sqrt{p_{b_l,a_{l,m}}}\boldsymbol{v}_{b_l,a_{l,m}}\hat{\boldsymbol{g}}_{b_l,a_{l,m}}^{\mathrm{H}}\boldsymbol{w}_{b_l,a_{l,m}}\boldsymbol{x}_{b_l,a_{l,m}}$ is the desired signal and the interference from other antennas is represented as $\mathcal{I}_{b_l,a_{l,m'}} = \sum_{m'\neq m}^{M}\sqrt{p_{b_l,a_{l,m'}}}\boldsymbol{v}_{b_l,a_{l,m}}\boldsymbol{g}_{b_l,a_{l,m}}^{\mathrm{H}}\boldsymbol{w}_{b_l,a_{l,m'}}$.

The other parts $\sum_{i\neq l}^{L}\mathcal{I}_{b_i,a_i}$, $\sum_{k=1}^{K_{\mathrm{D}}}\mathcal{I}_{a,d_k}^{b}$, $\sum_{i=1}^{L}\mathcal{I}_{a_i,b_i}$ and $\sum_{k=1}^{K_{\mathrm{U}}}\mathcal{I}_{u_k}^{b}$ are the same as (5). The signal received by m-th AN contains the cross-tier interference caused by DL ANs, which can be eliminated by an interference cancelation scheme based on beamforming training. The scheme is described in the next part specifically for it has a more significant effect on UL transmissions. With the interference cancelation scheme, CLI caused by DL users can be decoded to a level of white noise, therefore, $\sum_{k=1}^{K_{\mathrm{D}}}\mathcal{I}_{a,d_k}^{b} = \sum_{k=1}^{K_{\mathrm{D}}}\sqrt{p_{a,d_k}}\boldsymbol{v}_{b_l,a_{l,m}}\boldsymbol{g}_{a,a_{l,m}}^{\mathrm{H}}\boldsymbol{w}_{a,d_k}x_{a,d_k}$ can be seen as white noise with i.i.d. $\mathcal{CN}(0,\sigma_{\boldsymbol{I},a_{l,m}}^2)$ elements.

The ergodic achievable DL rates of m-th AN connected with l-th BN can be expressed as

$$R_m^{\mathrm{dl,mrc}} = \log_2\left|\boldsymbol{I}_{M_{\mathrm{r}}^{\mathrm{A}}} + \frac{\bar{\boldsymbol{H}}_m^{\mathrm{H}}\bar{\boldsymbol{H}}_m}{\bar{\Xi}_m - \bar{\boldsymbol{H}}_m^{\mathrm{H}}\bar{\boldsymbol{H}}_m}\right|, \tag{27}$$

where $\bar{\boldsymbol{H}}_m \triangleq \sqrt{p_{b_l,a_l}}\mathbb{E}\{\boldsymbol{v}_{b_l,a_l}\boldsymbol{g}_{b_l,a_l}^{\mathrm{H}}\boldsymbol{w}_{b_l,a_l}\}$ and $\bar{\Xi}_m \triangleq \mathcal{L}_{b_l,a_{l,m'}} + \mathcal{L}_{b_i,a_i} + \mathcal{L}_{a_i,b_i} + \mathcal{L}_{u_k,a_l} + \mathcal{L}_{\boldsymbol{I},a_{l,m}}$.

Specific expressions of each part of $\bar{\Xi}_m$ are expressed in Appendix C. Considering that the antenna number of BN is far larger than ANs', the closed-form expression of the ergodic achievable DL rates of m-th AN connected with l-th BN is derived as

$$R_m^{\mathrm{dl,mrc}} = M_{\mathrm{r}}^{\mathrm{A}}\log_2\left(1+\frac{M_{\mathrm{t}}^{\mathrm{B}}\left(M_{\mathrm{t}}^{\mathrm{B}}+M_{\mathrm{t}}^{\mathrm{B}}-M_{\mathrm{r}}^{\mathrm{A}}\right)\beta_{b_l,a_{l,m}}^2}{\mathcal{I}_{b_l,a_{l,m}}+\mathcal{I}_{b_{l'},a_{l,m}}+\mathcal{I}_{u,a}+\mathcal{I}_{a,b}+\sigma_{\boldsymbol{I},a_{l,m}}^2+\sigma_{a_l}^2}\right), \tag{28}$$

where the omitted parts are represented as $\mathcal{I}_{b_l,a_{l,m}} = M_{\mathrm{t}}^{\mathrm{B}}(M-1)(1+\frac{M_{\mathrm{t}}^{\mathrm{A}}-1}{M_{\mathrm{t}}^{\mathrm{B}}-1})\lambda_{b_l,a_{l,m}}$, $\mathcal{I}_{b_{l'},a_{l,m}} = M_{\mathrm{t}}^{\mathrm{B}}M_{\mathrm{r}}^{\mathrm{A}}\beta_{b_l,a_{l,m}}^2 + M_{\mathrm{t}}^{\mathrm{B}}(1+\frac{M_{\mathrm{r}}^{\mathrm{A}}-1}{M_{\mathrm{t}}^{\mathrm{B}}-1})(\lambda_{b_l,a_{l,m}}-\beta_{b_l,a_{l,m}}^2)$, $\mathcal{I}_{u,a} = M_{\mathrm{t}}^{\mathrm{B}}M_{\mathrm{r}}^{\mathrm{A}}\sum_{k=1}^{K_{\mathrm{U}}}\lambda_{u_k,a_{l,m}}$, and the final part is $\mathcal{I}_{a,b} = M_{\mathrm{r}}^{\mathrm{A}}M_{\mathrm{t}}^{\mathrm{B}}(1+\frac{M_{\mathrm{t}}^{\mathrm{A}}-1}{M_{\mathrm{t}}^{\mathrm{B}}-1})(\sum_{i\neq l}^{L}p_{b_i,a_i}M_i\lambda_{b_i,a_{l,m}}+\sum_{i\neq l}^{L}p_{a_i,b_i}M_i\lambda_{b_i,a_{l,m}})$.

The proof of (28) is given in Appendix C.

In summary, Eqs. (24) and (28) represent the average rates of DL users and ANs in DL transmission, respectively. We analyze the spectral efficiency of UL transmissions in the next part.

4.2 Uplink spectral efficiency

For UL transmissions, ANs jointly receive data signals from UL terminals. Simultaneously, ANs transmit UL data to BN it connects with. UL signal models are shown in (6) and (7). The MRC receiver vectors are represented as $\boldsymbol{v}_{u_k,a} = \frac{\hat{\boldsymbol{g}}_{u_k,a}}{\|\hat{\boldsymbol{g}}_{u_k,a}\|}$, the received signal from k-th UL user is given in (6), and the equation can be rewritten as

$$y_{u_k} = \sqrt{p_{u_k,a}}\boldsymbol{v}_{u_k,a}\boldsymbol{g}_{u_k,a}^{\mathrm{H}}x_{u_k,a} + \sum_{j\neq k}^{K_{\mathrm{U}}}\sqrt{p_{u_j,a}}\boldsymbol{v}_{u_k,a}\boldsymbol{g}_{u_j,a}^{\mathrm{H}}x_{u_j,a} + \sum_{j\neq k}\mathcal{I}_{u_j,a}$$

$$+\sum_{j=1}^{K_{\mathrm{D}}}\mathcal{I}_{a,d_j}+\sum_{l=1}^{L}\mathcal{I}_{a_l,b_l}+\sum_{l=1}^{L}\mathcal{I}_{b_l,a_l}+n_{u_k}. \tag{29}$$

The ergodic achievable rates of k-th UL user can be expressed as

$$R_k^{\mathrm{ul,mrc}}=\mathbb{E}\left[\log_2\left(1+\gamma_k^{\mathrm{ul,mrc}}\right)\right], \tag{30}$$

$$\gamma_k^{\mathrm{ul,mrc}}=\frac{\left|\boldsymbol{v}_{u_k,a}\hat{\boldsymbol{g}}_{u_k,a}^{\mathrm{H}}\right|^2}{\mathbb{E}\left[\left(\mathcal{N}_k\right)|\hat{\boldsymbol{g}}_{u,a}\right]}, \tag{31}$$

where $\mathcal{N}_k=\mathcal{N}_{u_j,a}+\mathcal{N}_{u_j,a,\mathrm{est}}+\mathcal{N}_{a,d_j}+\mathcal{N}_{a,b}+\mathcal{N}_{b,a}+\mathcal{N}_{\mathrm{n}}$, and specific expressions of each part are $\mathcal{N}_{u_j,a}=\mathbb{E}[\sum_{j\neq k}|\boldsymbol{v}_{u_k,a}\hat{\boldsymbol{g}}_{u_j,a}^{\mathrm{H}}|^2]$, $\mathcal{N}_{u_j,a,\mathrm{est}}=\mathbb{E}[\sum_{j=1}^{K_{\mathrm{U}}}|\boldsymbol{v}_{u_k,a}\tilde{\boldsymbol{g}}_{u_j,a}^{\mathrm{H}}|^2]$, $\mathcal{N}_{a,d_j}=\mathbb{E}[\sum_{j=1}^{K_{\mathrm{D}}}|\boldsymbol{v}_{u_k,a}\boldsymbol{g}_{a,a}^{\mathrm{H}}\boldsymbol{w}_{a,d_j}|^2]$, $\mathcal{N}_{a,b}=\mathbb{E}[\sum_{l=1}^{L}|\boldsymbol{v}_{u_k,a}\boldsymbol{g}_{a_l,a}^{\mathrm{H}}\boldsymbol{w}_{a_l,b_l}|^2]$, $\mathcal{N}_{b,a}=\mathbb{E}[\sum_{l=1}^{L}|\boldsymbol{v}_{u_k,a}\boldsymbol{g}_{b_l,a}^{\mathrm{H}}\boldsymbol{w}_{b_l,a_l}|^2]$, and $\mathcal{N}_{\mathrm{n}}=\frac{1}{\gamma_{u_k}}\|\boldsymbol{v}_{u_k,a}\|^2$. The definition of UL access link spectral efficiency is the same as DL.

