Downlink ergodic sum capacity maximisation for massive distributed antenna systems with SWIPT protocol

Wenfeng Sun$^{1,2}$ | Chen Liu$^1$ | Mujun Qian$^1$ | Shu Xu$^1$ | Yan Chen$^1$

$^1$ School of Electronic and Optical Engineering, Nanjing University of Posts and Telecommunications, Nanjing, People's Republic of China
$^2$ School of Science and Information, Qingdao Agricultural University, Qingdao, People's Republic of China

Correspondence
Chen Liu, School of Electronic and Optical Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, People's Republic of China.
Email: liuch@njupt.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 61771257

Abstract
The paper investigates simultaneous wireless information and power transfer (SWIPT) in massive distributed antenna systems, where the remote radio heads (RRH) equipped with a large number of antennas are arbitrarily distributed over the coverage area and each terminal operates via a power-splitting (PS) device. Utilising the harvested energy, each terminal transmits its pilot signals for channel estimation. A closed-form lower bound expression on the downlink ergodic capacity is derived for each terminal and then downlink ergodic sum capacity maximisation problem with respect to RRH-terminal association, channel estimation duration, PS ratios and power allocation (PA) is designed. The designed problem is proved to be a mixed-inter nonlinear programming with combinatorial variables, which is generally hard to solve due to its NP-hardness. To this end, a hierarchical iterative algorithm is developed by utilising decomposition technique. Moreover, the convergence and computational complexity of this algorithm are analysed as well. Numerical results manifest the feasibility and effectiveness of this algorithm, and also show that it can achieve the maximum downlink ergodic sum capacity for massive distributed antenna systems with SWIPT protocol.

1 | INTRODUCTION

In recent years, with the rapid growth of Internet of Things (IoT) and Mobile Internet, high energy consumption has become a huge challenge in their further developments. Motivated by this fact, the technologies of green wireless communications have received extensive attention. Among them, the technology of simultaneous wireless information and power transfer (SWIPT) has attracted increasing research interest in both academia and industry [1–4]. The basic idea of SWIPT is that a transmitter can transfer both information and energy simultaneously to the terminals. As energy harvesting operation on wireless signals can destroy the information content carried by the signals, terminal architectures should be redesigned in order to practically achieve information decoding (ID) and energy harvesting (EH). Mainly, two typical terminal architectures were introduced in the existing literatures, namely, time switching (TS) [5] and power splitting (PS) [6]. The terminal used in the TS architecture divides the receiving time into two time slots based on a TS ratio between ID and EH circuits, that is, the terminal performs ID in one time slot and then does EH in the other time slot. The terminal used in the PS architecture divides the received signal power into two power streams with a PS ratio and then both power streams are sent to information decoder and energy harvester, respectively, to perform simultaneously ID and EH [7, 8].

It is very necessary to note that the TS ratio or PS ratio is an important factor as it can directly affect system performance. The optimal TS ratio was investigated to maximise the information rate in single-user case [9] and multiple-user case [10]. The PS ratios were optimally designed to achieve the best ergodic capacity performance while maintaining a required EH rate in a point-to-point system [11]. In [12], the authors jointly optimised transmit beamforming vector, PS ratios and transmit power value to minimise the weighted sum transmit power in full-duplex (FD) point-to-point system. In [13], the authors maximised the weighted sum-rate over all users by varying the time/frequency power allocation (PA) and either TS or PS ratio in multiuser orthogonal frequency division multiplexing (OFDM) systems.
However, SWIPT faces several challenges as well. First, it is very susceptible to shadow fading and path loss, resulting in a lower power transfer efficiency. Second, it gives rise to the issue of how to improve spectral efficiency and energy efficiency simultaneously, as the two goals are in conflict [14]. To overcome these bottlenecks, massive multiple-input multiple output (MIMO) are introduced to combine with SWIPT as it can bring some benefits. For example, it can provide energy beamforming to help restore the energy transfer efficiency by exploiting spatial degree of freedom [15, 16]. Moreover, massive MIMO can obtain high spectral efficiency and energy efficiency when a large number of antennas are equipped at the transmitter [17, 18]. Thus, many research activities have been done on the integration of SWIPT with massive MIMO.

Among them, [19] investigated SWIPT in downlink massive MIMO systems and proposed an iterative algorithm to maximise the minimum achievable rate among all terminals. [20] studied SWIPT in 3D massive MIMO systems and proposed an iterative algorithm to obtain optimal PA and PS ratios for minimising the transmit power with the constraints of signal-to-interference-plus-noise ratio (SINR) and harvested power. [21] considered hybrid TS and PS SWIPT in full-duplex massive MIMO systems and maximised the system achievable sum rate by optimising transmit power and TS ratio. In [22], max-min fairness optimal rate-energy trade-off of SWIPT was derived with imperfectly estimated channels in massive MIMO systems. In [23], the authors investigated massive MIMO enabled SWIPT systems in Rician fading channels and maximised the achievable system sum rate and the minimum rate among terminals, respectively, by optimising channel estimation duration, PS ratios and PA.

Nevertheless, the common assumptions of the above-mentioned studies have been limited to the centralised version of massive MIMO systems, where lots of antennas are co-located in a compact area at the transmitter. A major issue for centralised version of massive MIMO systems is that the power transfer efficiency of each terminal is closely related to its location, i.e., the terminal close to the transmitter can harvest higher power and the terminal far from the transmitter harvests less power. Fortunately, by distributing arbitrarily some remote radio heads (RRH) over the coverage area, massive distributed antenna systems can effectively solve this issue and make SWIPT more feasible [24]. Based on this observation, in this paper, we further take one step from [19] and [23] to investigate SWIPT in massive distributed antenna systems. To the best of our knowledge, such the study has not been found yet in the existing literatures. Specifically, the main contributions of this paper are summarised as follows:

1. We extend SWIPT to massive distributed antenna systems on the basis of [19] and [23]. In the scheme, each terminal utilised its harvested energy to transmit uplink pilot signals for channel estimation. A closed-form lower bound expression on the downlink ergodic capacity for each terminal is derived based on maximum ratio transmission (MRT) precoding. In particular, the RRH-terminal association is considered.

2. We formulate an ergodic sum capacity maximisation problem under the constraints of RRH-terminal association, channel estimation duration, PS ratios and PA. The formulated problem is shown to be a mixed-inter nonlinear programming with combinatorial variables, which is a huge challenge for solving it. To this end, we divide it into four subproblems and propose an iterative optimisation algorithm. Moreover, the convergence and computational complexity of the proposed algorithm are discussed as well.

3. The feasibility and effectiveness of the proposed algorithm are demonstrated by numerical results. These indicate that RRH-terminal association, channel estimation duration, PS ratios and PA are all effective means to improve ergodic sum capacity in massive distributed antenna systems with SWIPT protocol.

The rest of this paper is organised as follows. Section 2 presents the massive distributed antenna system model. The closed-form lower bound expression on ergodic capacity for each terminal is derived and then an ergodic sum capacity maximisation problem is also proposed in Section 3. Section 4 describes an iterative optimisation algorithm to solve the problem. Furthermore, numerical results are conducted to demonstrate this algorithm in Section 5. Finally, we conclude the whole paper in Section 6.

Notations: Scalars are denoted by lowercase or uppercase letters. Vectors and matrices are denoted by bold lowercase and bold uppercase letters, respectively. The meanings of other notations are summarised in Table 1.

### 2 SYSTEM MODEL

A schematic diagram of system model is shown in Figure 1. We consider a massive distributed antenna system with \( L \) RRHs distributed evenly over the coverage area and connected to a baseband processing unit (BPU) using high-quality bidirectional wired or wireless links [24]. Each RRH is equipped with \( M \) antennas and all RRHs serve \( K \) single antenna terminals together in the same time-frequency resource. Each terminal is
equipped with a PS device to coordinate the processes of ID and EH from the received signals. Note that synchronisation overhead among RRHs and backhaul management are heavy burdens for massive distributed antenna system [25]. Thus, in our work, each terminal is allowed to associate with only one RRH.