According to [1], the simultaneous occurrence of DL transmission and UL reception leads to the interference between ANs. The reason is that ANs also receive the DL data transmitted when receiving UL data through access link. It is confirmed that the interference is the major factor degrading the UL spectral of FD distributed massive MIMO system. In this paper, we adopt an interference cancelation scheme based on beamforming training to eliminate the interference between ANs. First, the core network estimates the effective interference channel $\boldsymbol{f}_{\mathrm{d}_k}$. Second, the core network regenerates the interference with $\boldsymbol{f}_{\mathrm{d}_k}x_{a,\mathrm{d}_k}$ since the transmitted signals for DL users are known at the core network. Third, the core network subtracts the regenerated interference from the received signal. Finally, the core network decodes the signal transmitted by UL users.

The interference cancelation scheme based on beamforming training is adopted in the analysis of UL achievable rates and DL achievable rates for backhaul links in (28). With MRC receivers, the closed-form expressions of UL ergodic achievable rates of k-th UL user can be expressed as

$$R_k^{\mathrm{ul,mrc}}=\log_2\left(1+\gamma_k^{\mathrm{ul}}\right), \tag{32}$$

$$\gamma_k^{\mathrm{ul}}=\frac{p_{u_k,a}\hat{k}_{u_k,a}\hat{\theta}_{u_k,a}}{\mathcal{I}_{u_j,a}+\mathcal{I}_{u_j,a,\mathrm{est}}+\mathcal{I}_{a,d_j}+\mathcal{I}_{a,b}+\mathcal{I}_{b,a}+\sigma_{u_k}^2}, \tag{33}$$

where the omitted parts are $\mathcal{I}_{u_j,a}=\sum_{j\neq k}\frac{p_{u_j,a}}{MM_{\mathrm{r}}^{\mathrm{A}}}\hat{k}_{u_j,a}\hat{\theta}_{u_j,a}$, $\mathcal{I}_{u_j,a,\mathrm{est}}=\sum_{j=1}^{K_{\mathrm{U}}}\frac{p_{u_j,a}}{MM_{\mathrm{r}}^{\mathrm{A}}}\tilde{k}_{u_j,a}\tilde{\theta}_{u_j,a}$, $\mathcal{I}_{a,d_j}=\sum_{j=1}^{K_{\mathrm{D}}}\frac{p_{a,d_j}}{M_lM_l}\sum_{i=1}^{M_l}\sum_{m=1}^{M_l}\lambda_{i,m}$, $\mathcal{I}_{a,b}=\sum_{l=1}^{L}\frac{p_{a_l,b_l}}{M_l}\sum_{i=1}^{M_l}\sum_{m=1}^{M_l}\lambda_{i,m}$, $\mathcal{I}_{b,a}=\sum_{l=1}^{L}\frac{p_{b_l,a_l}M_{\mathrm{r}}^{\mathrm{B}}}{M_lM_{\mathrm{r}}^{\mathrm{A}}}\sum_{m=1}^{M_l}\lambda_{b_l,a_{l,m}}$.

The proof of (32) is shown in Appendix D.

Remark 3. In FD densely distributed massive MIMO system, interference between ANs caused by the FD technique has a great negative impact on UL spectral efficiency. With the beamforming training scheme, the core network can estimate the effective CSI of the interference channel for interference cancelation, which greatly improves the UL spectral efficiency. To simplify the expressions, the eliminated interference can be seen as a white gaussian noise with i.i.d. $\mathcal{CN}\left(0,\sigma_{\boldsymbol{I}}^2\right)$ elements. $\sigma_{\boldsymbol{I}}^2=\delta_{l,i}^2$ represents the elimination level of the scheme. The closed-form expressions of UL achievable rate without interference cancelation and with perfect interference cancelation can be obtained by substituting $\delta_{l,i}^2$ with $\sum_{j=1}^{K_{\mathrm{D}}}\frac{p_{a,d_j}}{M_lM_l}\sum_{i=1}^{M_l}\sum_{m=1}^{M_l}\lambda_{i,m}$ and with 0 in (33), respectively.

For UL transmission over backhaul network, ANs are regarded as multi-antenna users. We analyze the performance of each antenna. AN performs the MRT beamforming scheme and BN uses MRC receivers, $\boldsymbol{w}_{a_{l,m,n},b_l}=\frac{\hat{\boldsymbol{g}}_{a_{l,m,n},b_l}}{\|\hat{\boldsymbol{g}}_{a_{l,m,n},b_l}\|}$, $\boldsymbol{v}_{a_{l,m,n},b_l}=\frac{\hat{\boldsymbol{g}}_{a_{l,m,n},b_l}}{\|\hat{\boldsymbol{g}}_{a_{l,m,n},b_l}\|}$. $\hat{\boldsymbol{g}}_{a_{l,m,n},b_l}\in\mathbb{C}^{1\times M_{\mathrm{r}}^{\mathrm{B}}}$ is the channel vector between n-th antenna of m-th AN in region l and BN in the same region. With an interference elimination scheme, the interference from data from ANs to DL users is seen as noise with i.i.d $\mathcal{CN}\left(0,\sigma_{\mathrm{I,BN}}^2\right)$ elements. The signal BN received from m-th AN can be written as

$$y_{a_{l,m},b_l}=\mathcal{D}_{a_{l,m},b_l}+\mathcal{I}_{a_{l,m'},b_l}+\sum_{i\neq l}^{L}\mathcal{I}_{a_i,b_i}+\sum_{i=1}^{L}\mathcal{I}_{a_i,b_i}+\sum_{k=1}^{K_{\mathrm{U}}}\mathcal{I}_{u_k}+n_{K_{\mathrm{D}}}+n_{b_l}, \tag{34}$$

where $\mathcal{D}_{a_{l,m},b_l}=\sqrt{p_{a_{l,m},b_l}}\boldsymbol{v}_{a_{l,m},b_l}\boldsymbol{g}_{a_{l,m},b_l}^{\mathrm{H}}\boldsymbol{w}_{a_{l,m},b_l}\boldsymbol{x}_{a_{l,m},b_l}$ is the desired part and the interference caused by other antennas is $\mathcal{I}_{a_{l,m'},b_l}=\sum_{m'\neq m}^{M}\sqrt{p_{a_{l,m'},b_l}}\boldsymbol{v}_{a_{l,m},b_l}\boldsymbol{g}_{a_{l,m'},b_l}^{\mathrm{H}}\boldsymbol{w}_{a_{l,m'},b_l}\boldsymbol{x}_{a_{l,m'},b_l}$. The omitted parts

$\sum_{i\neq l}\mathcal{I}_{a_i,b_i}$, $\sum_{i=1}^{L}\mathcal{I}_{a_i,b_i}$, and $\sum_{k=1}^{K_{\mathrm{U}}}\mathcal{I}_{u_k}$ are the same as (7). The ergodic achievable rates of m-th AN connected with l-th BN can be expressed as

$$R_{l,m}^{\mathrm{BN,mrc}}=\sum_{n=1}^{M_{\mathrm{t}}^{\mathrm{A}}}\mathbb{E}\left[\log_2\left(1+\gamma_{l,m,n}^{\mathrm{BN,mrc}}\right)\right], \tag{35}$$

$$\gamma_{l,m,n}^{\mathrm{BN,mrc}}=\frac{\left|\sqrt{p_{a_{l,m,n},b_l}}\boldsymbol{v}_{a_{l,m,n},b_l}\hat{\boldsymbol{g}}_{a_{l,m,n},b_l}^{\mathrm{H}}\boldsymbol{w}_{a_{l,m,n},b_l}\right|^2}{\mathcal{N}_{l,m,n}}, \tag{36}$$