It is assumed that the channels are constant and frequency-flat in a frame of length \( T \) seconds and the system operates in time-division duplex (TDD) mode. Specially, we only focus on the downlink transmission in this paper and leave uplink transmission in improving system performance, it has to be estimated from uplink pilot signals in every frame. To avoid pilot contamination, the pilot sequences assigned to all terminals must be mutually orthogonal such that

\[
\phi_{k}^{l} = \begin{cases} 1 & k = t \\ 0 & k \neq t \end{cases} \quad \forall k, t.
\]  

In the first phase of every frame, all terminals simultaneously transmit uplink pilot signals for acquisition of CSI. The received pilot signals at RRH \( l \) can be expressed as

\[
Y_{i} = G_{i}P^{1/2}h^{T} + N_{i} \in \mathbb{C}^{M \times T_{p}}.
\]  

Similarly as [26], by employing the minimum mean square error (MMSE) estimation, the estimated channel \( \hat{g}_{lk} \) of the channel \( g_{lk} \) is

\[
\hat{g}_{lk} = \frac{\sqrt{\tau_{p}}h_{lk}}{\tau_{p}q_{k} \sqrt{\beta_{lk}} + \sigma_{ul}^{2}} (\tau_{p}q_{k}g_{lk} + N_{i} \phi_{k})
\]

and the estimation error \( \tilde{g}_{lk} \) of the channel \( g_{lk} \) is

\[
\tilde{g}_{lk} = \hat{g}_{lk} - g_{lk}.
\]

Consequently, \( \hat{g}_{lk} \) and \( \tilde{g}_{lk} \) are independent and distributed as

\[
\hat{g}_{lk} \sim \mathcal{CN}(0, \gamma_{lk}I_{M}),
\]

\[
\tilde{g}_{lk} \sim \mathcal{CN}(0, (\beta_{lk} - \gamma_{lk})I_{M}),
\]

where

\[
\gamma_{lk} = \frac{\tau_{p}q_{k} \beta_{lk}^{2}}{\tau_{p}q_{k} \beta_{lk}^{2} + \sigma_{ul}^{2}}.
\]
With the estimated uplink channels, the downlink CSI can be also obtained easily by exploiting channel reciprocity.

### 2.2 Downlink SWIPT phase

During downlink SWIPT, each RRH transmits information signals to those terminals it serves. Denote \( S_l \) as the set of terminals served by RRH \( l \) and \(|S_l|\) as the number of elements in set \( S_l \). Then, the transmitted signal \( x_l \) at RRH \( l \) is expressed as

\[
x_l = \sum_{i=1}^{\left| S_l \right|} \sqrt{p_{l,i}} w_{l,i} s_{l,i}.
\]

(10)

Here, the data symbol \( s_{l,i} \), which RRH \( l \) intends to transmit to terminal \( t \), has unit power \( E\{\left| s_{l,i} \right|^2\} = 1 \) and \( p_{l,i} \) stands for the transmit power from RRH \( l \) to terminal \( t \) in set \( S_l \). Note that \( p_{l,i} > 0 \) if terminal \( t \) is in set \( S_l \) and \( p_{l,i} = 0 \) otherwise. In addition, \( w_{l,i} \in \mathbb{C}^{M} \) is the corresponding MRT precoding vector as it is the optimal precoding for energy transfer in the context of massive antennas \([27]\). Based on the estimated CSI, \( w_{l,i} \) is described as

\[
w_{l,i} = \frac{\hat{g}_{l,i}}{\sqrt{E\{\|\hat{g}_{l,i}\|^2\}}}.
\]

(11)

After transmission, the received signal at terminal \( k \) in set \( S_k \) is modelled as

\[
y_{l,k} = g_{l,k}^H \sum_{i=1}^{\left| S_l \right|} \sqrt{p_{l,i}} w_{l,i} s_{l,i} + \sqrt{p_{j,k}} \sum_{j \neq i}^{\left| S_j \right|} w_{j,k} s_{j,k} + n_k,
\]

\[
\text{where } n_k \text{ is AWGN and follows } \mathcal{CN}(0, \sigma_k^2). \tag{12}
\]

To achieve SWIPT, each terminal is equipped with a power splitter to coordinate the processes of ID and EH from the received signal. In detail, the \( \rho_{l,k} \in [0, 1] \) portion of the signal power is used for ID and the remaining \( 1 - \rho_{l,k} \) portion is used for EH. Then, the signals split for ID and EH are, respectively, modelled as

\[
J_{l,k}^{\text{ID}} = \sqrt{\rho_{l,k}} \left( g_{l,k}^H \sum_{i=1}^{\left| S_l \right|} \sqrt{p_{l,i}} w_{l,i} s_{l,i} + \right. \\
+ g_{l,k}^H \sum_{j \neq i}^{\left| S_j \right|} \sqrt{p_{j,k}} w_{j,k} s_{j,k} + n_k \right) + n_{l,k}, \tag{13}
\]

\[
J_{l,k}^{\text{EH}} = \sqrt{1 - \rho_{l,k}} \left( g_{l,k}^H \sum_{i=1}^{\left| S_l \right|} \sqrt{p_{l,i}} w_{l,i} s_{l,i} + \\
+ g_{l,k}^H \sum_{j \neq i}^{\left| S_j \right|} \sqrt{p_{j,k}} w_{j,k} s_{j,k} + n_k \right), \tag{14}
\]

where \( n_{l,k} \) is the additional AWGN introduced by baseband conversion \([3]\) and follows \( \mathcal{CN}(0, \sigma_{p}^2) \).

### 3 OPTIMISATION PROBLEM FORMULATION

In this section, we first analyse ergodic capacity and harvested energy for each terminal and then design an ergodic sum capacity maximisation problem.

#### 3.1 Ergodic capacity analysis

In order to obtain an insightful expression that can be used for sum capacity maximisation, a lower bound expression on ergodic capacity for each terminal is derived by utilising the technique of \([28]\), in which the received signal is rewritten as a desired signal, plus an uncorrelated effective noise.

**Theorem 1.** From (11) and (13), by considering Gaussian noise as the worst case distribution of the uncorrelated noise, a lower bound on the downlink ergodic capacity between RRH \( l \) and terminal \( k \) is

\[
R_{l,k} = (1 - \tau) \log_2(1 + \chi_{l,k}), \tag{15}
\]

where \( \chi_{l,k} \) is SINR and given as

\[
\chi_{l,k} = \frac{\rho_{l,k} \gamma_{l,k} y_{l,k} p_{l,k}}{\rho_{l,k} \left( \sum_{j=1}^{L} \sum_{t \neq k} \beta_{j,t} + \sigma_d^2 + \sigma_{p}^2 \right)}, \tag{16}
\]

**Proof.** The proof is provided in Appendix A. \(\square\)

In order to investigate the impact of RRH-terminal association on system performance, one set of integer binary variables is introduced and is modelled as

\[
\omega_{l,k} = \begin{cases} 1 & k \in S_l \\ 0 & k \notin S_l \end{cases} \forall l, k. \tag{17}
\]

Then a lower bound on the ergodic capacity of an arbitrary terminal \( k \) can be expressed as

\[
R_k = (1 - \tau) \sum_{l=1}^{L} \omega_{l,k} \log_2(1 + \chi_{l,k}). \tag{18}
\]

Furthermore, a lower bound on ergodic sum capacity of all the terminals can be expressed as

\[
R_s = (1 - \tau) \sum_{l=1}^{L} \sum_{k=1}^{K} \omega_{l,k} \log_2(1 + \chi_{l,k}). \tag{19}
\]
3.2 | Harvested energy analysis

In the following, we will analyze the harvested energy for each terminal in detail. According to [19], we know that the average harvested energy for each terminal has an asymptotic tight lower bound with respect to the number of antennas. From (14) and resorting to the results of (A.2), (A.3), (A.4) and (A.5), we can easily obtain the lower bound on average harvested energy between RRH $l$ and terminal $k$, which can be expressed as

$$E_{l,k} = (1 - \tau)T(1 - \rho_{l,k})M\gamma_{l,k}p_{l,k}\eta_{k}, \quad (20)$$

where $\eta_{k} \in [0,1]$ represents the energy conversion efficiency of EH.

Similarly as [19], $E_{l,k}$ is used to transmit the uplink pilots, namely,

$$\tau_{pq} = \frac{E_{l,k}}{T} = \frac{(1 - \tau)(1 - \rho_{l,k})M\gamma_{l,k}p_{l,k}\eta_{k}}{\tau}. \quad (21)$$

Combining (9), (20) and (21), we can obtain

$$E_{l,k} = \frac{(1 - \tau)T(1 - \rho_{l,k})M\beta_{l,k}p_{l,k}\eta_{k} - \tau T\sigma_{al}^2}{\beta_{l,k}}. \quad (22)$$

Obviously, $E_{l,k} > 0$ can be easily satisfied as long as the relevant parameters are carefully designed such as $\tau, \rho_{l,k}, M,$ and $p_{l,k}$.