where $\mathcal{N}_{l,m,n}=\mathcal{N}_{a_{l,m},b_l}+\mathcal{N}_{a_{l,m'},b_l}+\mathcal{N}_{\mathrm{est}}+\mathcal{N}_{u_k}+\mathcal{N}_{a,b}+\mathcal{N}_n$, let $\boldsymbol{V}$ represent the receiver vector $\boldsymbol{v}_{a_{l,m,n},b_l}$ and $\bar{m}'$ represent subscript $a_{l,m',i},b_l$; then specific expressions of each part in $\mathcal{N}_{l,m,n}$ are $\mathcal{N}_{a_{l,m},b_l}=\mathbb{E}\{|\sum_{i=1,i\neq n}^{M_{\mathrm{t}}^{\mathrm{A}}}\boldsymbol{V}\hat{\boldsymbol{g}}_{a_{l,m,i},b_l}^{\mathrm{H}}\boldsymbol{w}_{a_{l,m,i},b_l}|^2\}$, $\mathcal{N}_{a_{l,m'},b_l}=\mathbb{E}\{|\sum_{m'\neq m}^{M}\sum_{i=1}^{M_{\mathrm{t}}^{\mathrm{A}}}\boldsymbol{V}\hat{\boldsymbol{g}}_{\bar{m}',b_l}^{\mathrm{H}}\boldsymbol{w}_{\bar{m}',b_l}|^2\}$, $\mathcal{N}_{\mathrm{est}}=\mathbb{E}\{|\sum_{i\neq l}\sqrt{p_{a_i,b_i}}\boldsymbol{V}\boldsymbol{g}_{a_i,b_l}^{\mathrm{H}}\boldsymbol{w}_{a_i,b_i}|^2\}$, $\mathcal{N}_{u_k}=\mathbb{E}\{|\sum_{k=1}^{K_{\mathrm{U}}}\boldsymbol{V}\boldsymbol{g}_{u_k,b_l}^{\mathrm{H}}|^2\}$, $\mathcal{N}_{a,b}=\mathbb{E}[\sum_{l=1}^{L}|\boldsymbol{V}\boldsymbol{g}_{a_l,a}^{\mathrm{H}}\boldsymbol{w}_{a_l,b_l}|^2]$, and $\mathcal{N}_n=\mathbb{E}\{|\boldsymbol{V}(\boldsymbol{n}_{b_l}+\boldsymbol{n}_{K_{\mathrm{D}}})|^2\}$. Considering that the antenna number of BN is far larger than the number of AN and its antenna, the closed-form expression of the UL ergodic achievable rates of m-th AN connected with l-th BN is derived as

$$R_{l,m}^{\mathrm{BN,mrc}}=\sum_{n=1}^{M_{\mathrm{t}}^{\mathrm{A}}}\log_2\left(1+\gamma_{l,m,n}^{\mathrm{BN,mrc}}\right), \tag{37}$$

$$\gamma_{l,m,n}^{\mathrm{BN,mrc}}=\frac{M_{\mathrm{r}}^{\mathrm{B}}\left(M_{\mathrm{r}}^{\mathrm{B}}+1\right)\beta_{a_{l,m,n},b_l}^2}{\mathcal{I}_{a_{l,m},b_l}+\mathcal{I}_{a_{l,m'},b_l}+\mathcal{I}_{u_k}+\mathcal{I}_{b_l,a_{l,m}}+\mathcal{I}_n}, \tag{38}$$

where $\mathcal{I}_{a_{l,m},b_l}=(M_{\mathrm{t}}^{\mathrm{A}}-1)M_{\mathrm{r}}^{\mathrm{B}}p_{a_{l,m,i},b_l}\beta_{a_{l,m},b_l}^2$, $\mathcal{I}_{a_{l,m'},b_l}=M_{\mathrm{t}}^{\mathrm{A}}M_{\mathrm{r}}^{\mathrm{B}}\sum_{m'\neq m}^{M}p_{a_{l,m',i},b_l}\beta_{a_{l.m'},b_l}^2$, $\mathcal{I}_{u_k}=M_{\mathrm{r}}^{\mathrm{B}}\sum_{i\neq l}p_{a_i,b_i}M_iM_{\mathrm{t}}^{\mathrm{A}}\lambda_{a_i,b_l}$, $\mathcal{I}_{b_l,a_{l,m}}=M_{\mathrm{r}}^{\mathrm{B}}\sum_{k=1}^{K_{\mathrm{U}}}p_{u_k,a}\lambda_{u_k,b_l}$, and $\mathcal{I}_n=M_{\mathrm{r}}^{\mathrm{B}}\left(\sigma_{\mathrm{I,BN}}^2+\sigma_{b_l}^2\right)$.

The proof of (37) is given in Appendix E.

Remark 4. Noted that the closed-form expressions derived is under the case that BN configures with a large number of antennas, if the number of BN antenna is not much larger than the product of the number of AN antennas and the number of AN, the derived results may have some deviation.

The UL and DL spectral efficiency of the backhaul link is defined as

$$\mathrm{SE}_{l,m}=\left(1-\frac{\tau_{\mathrm{p}}+\tau_1}{\tau_{\mathrm{c}}}\right)R_{l,m}^{\mathrm{AN/BN,mrc}}. \tag{39}$$

5 Simulation results

In this section, the accuracy of the closed-form expressions for UL and DL achievable rates is verified by Monte Carlo simulation, the feasibility of two-phase channel estimation is confirmed, and the benefit of the interference cancelation scheme is shown. Moreover, the comparison of FD and HD systems on system spectral efficiency is given.

This paper assumes a circular area with a radius $R=1000$ m. To simplify the analysis, we set one BN located in the center of the area and M_l ANs uniformly located in the area. BN equips with $M_{\mathrm{t}}^{\mathrm{B}}$ transmitting antennas and $M_{\mathrm{r}}^{\mathrm{B}}$ receiving antennas. AN equips with $M_{\mathrm{t}}^{\mathrm{A}}$ transmitting antennas and $M_{\mathrm{r}}^{\mathrm{A}}$ receiving antennas. Let $M_{\mathrm{r}}^{\mathrm{B}}=M_{\mathrm{t}}^{\mathrm{B}}$, $M_{\mathrm{r}}^{\mathrm{A}}=M_{\mathrm{t}}^{\mathrm{A}}$. We assume that there are 10 UL users and 10 DL users in the research area and they are uniformly distributed in the area with a minimum access distance $r_0=30$ m to ANs. According to [17], the path loss exponent of both access link and backhaul link is set as $\alpha_{\mathrm{ul}}=\alpha_{\mathrm{dl}}=\alpha_{\mathrm{AN}}=\alpha_{\mathrm{BN}}=3.7$, the path loss exponent between ANs is set as $\alpha_{\mathrm{I}}=3$, and the path loss exponent of other link is set as $\alpha=4$. We choose $\sigma^2=-83$ dBm. The coherence time in symbols is $T=196$, the length of uplink pilot sequence is $\tau_1=K$, and $\tau_{\mathrm{p}}=M_lM_{\mathrm{t}}^{\mathrm{A}}$. The length of data in access link and backhaul is $T_{\mathrm{data}}=T-K$ and $T_{\mathrm{data}}=T-M_lM_{\mathrm{t}}^{\mathrm{A}}$, respectively.

First, the closed-form expressions given in (24), (28), (32), and (37) are validated by Monte Carlo simulations. We assume that all users have the same UL data transmitted power as $p_{u_k,a}=100$ mW, and ANs have the same DL data transmitted power as $p_{a,d_k}=500$ mW. The data power of data transmitted by ANs in wireless backhaul link is set as 500 mW and the power of data transmitted by BNs in wireless backhaul link is set as 1 W. We assume that the uplink pilot transmitted power as

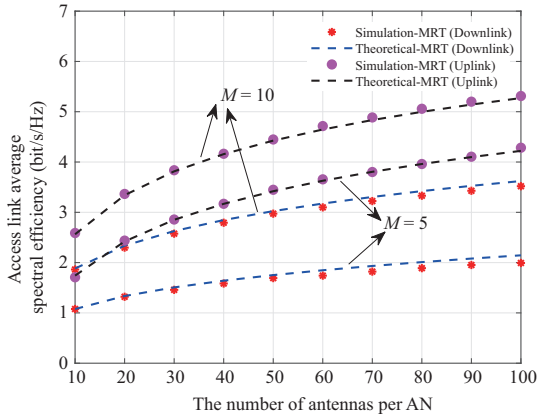

Figure 2 (Color online) Access link average spectral efficiency against the number of antennas per AN.

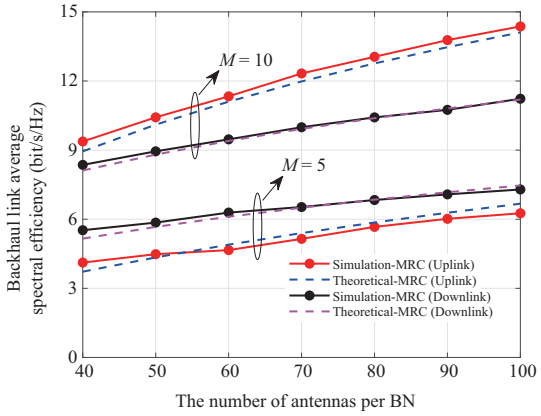

Figure 3 (Color online) Backhaul link average spectral efficiency against the number of antennas per BN.