Next, by substituting (22) into (21), we rewrite $\gamma_{l,k}$ as

$$\gamma_{l,k} = \frac{\sigma_{al}^2}{(1 - \tau)(1 - \rho_{l,k})M\gamma_{l,k}p_{l,k}\eta_{k} - \tau T\sigma_{al}^2}, \quad (23)$$

and then by substituting (23) into (16), we rewrite (16) as

$$\chi_{l,k} = \frac{\rho_{l,k}[(1 - \tau)(1 - \rho_{l,k})M\gamma_{l,k}p_{l,k}\eta_{k} - \tau T\sigma_{al}^2]}{[(1 - \tau)(1 - \rho_{l,k})\gamma_{l,k}\beta_{l,k}]} \left[\rho_{l,k} \left(\sum_{j=1}^{L} \sum_{m=1}^{K} p_{l,m}\beta_{l,k}^j + \sigma_p^2\right) + \sigma_p^2\right]. \quad (24)$$

3.3 | Ergodic sum capacity maximisation problem

Based on the above analyses, the ergodic sum capacity maximisation problem can be formulated as

$$\mathcal{P}1 : \max_{\tau, \{\rho_{l,k}\}} R_s \quad (25a)$$

$$\text{s.t.} \quad 0 < \tau < 1, \quad (25b)$$

$$0 < \rho_{l,k} < 1 \quad \forall l, k, \quad (25c)$$

$$p_{l,k} > 0 \quad \forall l, k, \quad (25d)$$

$$\sum_{k=1}^{K} p_{l,k} \leq P_{\text{max}} \quad \forall l, \quad (25e)$$

$$\omega_{l,k} \in \{0, 1\} \quad \forall l, k, \quad (25f)$$

$$\sum_{j=1}^{L} \omega_{l,k} = 1 \quad \forall k, \quad (25g)$$

where (25b) and (25c) represent the constraints of the channel estimation duration for all terminals and PS ratio for each terminal, respectively. (25d) denotes the allocated power for terminal $k$ from RRH $l$ and (25e) guarantees that the total transmit power of RRH $l$ does not exceed its maximum capacity. (25f) and (25g) ensure that each terminal can only associate with one RRH.

We note that problem $\mathcal{P}1$ is a mixed integer nonlinear programming, which usually is hard to solve due to its NP-hardness. Therefore, we will develop an algorithm in the next section by transforming $\mathcal{P}1$ into a solvable one.

4 | PROPOSED OPTIMISATION ALGORITHM

As described in problem $\mathcal{P}1$, the optimisation variables are mutually coupled in the objective function, which is a huge challenge for solving it. To overcome this problem, we first divide the original problem into four subproblems and then focus on the optimisation of only one variable in each subproblem. At last, an iterative optimisation algorithm is designed for all variables. Specially, to develop the algorithm, the binary integer variables are relaxed to continuous variables. As a result, (25f) and (25g) are rewritten as

$$\sum_{j=1}^{L} \omega_{l,k} = 1, \omega_{l,k} \in [0,1], \forall l, k. \quad (26)$$

After relaxation, the fractional RRH-terminal association variables are regarded as partial association with different RRHs.

4.1 | Optimal channel estimation duration

Although longer duration for channel estimation results in more accurate CSI, which can effectively exploit benefits of massive antennas, this also leads to reduced duration for SWIPT. Hence, the trade-off for durations between uplink pilot transmission and downlink SWIPT should be considered.

**Lemma 1.** By assuming that the sampling period is $T_s$, the optimal channel estimation duration with fixed RRH-terminal association, PS ratios
and PA for maximising ergodic sum capacity can be given as

$$\tau^* = \frac{T_s}{T_p}.$$  \hspace{1cm} (27)

**Proof.** The proof is provided in Appendix B.  \hspace{1cm} \square

### 4.2 Optimal power-splitting ratios

When RRH-terminal association, channel estimation duration and PA are fixed, on one hand, higher PS ratios result in higher ergodic sum capacity, but on the other hand, this also degrades the power for EH and makes CSI more inaccurate, which lowers ergodic sum capacity. Thus, PS ratios need to be selected carefully.

**Lemma 2.** With fixed RRH-terminal association, channel estimation duration and PA, the optimal PA ratio which can maximize the ergodic sum capacity, is given as

$$\rho_{j,k}^* = \begin{cases} 
\frac{\xi_{j,1} - \xi_{j,2}}{2\xi_{j,1}} & \xi_{j,0} = \xi_{j,2}\xi_{j,3}, \\
\frac{2\xi_{j,1}\xi_{j,4} - \sqrt{\Delta}}{2(\xi_{j,1}\xi_{j,4} - \xi_{j,2}\xi_{j,3})} & \xi_{j,1}\xi_{j,4} \neq \xi_{j,2}\xi_{j,3},
\end{cases}$$

where

$$\xi_{j,1} = (1 - \tau)M_{p_{j,k}}\eta_{j,k}\beta_{j,k}^2,$$

$$\xi_{j,3} = \sum_{j=1}^{L} \sum_{l=1}^{K} p_{l,j}\beta_{j,k}^4 + \sigma_{d,l}^2,$$

and

$$\Delta = 4\xi_{j,2}\xi_{j,4}(\xi_{j,1}\xi_{j,4} + \xi_{j,1}\xi_{j,3} - \xi_{j,2}\xi_{j,3}).$$

**Proof.** The proof is provided in Appendix C.  \hspace{1cm} \square

### 4.3 Optimal power allocation

The optimal PA strategy is provided in the following, subject to fixed RRH-terminal association, channel estimation duration and PS ratios. Then, $P1$ is equivalent to $P2$ as follows

$$P2: \hspace{1cm} \max_{p_{j,k}} R_s$$

s.t.  \hspace{1cm} $p_{j,k} > 0$ \hspace{1cm} $\forall j, k$  \hspace{1cm} (29a)

$$\sum_{j=1}^{L} \sum_{k=1}^{K} p_{j,k} \leq P_{j}^{\max} \hspace{1cm} \forall j.$$  \hspace{1cm} (29b)

Since $\log_2(\cdot)$ is a monotonic increasing function, $P2$ can be equivalently rewritten as

$$P2.1: \hspace{1cm} \min_{p_{j,k}} \frac{\prod_{j=1}^{L} \prod_{k=1}^{K} (1 + \chi_{j,k})\rho_{j,k}}$$

s.t.  \hspace{1cm} $p_{j,k} > 0$  \hspace{1cm} $\forall j, k$  \hspace{1cm} (30a)

$$\sum_{j=1}^{L} \sum_{k=1}^{K} p_{j,k} \leq P_{j}^{\max} \hspace{1cm} \forall j.$$  \hspace{1cm} (30b)

$$\chi_{j,k} \leq \frac{\rho_{j,k}(1 - \tau)(1 - \rho_{j,k})\eta_{j,k}\beta_{j,k}^2 - \tau\rho_{j,k}}{(1 - \tau)(1 - \rho_{j,k})\eta_{j,k}\beta_{j,k}^2 + \sigma_{d,l}^2 + \sigma_p^2}.$$  \hspace{1cm} (30c)

Note that we have replaced “=“ in (24) with “≤” in (30d). However, this does not change or relax the original problem $P2$ since the objective function $P2.1$ is decreasing with respect to $\chi_{j,k}$. It can be observed from $P2.1$ that all constraints can be transformed into posynomial functions, hence, if the objective function is a monomial function, $P2.1$ becomes a geometric programming (GP) problem and then can be solved with standard optimisation tools such CVX. According to [29], we can transform the objective function into a monotonic function. The key idea is to use a monomial function $\varphi_{j,k}(\chi_{j,k})\rho_{j,k}$ to approximate $(1 + \chi_{j,k})$ near an arbitrary point $\tilde{\chi}_{j,k} > 0$, where $\tilde{\varphi}_{j,k} = \tilde{\chi}_{j,k}(1 + \tilde{\chi}_{j,k})^{-1}$ and $\varphi_{j,k} = (\tilde{\chi}_{j,k})^{-1} \tilde{\varphi}_{j,k}$. Consequently, the objective function can be approximated as

$$\left[\prod_{j=1}^{L} \prod_{k=1}^{K} (\varphi_{j,k}(\chi_{j,k})\rho_{j,k})^{\tilde{\varphi}_{j,k}}\right]^{-1},$$

which is monotonic function. In this way, $P2.1$ is transformed into a GP problem by approximation. Similar to [29], a successive approximation algorithm for PA is proposed in Algorithm 1.