$p_{\mathrm{up},k} = 500$ mW. Figure 2 illustrates that the simulated results of DL and UL spectral efficiency in access link agree well with the theoretical results, which verifies the accuracy of (24) and (32) when BN's antenna number is set as 50. The spectral efficiency of access link increases when more ANs are distributed in the system as well as the number of ANs' antennae, which verifies the benefits of system scalability. Noted that in a practical situation, the antenna number per AN is not larger than that of BN, the setting in Figure 2 is to illustrate the trend of the spectral efficiency of access link with the increasing of ANs' antenna.

Figure 3 illustrates the simulated results of DL and UL backhaul link spectral efficiency agree well with the theoretical results, which verifies the accuracy of (28) and (37). Different from access link, the DL spectral efficiency of backhaul link is smaller than UL spectral efficiency when the number of AN is 10, while when the number of AN is 5, the UL spectral efficiency is smaller than DL's. The reason is that with the increasing of AN number, the more interference caused by other ANs appears, which decreases the achievable rate of DL transmission compared with UL transmission. Moreover, when AN's number is fixed, with the increasing of antenna number per BN, the more ANs connect with BN, the greater the backhaul spectral efficiency improves.

In wireless backhaul systems, consider the backhaul link from BN to ANs and the access link from ANs to users, the end-to-end achievable rate of the user is given by the lower value among backhaul and access link. Comparing the simulation results in Figures 2 and 3, in both UL and DL transmission, the average spectral efficiency of access link is lower than the average spectral efficiency of backhaul link. Therefore, the spectral performance of FD densely distributed MIMO system with wireless backhaul is mainly limited by access link.

The UL average spectral efficiency in access link with and without interference cancelation scheme against the number of antennas per AN is given in Figure 4. It can be seen that the DL access link interference is the dominant factor to degrade the UL spectral efficiency. However, when the beamforming training scheme is adopted at ANs to estimate the effective interference CSI and perform the interference cancelation scheme, the UL average spectral efficiency increases significantly and it is close to that with perfect interference cancellation, which indicates the necessity and effectiveness of beamforming training based interference cancelation scheme in improving UL spectral efficiency.

The DL average spectral efficiency of access link with perfect and imperfect backhaul link estimation is shown in Figure 5. From the figure, it can be seen that when backhaul link performs perfect channel estimation, the spectral efficiency is higher than when backhaul link performs imperfect channel estimation and the increasing level of spectral efficiency gets larger with the increasing of the number of antennas per AN. In conclusion, the channel estimation for two wireless backhaul is important in DL transmission. The same conclusion is obtained in UL transmission.

The DL average spectral efficiency of access link with different DL transmission power at AN is shown in Figure 6. The antenna number of BN is set as 30. From the comparison of the red and blue line, it can be seen that when the total antenna number (the product of AN number and antenna number per AN) is the same, by adjusting the antenna number per AN and AN number, the DL average spectral efficiency of access link is nearly equal, however, in the case of more ANs, the system performance will be slightly

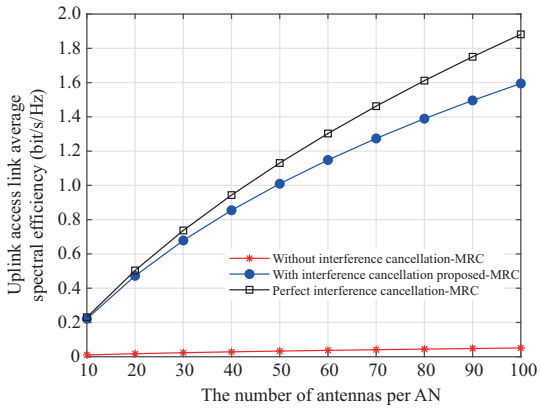

Figure 4 (Color online) Uplink access link average spectral efficiency against the number of antennas per AN.

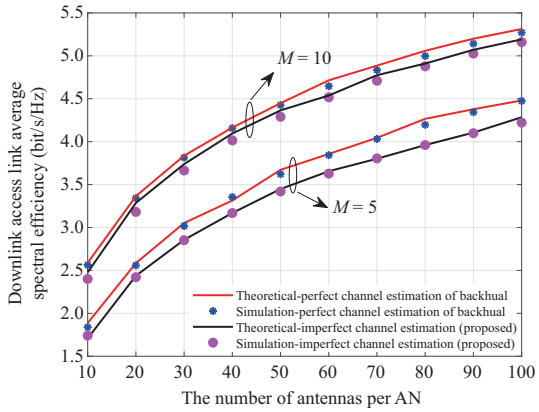

Figure 5 (Color online) Downlink access link average spectral efficiency against the number of antennas per AN.

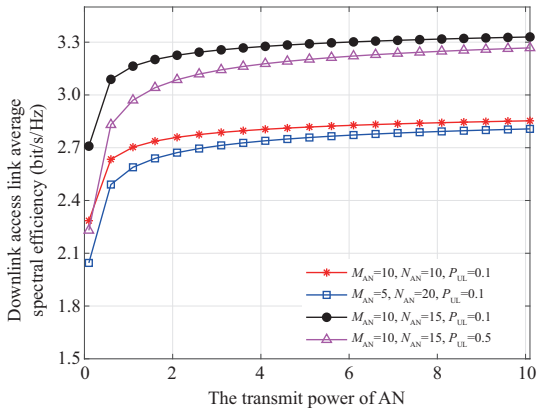

Figure 6 (Color online) Downlink access link average spectral efficiency against the power per AN.

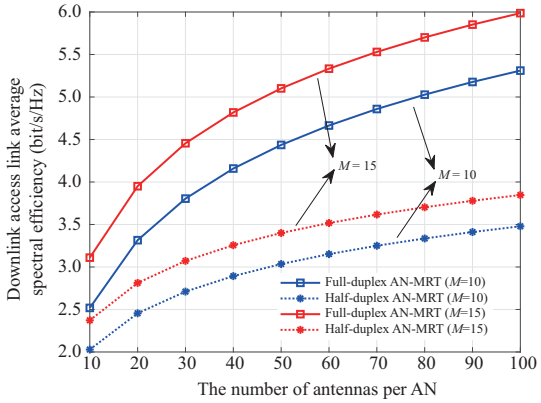

Figure 7 (Color online) Downlink access link average spectral efficiency against the number of antennas per AN.

better. When there is the same AN number, the more antenna per AN, the higher the spectral efficiency, and when the spectral efficiency is saturated, the system performance with the antennas number 15 is about 17.5% higher than that with the antennas number 10. When the AN configuration is the same, the increasing of UL users' power causes a decline in the DL average spectral efficiency of access link. This is because the interference caused by UL user is severer. The conclusion shows that by controlling UL interference from UL users, the DL spectral efficiency improves.

The performance comparison between HD densely distributed MIMO system and FD densely distributed MIMO system with wireless backhaul is presented in Figure 7. ANs in former systems work in HD mode and the latter in FD mode. From Figure 7, it can be seen that compared with the HD system, the FD system can achieve better performance. When the number of AN antennae is 50 and adopts MRT beamforming scheme, the DL spectral efficiency of the FD system can improve by about 48.3% compared with the HD system, and the performance gain of the FD system becomes more significant with the increase of the number of AN antennas. At the same time, the more AN distributed in systems, the advantage of FD system becomes more visible. The same conclusion is obtained for UL access link. The figure illustrates the benefits of an FD system.

6 Conclusion

In this paper, we investigate FD distributed massive MIMO system with wireless backhaul. First, we described the system model, and UL and DL signal transmission models. Under the assumption of imperfect CSI, we proposed a two-phase channel estimation scheme, which aims to get CSI of two wireless links respectively and utilize a beamforming training scheme to perform interference cancelation at ANs.

The spectral efficiency was analyzed to show the advantages of the two-phase channel estimation scheme. With the estimated CSI, we derived the closed-form expressions for DL and UL achievable rates with MRT beamforming and MRC receivers methods, respectively. Based on these expressions, we analyzed the spectral efficiencies of the system and the benefits of the channel estimation scheme proposed. Simulation results verified the accuracy of the theoretical results. It is demonstrated that the imperfect CSI in wireless backhaul link will degrade the achievable rates of access link and the two-phase channel estimation scheme we utilize for channel estimation and interference cancelation can improve the system spectral efficiency greatly. Moreover, it is verified that FD massive MIMO systems with interference cancelation can provide better performance than HD systems.

Acknowledgements This work was supported in part by National Key R&D Program of China (Grant No. 2019YFE0113400) and National Natural Science Foundation of China (Grant Nos. 61971127, 61871465, 61871122).