Note that the parameter $\mu$ is used to control approximation accuracy. The approximation accuracy is higher when $\mu$ is close to 1, but the convergence rate is lower, and vice versa. As discussed in [29], $\mu = 1.1$ is a good trade off in most practical cases.

### 4.4 Optimal RRH-terminal association

In this section, we only focus on RRH-terminal association. Combining (26), $P1$ can be equivalently rewritten as

$$P3: \hspace{1cm} \max_{\tilde{\omega}_{j,k}} (1 - \tau) \sum_{j=1}^{L} \sum_{k=1}^{K} \tilde{\omega}_{j,k}\log_2(1 + \chi_{j,k}),$$

s.t.  \hspace{1cm} $\tilde{\omega}_{j,k} \in [0, 1]$  \hspace{1cm} $\forall j, k$  \hspace{1cm} (32a)

$$\sum_{j=1}^{L} \tilde{\omega}_{j,k} = 1 \hspace{1cm} \forall k.$$  \hspace{1cm} (32b)
**Algorithm 1** Optimal power allocation algorithm for $P2.1$

Input: Given the tolerance $\varepsilon_1 > 0$, the parameter $\mu > 1$, and the maximum number of iterations $N_1$. Set $m = 1$. Select the initial values $\mathcal{X}_{l,k}^m$ for $\mathcal{X}_{l,k}$, $\forall l, k$.

Output: Denote $\omega_{l,k}$ as the optimal solutions.

1: Repeat
2: (1) calculate:
3: $\mathcal{X}_{l,k}^n = \left( \mathcal{X}_{l,k}^m \right)^{-1}$
4: $\varphi_{l,k}^n = \left( 1 + \mathcal{X}_{l,k}^n \right)$
5: (2) solve the GP:
6: $\min_{\{p_{l,k}, \mathcal{X}_{l,k}\}} \left[ \prod_{l=1}^{L} \prod_{k=1}^{K} \right]^{-1}
7: \text{s.t. } (30b), (30c), (30d),
8: $\mu^{-1} \mathcal{X}_{l,k}^n \leq \mathcal{X}_{l,k} \leq \mu \mathcal{X}_{l,k}^n$
9: (3) update the initial values:
10: Set $m = m + 1$ and update $\mathcal{X}_{l,k}^{(m)}$, where $\mathcal{X}_{l,k}$ are obtained based on the solutions of GP.
11: Until: $\max_{l,k} \left\| \frac{\mathcal{X}_{l,k}^{(m)} - \mathcal{X}_{l,k}^{(m-1)}}{\mathcal{X}_{l,k}^{(m-1)}} \right\| < \varepsilon_1$, or $m = N_1$.

We observe from $P3$ that the objective function is linear combination of $\omega_{l,k} \forall l,k$. Thus, $P3$ can be effectively solved by linear program. After the optimal continuous variables $\omega_{l,k} \forall l,k$ are obtained, we further restore them to binary integer variables by

$$\omega_{l,k} = \begin{cases} 1 & \text{if } \omega_{l,k} = \text{max} (\omega_{l,k}) \\
0 & \text{if } \omega_{l,k} \neq \text{max} (\omega_{l,k}) \end{cases} \forall l,k. \quad (33)$$

In this way, each terminal can only associate with one RRH.

**4.5 Iterative optimisation algorithm**

Based on the above analyses for four subproblems, we are now ready to design a feasible iterative framework for original problem $P1$, which includes the following three aspects.

1. The optimal channel estimation duration is first calculated based on Lemma 1. As a result, it is irrelevant to RRH-terminal association, PS ratios and PA. Thus, it does not need to be updated during iterations.
2. Inner loop: calculate the optimal PS ratios based on Lemma 2 and then update PA via Algorithm 1 for fixed RRH-terminal association.
3. Outer loop: update the RRH-terminal association for fixed PS ratios and PA.