References

1 Li J M, Lv Q, Zhu P C, et al. Network-assisted full-duplex distributed massive MIMO systems with beamforming training based CSI estimation. IEEE Trans Wireless Commun, 2021, 20: 2190–2204
2 Li J M, Wang D M, Zhu P C, et al. Downlink spectral efficiency of distributed massive MIMO systems with linear beamforming under pilot contamination. IEEE Trans Veh Technol, 2018, 67: 1130–1145
3 You X H, Wang D M, Sheng B, et al. Cooperative distributed antenna systems for mobile communications [coordinated and distributed MIMO. IEEE Wireless Commun, 2010, 17: 35–43
4 Zhu H. Performance comparison between distributed antenna and microcellular systems. IEEE J Sel Areas Commun, 2011, 29: 1151–1163
5 Dai L. A comparative study on uplink sum capacity with co-located and distributed antennas. IEEE J Sel Areas Commun, 2011, 29: 1200–1213
6 Ge X, Tu S, Mao G, et al. 5G ultra-dense cellular networks. IEEE Wireless Commun, 2016, 23: 72–79
7 Siddique U, Tabassum H, Hossain E. Downlink spectrum allocation for in-band and out-band wireless backhauling of full-duplex small cells. IEEE Trans Commun, 2017, 65: 3538–3554
8 Kramer G, Gastpar M, Gupta P. Cooperative strategies and capacity theorems for relay networks. IEEE Trans Inform Theor, 2005, 51: 3037–3063
9 Siddique U, Tabassum H, Hossain E, et al. Wireless backhauling of 5G small cells: challenges and solution approaches. IEEE Wireless Commun, 2015, 22: 22–31
10 Forouzan N, Rabiei A M, Vehkapera M, et al. A distributed resource allocation scheme for self-backhauled full-duplex small cell networks. IEEE Trans Veh Technol, 2021, 70: 1461–1473
11 Sabharwal A, Schniter P, Guo D, et al. In-band full-duplex wireless: challenges and opportunities. IEEE J Sel Areas Commun, 2014, 32: 1637–1652
12 Thilina K M, Tabassum H, Hossain E, et al. Medium access control design for full duplex wireless systems: challenges and approaches. IEEE Commun Mag, 2015, 53: 112–120
13 Duarte M, Sabharwal A, Aggarwal V, et al. Design and characterization of a full-duplex multiantenna system for WiFi networks. IEEE Trans Veh Technol, 2014, 63: 1160–1177
14 Wang D M, Wang M, Zhu P C, et al. Performance of network-assisted full-duplex for cell-free massive MIMO. IEEE Trans Commun, 2020, 68: 1464–1478
15 Tabassum H, Sakr A H, Hossain E. Analysis of massive MIMO-enabled downlink wireless backhauling for full-duplex small cells. IEEE Trans Commun, 2016, 64: 2354–2369
16 Hu X, Zhong C, Chen X, et al. Cell-free massive MIMO systems with low resolution ADCs. IEEE Trans Commun, 2019, 67: 6844–6857
17 Akbar S, Deng Y, Nallanathan A, et al. Massive multiuser MIMO in heterogeneous cellular networks with full duplex small cells. IEEE Trans Commun, 2017, 65: 4704–4719
18 Korpi D, Riihonen T, Sabharwal A, et al. Transmit power optimization and feasibility analysis of self-backhauling full-duplex radio access systems. IEEE Trans Wireless Commun, 2018, 17: 4219–4236
19 Sharma A, Ganti R K, Milleth J K. Joint backhaul-access analysis of full duplex self-backhauling heterogeneous networks. IEEE Trans Wireless Commun, 2017, 16: 1727–1740
20 Li J M, Wang D M, Zhu P C, et al. Uplink spectral efficiency analysis of distributed massive MIMO with channel impairments. IEEE Access, 2017, 5: 5020–5030
21 Anokye P, Ahiadormey R K, Song C, et al. Achievable sum-rate analysis of massive MIMO full-duplex wireless backhaul links in heterogeneous cellular networks. IEEE Access, 2018, 6: 23456–23469
22 Ngo H Q, Larsson E G, Marzetta T L. Massive MU-MIMO downlink TDD systems with linear precoding and downlink pilots. In: Proceedings of the 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton), Monticello, 2013. 293–298
23 Li J M, Wang D M, Zhu P C, et al. Benefits of beamforming training scheme in distributed large-scale MIMO systems. IEEE Access, 2018, 6: 7432–7444
24 Mahyiddin W A, Zakaria N A B, Dimyati K, et al. Downlink rate analysis of training-based massive MIMO systems with wireless backhaul networks. IEEE Access, 2018, 6: 45086–45099
25 Hu N Z, Xiao S H, Pan W S, et al. Performance analysis of the nonlinear self-interference cancellation for full-duplex communications. Sci China Inf Sci, 2022, 65: 212301
26 Tong J P, Zhong C J. Full-duplex two-way AF relaying systems with imperfect interference cancellation in Nakagamim fading channels. Sci China Inf Sci, 2021, 64: 182310

Appendix A Preliminary results

We first provide two lemmas on the properties of Gamma distribution and projection.

Lemma A1. If a random vector $\boldsymbol{x}$ and each of its elements are distributed as $\boldsymbol{x} \sim \mathcal{CN}\left(0, \sigma^2 \boldsymbol{I}_M\right)$, then $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x} \sim \Gamma\left(M, \sigma^2\right)$. If a random vector $\boldsymbol{y}$ is distributed as $\boldsymbol{y} \sim \Gamma\left(a, b\right)$, the expectation and variance of $\boldsymbol{y}$ are $\mathbb{E}\left\{\boldsymbol{y}\right\} = ab$ and $\mathrm{var}\left\{\boldsymbol{y}\right\} = ab^2$[1)].

Lemma A2 ([2])**.** If a set of independent random vectors $\{x_i\}$ and each of them is distributed as $\Gamma\left(k_i, \theta_i\right)$, then the sum of their powers are Gamma distributed; i.e., $\sum_i \boldsymbol{x}_i^{\mathrm{H}}\boldsymbol{x}_i \sim \Gamma\left(k, \theta\right)$, where $k = \frac{(\sum_i k_i\theta_i)^2}{\sum_i k_i\theta_i^2}$, $\theta = \frac{\sum_i k_i\theta_i^2}{\sum_i k_i\theta_i}$.

Lemma A3. If $\boldsymbol{x}$ is an $N \times 1$ vector which elements are i.i.d., $\boldsymbol{A}$ is a constant matrix, then $\mathbb{E}[\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}] = \frac{\mathrm{tr}(\boldsymbol{A})}{N}$[2)].

Lemma A4. Let $\boldsymbol{B} = \boldsymbol{Y}^{\mathrm{H}}\boldsymbol{X}$, where $\boldsymbol{X}$, $\boldsymbol{Y}$ are $M \times N$ random matrix which elements are assumed to be i.i.d. $\mathcal{CN}\left(0, 1\right)$ and $\boldsymbol{C}$ is $N \times N$ matrix. Then $\mathbb{E}\{\boldsymbol{B}^{\mathrm{H}}\boldsymbol{C}\boldsymbol{B}\} = M\mathrm{tr}\left(\boldsymbol{C}\right)\boldsymbol{I}_N$[3)].

Appendix B Proof of equation (24)

We assumed that AN uses an MRT precoding scheme and calculates each part of (24).

First, the numerator $\mathbb{E}[\hat{\boldsymbol{g}}_{a,d_k}^{\mathrm{H}}\boldsymbol{w}_{a,d_k}]$ is calculated. According to Lemma A1, $\|\hat{\boldsymbol{g}}_{a,d_k}\|^2 \sim \Gamma(k_{a,d_k}, \theta_{a,d_k})$. According to the relationship between Gamma distribution and Nakagami distribution, $\|\hat{\boldsymbol{g}}_{a,d_k}\|$ satisfies $\|\hat{\boldsymbol{g}}_{a,d_k}\| \sim \mathrm{Nakagami}(k_{a,d_k}, k_{a,d_k}\theta_{a,d_k})$, based on the property of Nakagami distribution. We obtain $\mathbb{E}[\hat{\boldsymbol{g}}_{a,d_k}^{\mathrm{H}}\boldsymbol{w}_{a,d_k}] = \mathbb{E}[\|\hat{\boldsymbol{g}}_{a,d_k}\|] = \frac{\Gamma(\hat{k}_{a,d_k}+\frac{1}{2})}{\Gamma(\hat{k}_{a,d_k})}\theta_{a,d_k}^{\frac{1}{2}}$.

Next we obtain each part of the denominator. As for $\mathcal{V}$,

$$\begin{aligned}
\mathcal{V} &= \mathbb{E}\left\{\left|\sqrt{p_{a_l,d_k}}\left(\hat{\boldsymbol{g}}_{a_l,d_k}^{\mathrm{H}}\boldsymbol{w}_{a_l,d_k} - \mathbb{E}\left[\hat{\boldsymbol{g}}_{a_l,d_k}^{\mathrm{H}}\boldsymbol{w}_{a_l,d_k}\right]\right)\right|^2\right\} \\
&= p_{a_l,d_k}\mathbb{E}\left\{\left[\left|\boldsymbol{g}_{a,d_k}^{\mathrm{H}}\boldsymbol{w}_{a,d_k}\right|^2\right] - \left(\mathbb{E}\left[\hat{\boldsymbol{g}}_{a_l,d_k}^{\mathrm{H}}\boldsymbol{w}_{a_l,d_k}\right]\right)^2\right\} \\
&= p_{a_l,d_k}\left[\hat{k}_{a,d_k}\hat{\theta}_{a,d_k} + \frac{1}{MM_{\mathrm{t}}^{\mathrm{A}}}\tilde{k}_{a,d_k}\tilde{\theta}_{a,d_k} - \xi\left(\hat{k}_{a,d_k}\right)\hat{\theta}_{a,d_k}\right] \\
&= p_{a_l,d_k}\left[\left(\hat{k}_{a,d_k} - \xi\left(\hat{k}_{a,d_k}\right)\right)\hat{\theta}_{a,d_k} + \frac{1}{MM_{\mathrm{t}}^{\mathrm{A}}}\tilde{k}_{a,d_k}\tilde{\theta}_{a,d_k}\right].
\end{aligned}$$

As for $\mathcal{N}_{a,d_j} = \sum_{j\neq k}\mathbb{E}[|\sqrt{p_{a,d_j}}\boldsymbol{g}_{a_l,d_k}^{\mathrm{H}}\boldsymbol{w}_{a_l,d_j}|^2]$, since $\hat{\boldsymbol{g}}_{a,d_k}$ and $\boldsymbol{w}_{a,d_j}$ are independent, from Lemma A1 in Appendix A, $|\boldsymbol{g}_{a,d_k}^{\mathrm{H}}\boldsymbol{w}_{a,d_j}|^2 \sim \Gamma(\frac{1}{MM_{\mathrm{t}}^{\mathrm{A}}}\hat{k}_{a,d_k}, \hat{\theta}_{a,d_k})$, we obtain

$$\begin{aligned}
\mathcal{N}_{a,d_j} &= p_{a,d_j}\sum_{j\neq k}\left[\left|\boldsymbol{g}_{a_l,d_k}^{\mathrm{H}}\boldsymbol{w}_{a_l,d_j}\right|^2\right] \\
&= p_{a,d_j}\sum_{j\neq k}\left[\left|\left(\hat{\boldsymbol{g}}_{a_l,d_k}^{\mathrm{H}} + \tilde{\boldsymbol{g}}_{a_l,d_k}^{\mathrm{H}}\right)\boldsymbol{w}_{a_l,d_j}\right|^2\right] \\
&= p_{a,d_j}\left(\frac{K_{\mathrm{D}}-1}{MM_{\mathrm{t}}^{\mathrm{A}}}\hat{k}_{a,d_k}\hat{\theta}_{a,d_k} + \frac{K_{\mathrm{D}}-1}{MM_{\mathrm{t}}^{\mathrm{A}}}\tilde{k}_{a,d_k}\tilde{\theta}_{a,d_k}\right).
\end{aligned}$$

Next, the cross-tier interference in different regions $\mathcal{N}_{a,b} = \mathbb{E}[|\sqrt{p_{a_i,b_i}}\boldsymbol{g}_{a_i,d_k}^{\mathrm{H}}\boldsymbol{w}_{a_i,b_i}|^2]$ and $\mathcal{N}_{b,a} = \sum_{i=1}^{L}\mathbb{E}[|\sqrt{p_{b_i,a_i}}\boldsymbol{g}_{b_l,d_k}^{\mathrm{H}}\boldsymbol{w}_{b_i,a_i}|^2]$ is calculated. Use the theorems of the random matrix and get approximate results: $\mathcal{N}_{a,b} = \sum_{i=1}^{L}M_iM_{\mathrm{t}}^{\mathrm{A}}p_{a_i,b_i}\lambda_{a_i,d_k}$ and $\mathcal{N}_{b,a} = M_{\mathrm{t}}^{\mathrm{B}}\sum_{i=1}^{L}p_{b_i,a_i}\lambda_{b_i,d_k}$.

Finally, according to $\left|\boldsymbol{g}_{u_i,d_k}\right|^2 \sim \Gamma\left(1, \lambda_{u_i,d_k}\right)$, we obtain

$$\mathcal{N}_{u,d} = \sum_{i=1}^{K_{\mathrm{U}}}\mathbb{E}\left[\left|\sqrt{p_{u_i,d_k}}\boldsymbol{g}_{u_i,d_k}^{\mathrm{H}}\right|^2\right] = \sum_{i=1}^{K_{\mathrm{U}}}p_{u_i,d_k}\lambda_{u_i,d_k}.$$

In summary, based on the conclusion above, the closed-form expression of downlink access link ergodic achievable rate is shown in (24).

Appendix C Proof of equation (28)

During downlink backhaul transmission, we assume that ANs use MRC receivers and BNs adopt the MRT precoding scheme. We calculate each part of (28).

First, the numerator $\mathbb{E}\{\boldsymbol{v}_{b_l,a_{l,m}}\hat{\boldsymbol{g}}_{b_l,a_{l,m}}^{\mathrm{H}}\boldsymbol{w}_{b_l,a_{l,m}}\}$ is analyzed, according to reference[4)] and Lemma 2.9[5)]. We obtain the numerator equals $M_{\mathrm{t}}^{\mathrm{B}}(M_{\mathrm{t}}^{\mathrm{B}} + M_{\mathrm{t}}^{\mathrm{B}} - M_{\mathrm{r}}^{\mathrm{A}})\beta_{b_l,a_{l,m}}^2\boldsymbol{I}_{M_{\mathrm{r}}^{\mathrm{A}}}$.

Next, the first part of denominator $\mathcal{L}_{b_l,a_{l,m'}} = \mathbb{E}\{|\sum_{m'=1}^{M}\boldsymbol{v}_{b_l,a_{l,m}}\boldsymbol{g}_{b_l,a_{l,m}}^{\mathrm{H}}\boldsymbol{w}_{b_l,a_{l,m'}}|^2\}$ can be decomposed into $\mathbb{E}\{|\sum_{m'\neq m}^{M}\boldsymbol{v}_{b_l,a_{l,m}}\boldsymbol{g}_{b_l,a_{l,m}}^{\mathrm{H}}\boldsymbol{w}_{b_l,a_{l,m'}}|^2\}$ and $\mathbb{E}\{|\boldsymbol{v}_{b_l,a_{l,m}}\boldsymbol{g}_{b_l,a_{l,m}}^{\mathrm{H}}\boldsymbol{w}_{b_l,a_{l,m}}|^2\}$. Calculating each part respectively, according to [26],

1) Heath J R W, Wu T, Kwon Y H, et al. Multiuser MIMO in distributed antenna systems with out-of-cell interference. IEEE Trans Signal Process, 2011, 59: 4885–4899.

2) Lv Q, Li J M, Zhu P C, et al. Spectral efficiency analysis for the bidirectional dynamic network with massive MIMO under imperfect CSI. IEEE Access, 2018, 6: 43600–43671.

3) Mai T C, Ngo H Q, Duong T Q. Cell-free massive MIMO systems with multi-antenna users. In: Proceedings of 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Anaheim, 2018. 828–832.

4) Mai T C, Ngo H Q, Duong T Q. Uplink spectral efficiency of cell-free massive MIMO with multi-antenna users. In: Proceedings of 2019 3rd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom), Ha Noi, 2019. 126–129

5) Tulino A, Verdú S. Random matrix theory and wireless communications. 2004. doi: 10.1561/0100000001.

each antenna is unrelated and we obtain

$$\mathbb{E}\left\{\left|\sum_{m'\neq m}^{M}\boldsymbol{v}_{b_l,a_{l,m}}\boldsymbol{g}^{\mathrm{H}}_{b_l,a_{l,m}}\boldsymbol{w}_{b_l,a_{l,m'}}\right|^2\right\}=M_{\mathrm{t}}^{\mathrm{B}}\left(M-1\right)\left(1+\frac{M_{\mathrm{r}}^{\mathrm{A}}-1}{M_{\mathrm{t}}^{B}-1}\right)\lambda_{b_l,a_{l,m}}\boldsymbol{I}_{M_{\mathrm{r}}^{A}}. \tag{C1}$$

The second part of $\mathcal{L}_{b_l,a_{l,m'}}$ can be composed of $\mathcal{L}_{b_l,a_{l,m'},\mathrm{I}}$ and $\mathcal{L}_{b_l,a_{l,m'},\mathrm{II}}$.

$$\mathcal{L}_{b_l,a_{l,m'},\mathrm{I}}=\mathbb{E}\left\{\frac{\hat{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}}{\left\|\hat{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}\right\|}\hat{\boldsymbol{g}}_{b_l,a_{l,m}}\frac{\hat{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}}{\left\|\hat{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}\right\|}\frac{\hat{\boldsymbol{g}}_{b_l,a_{l,m}}}{\left\|\hat{\boldsymbol{g}}_{b_l,a_{l,m}}\right\|}\hat{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}\frac{\hat{\boldsymbol{g}}_{b_l,a_{l,m}}}{\left\|\hat{\boldsymbol{g}}_{b_l,a_{l,m}}\right\|}\right\},$$

$$\mathcal{L}_{b_l,a_{l,m'},\mathrm{II}}=\mathbb{E}\left\{\frac{\hat{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}}{\left\|\hat{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}\right\|}\tilde{\boldsymbol{g}}_{b_l,a_{l,m}}\frac{\hat{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}}{\left\|\hat{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}\right\|}\frac{\hat{\boldsymbol{g}}_{b_l,a_{l,m}}}{\left\|\hat{\boldsymbol{g}}_{b_l,a_{l,m}}\right\|}\tilde{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}\frac{\hat{\boldsymbol{g}}_{b_l,a_{l,m}}}{\left\|\hat{\boldsymbol{g}}_{b_l,a_{l,m}}\right\|}\right\}.$$

According to Lemma A4 in Appendix A and the relativity of each vector, we have

$$\mathcal{L}_{b_l,a_{l,m'},\mathrm{I}}=M_{\mathrm{t}}^{\mathrm{B}}\left(M_{\mathrm{t}}^{\mathrm{B}}+M_{\mathrm{t}}^{\mathrm{B}}\right)\beta^2_{b_l,a_{l,m}}\boldsymbol{I}_{M_{\mathrm{r}}^{\mathrm{A}}}, \tag{C2}$$

$$\mathcal{L}_{b_l,a_{l,m'},\mathrm{II}}=M_{\mathrm{t}}^{\mathrm{B}}\left(1+\frac{M_{\mathrm{r}}^{\mathrm{A}}-1}{M_{\mathrm{t}}^{\mathrm{B}}-1}\right)\left(\lambda_{b_l,a_{l,m}}-\beta^2_{b_l,a_{l,m}}\right)\boldsymbol{I}_{M_{\mathrm{r}}^{\mathrm{A}}}. \tag{C3}$$

From the results above, $\mathcal{L}_{b_l,a_{l,m'}}$ equals (C1) plus (C2) plus (C3).

As for the calculation of the cross-tier interference $\mathcal{L}_{b_i,a_i}=\mathbb{E}\{|\sum_{i\neq l}\boldsymbol{v}_{b_l,a_{l,m}}\boldsymbol{g}^{\mathrm{H}}_{b_i,a_{l,m}}\boldsymbol{w}_{b_i,a_i}|^2\}$ and $\mathcal{L}_{a_i,b_i}=\mathbb{E}\{|\sum_{i=1}^{L}\boldsymbol{v}_{b_l,a_{l,m}}\boldsymbol{g}^{\mathrm{H}}_{a_i,a_{l,m}}\boldsymbol{w}_{a_i,b_i}|^2\}$, similar to the derived process of (C1), the results of two parts are represented by Ω_1 and Ω_2, which are

$$\Omega_1=M_{\mathrm{r}}^{\mathrm{A}}M_{\mathrm{t}}^{\mathrm{B}}\left(1+\frac{M_{\mathrm{r}}^{\mathrm{A}}-1}{M_{\mathrm{t}}^{\mathrm{B}}-1}\right)\sum_{i\neq l}^{L}p_{b_i,a_i}M_i\lambda_{b_i,a_{l,m}}\boldsymbol{I}_{M_{\mathrm{r}}^{\mathrm{A}}},$$

$$\Omega_2=M_{\mathrm{r}}^{\mathrm{A}}M_{\mathrm{r}}^{\mathrm{A}}\left(1+\frac{M_{\mathrm{r}}^{\mathrm{A}}-1}{M_{\mathrm{r}}^{\mathrm{A}}-1}\right)\sum_{i=1}^{L}p_{a_i,b_i}M_i\lambda_{a_i,a_{l,m}}\boldsymbol{I}_{M_{\mathrm{r}}^{\mathrm{A}}}.$$

Next, we calculate $\mathcal{L}_{u_k,a_l}=\mathbb{E}\{|\sum_{k=1}^{K_{\mathrm{U}}}\boldsymbol{v}_{b_l,a_{l,m}}\boldsymbol{g}^{\mathrm{H}}_{u_k,a_{l,m}}|^2\}$. Since $\boldsymbol{v}_{b_l,a_{l,m}}$ and $\boldsymbol{g}_{u_k,a_{l,m}}$ are unrelated and according to Lemma A3 in Appendix A, we have

$$\begin{aligned}\mathcal{L}_{u_k,a_l}&=\mathbb{E}\left\{\sum_{k=1}^{K_{\mathrm{U}}}\frac{\hat{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}}{\left\|\hat{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}\right\|}\boldsymbol{g}_{u_k,a_{l,m}}\boldsymbol{g}^{\mathrm{H}}_{u_k,a_{l,m}}\frac{\hat{\boldsymbol{g}}_{b_l,a_{l,m}}}{\left\|\hat{\boldsymbol{g}}_{b_l,a_{l,m}}\right\|}\right\}\\&=\mathbb{E}\left\{\sum_{k=1}^{K_{\mathrm{U}}}\frac{\hat{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}}{\left\|\hat{\boldsymbol{g}}^{\mathrm{H}}_{b_l,a_{l,m}}\right\|}\mathbb{E}\left\{\boldsymbol{g}_{u_k,a_{l,m}}\boldsymbol{g}^{\mathrm{H}}_{u_k,a_{l,m}}\right\}\frac{\hat{\boldsymbol{g}}_{b_l,a_{l,m}}}{\left\|\hat{\boldsymbol{g}}_{b_l,a_{l,m}}\right\|}\right\}\\&=M_{\mathrm{t}}^{\mathrm{B}}M_{\mathrm{r}}^{\mathrm{A}}\sum_{k=1}^{K_{\mathrm{U}}}\lambda_{u_k,a_{l,m}}\boldsymbol{I}_{M_{\mathrm{r}}^{\mathrm{A}}}.\end{aligned}$$

Next, the noise part equals $\mathcal{L}_{\boldsymbol{I},a_{l,m}}=\sigma^2_{\boldsymbol{I},a_{l,m}}\boldsymbol{I}_{M_{\mathrm{r}}^{\mathrm{A}}}+\sigma^2_{a_l}\boldsymbol{I}_{M_{\mathrm{r}}^{\mathrm{A}}}$.

In summary, based on the conclusion above, the closed-form expression of downlink backhaul link ergodic achievable rate is shown in (28).

Appendix D Proof of equation (32)

We assumed that ANs use MRC receivers and calculate each part of (32).

First, the numerator $\mathbb{E}[|\boldsymbol{v}_{u_k,a}\hat{\boldsymbol{g}}^{\mathrm{H}}_{u_k,a}|^2]$ is analyzed, according to the estimated channel and its property, $\hat{\boldsymbol{g}}^{\mathrm{H}}_{u_k,a}\hat{\boldsymbol{g}}_{u_k,a}\sim\Gamma(\hat{k}_{u_k,a},\hat{\theta}_{u_k,a})$. Similar to part (1) in Appendix B, we obtain

$$\mathbb{E}\left[\left|\boldsymbol{v}_{u_k,a}\hat{\boldsymbol{g}}^{\mathrm{H}}_{u_k,a}\right|^2\right]=\hat{k}_{u_k,a}\hat{\theta}_{u_k,a}.$$

Next, the first part of denominator is $\mathcal{N}_{u_j,a}=\mathbb{E}[\sum_{j\neq k}|\boldsymbol{v}_{u_k,a}\hat{\boldsymbol{g}}^{\mathrm{H}}_{u_j,a}|^2]$. According to Lemma A3 in Appendix A, we have

$$\mathcal{N}_{u_j,a}=\sum_{j\neq k}\frac{1}{MM_{\mathrm{r}}^{\mathrm{A}}}\mathrm{tr}\left(\mathbb{E}\left[\hat{\boldsymbol{g}}_{u_j,a}\hat{\boldsymbol{g}}^{\mathrm{H}}_{u_j,a}\right]\right)=\sum_{j\neq k}\frac{1}{MM_{\mathrm{r}}^{\mathrm{A}}}\hat{k}_{u_j,a}\hat{\theta}_{u_j,a}.$$

As for $\mathcal{N}_{u_j,a,\mathrm{est}}=\mathbb{E}[\sum_{j=1}^{K_{\mathrm{U}}}|\boldsymbol{v}_{u_k,a}\tilde{\boldsymbol{g}}^{\mathrm{H}}_{u_j,a}|^2]$, similar to the derived process of $\mathcal{N}_{u_j,a}$ in this part, we have

$$\mathcal{N}_{u_j,a,\mathrm{est}}=\sum_{j=1}^{K_{\mathrm{U}}}\frac{1}{MM_{\mathrm{r}}^{\mathrm{A}}}\mathrm{tr}\left(\mathbb{E}\left[\tilde{\boldsymbol{g}}_{u_j,a}\tilde{\boldsymbol{g}}^{\mathrm{H}}_{u_j,a}\right]\right)=\sum_{j=1}^{K_{\mathrm{U}}}\frac{1}{MM_{\mathrm{r}}^{\mathrm{A}}}\tilde{k}_{u_j,a}\tilde{\theta}_{u_j,a}.$$

Then $\mathcal{N}_{a,d_j} = \mathbb{E}[\sum_{j=1}^{K_\mathrm{D}} |\boldsymbol{v}_{u_k,a}\boldsymbol{g}_{a,a}^\mathrm{H}\boldsymbol{w}_{a,d_j}|^2]$ is calculated. Since $\boldsymbol{g}_{a,a}$ and $\boldsymbol{w}_{a,d_j}$ are unrelated, according to Lemma A3 in Appendix A, we obtain

$$\mathcal{N}_{a,d_j} = \sum_{j=1}^{K_\mathrm{D}} \frac{1}{M_l M_l} \sum_{i=1}^{M_l}\sum_{m=1}^{M_l} \lambda_{i,m}.$$

The cross-link interference caused by backhaul links is represented as $\mathcal{N}_{a,b} = \mathbb{E}[\sum_{l=1}^{L} |\boldsymbol{v}_{u_k,a}\boldsymbol{g}_{a_l,a}^\mathrm{H}\boldsymbol{w}_{a_l,b_l}|^2]$ and $\mathcal{N}_{b,a} = \mathbb{E}[\sum_{l=1}^{L} |\boldsymbol{v}_{u_k,a}\boldsymbol{g}_{b_l,a}^\mathrm{H}\boldsymbol{w}_{b_l,a_l}|^2]$. According to Lemma A4 in Appendix A and since $\hat{\boldsymbol{g}}_{u_k,a}$, $\boldsymbol{g}_{a_l,a}$, and $\boldsymbol{w}_{a_l,b_l}$ are unrelated, we have

$$\begin{aligned}\mathcal{N}_{a,b}+\mathcal{N}_{b,a} &= \mathbb{E}\left[\sum_{l=1}^{L}\frac{\hat{\boldsymbol{g}}_{u_k,a}^\mathrm{H}}{\|\hat{\boldsymbol{g}}_{u_k,a}^\mathrm{H}\|}\mathbb{E}\left[\left|\boldsymbol{g}_{a_l,a}^\mathrm{H}\boldsymbol{w}_{a_l,b_l}\right|^2\right]\frac{\hat{\boldsymbol{g}}_{u_k,a}}{\|\hat{\boldsymbol{g}}_{u_k,a}\|}\right] + \mathbb{E}\left[\sum_{l=1}^{L}\frac{\hat{\boldsymbol{g}}_{u_k,a}^\mathrm{H}}{\|\hat{\boldsymbol{g}}_{u_k,a}^\mathrm{H}\|}\mathbb{E}\left[\left|\boldsymbol{g}_{b_l,a}^\mathrm{H}\boldsymbol{w}_{b_l,a_l}\right|^2\right]\frac{\hat{\boldsymbol{g}}_{u_k,a}}{\|\hat{\boldsymbol{g}}_{u_k,a}\|}\right]\\ &= \sum_{l=1}^{L}\frac{1}{M_l}\sum_{i=1}^{M_l}\sum_{m=1}^{M_l}\lambda_{i,m} + \sum_{l=1}^{L}\frac{M_\mathrm{t}^\mathrm{B}}{M_l M_\mathrm{r}^\mathrm{A}}\sum_{m=1}^{M_l}\lambda_{b_l,a_{l,m}}.\end{aligned}$$

Finally, $\mathcal{N}_\mathrm{n} = \frac{1}{\gamma_{u_k}}\|\boldsymbol{v}_{u_k,a}\|^2 = \frac{1}{\gamma_{u_k}}$.

In summary, based on the conclusion above, the closed-form expression of uplink access link ergodic achievable rate is shown in (32).

Appendix E Proof of equation (37)

During uplink backhaul transmission, we assume that BNs use MRC receivers and ANs adopt the MRT precoding scheme. We calculate each part of (37). In the following equations, $\bar{m}'$ represents subscript $a_{l,m',i}, b_l$.

As for the numerator, similar to the process in Appendix D, we have

$$\mathbb{E}\left[\left|\boldsymbol{v}_{a_{l.m,n},b_l}\hat{\boldsymbol{g}}_{a_{l.m,n},b_l}^\mathrm{H}\boldsymbol{w}_{a_{l.m,n},b_l}\right|^2\right] = M_\mathrm{r}^\mathrm{B}\left(M_\mathrm{r}^\mathrm{B}+1\right)\beta_{a_{l,m,n},b_l}^2.$$

As for denominator, the first part is $\mathcal{N}_{a_{l,m},b_l} = \mathbb{E}\{|\sum_{i=1,i\neq n}^{M_\mathrm{t}^\mathrm{A}}\boldsymbol{v}_{a_{l,m,n},b_l}\hat{\boldsymbol{g}}_{a_{l,m,i},b_l}^\mathrm{H}\boldsymbol{w}_{a_{l,m,i},b_l}|^2\}$, when $i\neq n$, $\boldsymbol{v}_{a_{l,m,i},b_l}$ and $\hat{\boldsymbol{g}}_{a_{l,m,n},b_l}$ are unrelated, $\mathcal{N}_{a_{l,m},b_l}$ approximates to $(M_\mathrm{t}^\mathrm{A}-1)M_\mathrm{r}^\mathrm{B}\beta_{a_{l,m},b_l}^2$.

Next, $\mathcal{N}_{a_{l,m'},b_l} = \mathbb{E}\{|\sum_{m'\neq m}^{M}\sum_{i=1}^{M_\mathrm{t}^\mathrm{A}}\boldsymbol{v}_{a_{l,m,n},b_l}\hat{\boldsymbol{g}}_{\bar{m}',b_l}^\mathrm{H}\boldsymbol{w}_{\bar{m}',b_l}|^2\}$ is calculated. When $m'\neq m$, $\boldsymbol{v}_{a_{l,m,n},b_l}$ and $\hat{\boldsymbol{g}}_{\bar{m}',b_l}$ are unrelated, the result equals $M_\mathrm{t}^\mathrm{A}M_\mathrm{r}^\mathrm{B}\sum_{m'=1,m'\neq m}^{M}\beta_{a_{l.m'},b_l}^2$.

Next, $\mathcal{N}_\mathrm{est} = \mathbb{E}\{|\sum_{i\neq l}\sqrt{p_{a_i,b_i}}\boldsymbol{v}_{a_{l,m,n},b_l}\boldsymbol{g}_{a_i,b_l}^\mathrm{H}\boldsymbol{w}_{a_i,b_i}|^2\}$ is calculated. When $i\neq l$, $\boldsymbol{g}_{a_i,b_l}$ and $\boldsymbol{g}_{a_i,b_i}$ are unrelated, according to Lemma A4 in Appendix A, we have

$$\mathcal{N}_\mathrm{est} = \mathbb{E}\left\{\left|\sum_{i\neq l}\sqrt{p_{a_i,b_i}}\boldsymbol{v}_{a_{l,m,n},b_l}\boldsymbol{g}_{a_i,b_l}^\mathrm{H}\boldsymbol{w}_{a_i,b_i}\right|^2\right\} = M_\mathrm{r}^\mathrm{B}\sum_{i\neq l}p_{a_i,b_i}M_i M_\mathrm{t}^\mathrm{A}\lambda_{a_i,b_l}.$$

Next, $\mathcal{N}_{u_k} = \mathbb{E}\{|\sum_{k=1}^{K_\mathrm{U}}\boldsymbol{v}_{a_{l,m,n},b_l}\boldsymbol{g}_{u_k,b_l}^\mathrm{H}|^2\}$ is calculated. According to $|\boldsymbol{g}_{u_i,d_k}|^2 \sim \Gamma(M_\mathrm{r}^\mathrm{B},\beta_{u_k,b_l}^2)$, we have

$$\mathcal{N}_{u_k} = \mathbb{E}\left\{\left|\sum_{k=1}^{K_\mathrm{U}}\boldsymbol{v}_{a_{l,m,n},b_l}\boldsymbol{g}_{u_k,b_l}^\mathrm{H}\right|^2\right\} = M_\mathrm{r}^\mathrm{B}\sum_{k=1}^{K_\mathrm{U}}\beta_{u_k,b_l}^2.$$

Finally, by using an interference cancellation scheme, the noise is $\mathcal{N}_n = \mathbb{E}\{|\boldsymbol{v}_{a_{l,m,n},b_l}(\boldsymbol{n}_{b_l}+\boldsymbol{n}_{K_\mathrm{D}})|^2\}$, according to Lemma A4 in Appendix A as well as the receiver vector and noise which are independent, we have $\mathcal{N}_n = M_\mathrm{r}^\mathrm{B}(\sigma_{\mathrm{I,BN}}^2+\sigma_{b_l}^2)$.

In summary, based on the conclusion above, the closed-form expression of uplink backhaul link ergodic achievable rate is shown in (37).