The detailed procedures of the whole algorithm are summarised in Algorithm 2.

**Theorem 2.** The convergence of Algorithm 2 can be guaranteed.

---

**Algorithm 2** Ergodic sum capacity maximisation algorithm for $P1$

Input: Given tolerances $\varepsilon_2 > 0$ and $\varepsilon_3 > 0$, parameter $\mu > 1$, and maximum numbers of iterations $N_2$ and $N_3$. Set $\lambda = 1$ and $\vartheta = 1$. Select initial RRH-terminal association is $1/L$ and initial PA is equal power for each terminal.

Output: Denote $\tau, \rho_{l,k}, \alpha_{l,k}$ and $\phi_{l,k}$ as the optimal solutions.

1: calculate $\tau$ based on Lemma 1;
2: while $\vartheta < N_2$;
3: calculate $\rho_{l,k}$ based on Lemma 2;
4: while $\tau < N_3$;
5: update $\phi_{l,k}$ via Algorithm 1;
6: $\lambda = \lambda + 1$;
7: calculate $R_{l,k}^{(\alpha)}$ via (15) and (24);
8: break if $\max_{l,k} \left\| \frac{R_{l,k}^{(\alpha)} - R_{l,k}^{(\beta)}}{R_{l,k}^{(\beta)}} \right\| < \varepsilon_2$;
9: end
10: $\vartheta = \vartheta + 1$;
11: update $\phi_{l,k}$ via linear program based on $P3$;
12: break if $\max_{l,k} \left\| \frac{\phi_{l,k}^{(\beta)} - \phi_{l,k}^{(\alpha)}}{\phi_{l,k}^{(\alpha)}} \right\| < \varepsilon_3$;
13: end
14: restore $\omega_{l,k}$ to $\phi_{l,k}$ via (33);

---

**Proof.** First, since Algorithm 1 is a successive iterative approximation algorithm based on GP, its convergence is guaranteed [29]. Then, for inner loop, during the $\lambda^{th}$ iteration, $\{\rho_{l,k}\}$ is first calculated based on the PA $\{p_{l,k}\}^{(\beta)}$ from previous iteration. According to Lemma 2, we get

$$R_{l}(\{p_{l,k}\}^{(\beta)}, \{\rho_{l,k}\}^{(\beta)}) < R_{l}(\{p_{l,k}\}^{(\beta)}, \{\rho_{l,k}\}^{(\alpha)}). \quad (34)$$

With obtained $\{\rho_{l,k}\}^{(\beta)}$, we move forward to update $\{p_{l,k}\}^{(\beta)}$. According to Algorithm 1, we get

$$R_{l}(\{p_{l,k}\}^{(\beta)}, \{\rho_{l,k}\}^{(\beta)}) < R_{l}(\{p_{l,k}\}^{(\alpha)}, \{\rho_{l,k}\}^{(\beta)}). \quad (35)$$

By combining (34) and (35), we get

$$R_{l}(\{p_{l,k}\}^{(\alpha)}, \{\rho_{l,k}\}^{(\beta)}) < R_{l}(\{p_{l,k}\}^{(\beta)}, \{\rho_{l,k}\}^{(\beta)}). \quad (36)$$

Obviously, there exits an upper bound of ergodic sum capacity. Therefore, inner loop is convergent. Finally, outer loop is a successive iterative algorithm based on linear program and its convergence has been proved in many literatures such as [30]. Based on the above analyses, we conclude that the convergence of Algorithm 2 can be guaranteed.

We then discuss the complexity of Algorithm 2 which is mainly dependent on the complexities of three aspects involved, namely, outer loop, inner loop and Algorithm 1. Denote that $N_{\alpha}$ is the number of required iterations for outer loop, $N_{\beta}$ is the number of required iterations for inner loop and $(1/\varepsilon_1^2)N_{\alpha}LK$ is the complexity of Algorithm 1, where $N_{\alpha}$ is the number of iterations...
required iteration for GP. As a result, the complexity of Algorithm 2 is roughly expressed as \((1/\varepsilon_l) N_l N_p N_g L K\). Moreover, linear program and GP used in Algorithm 2 have high computational efficiency as they can be solved by standard optimisation tools. Thus, Algorithm 2 can converge quickly and has a low complexity, which is also confirmed by numerical results in Section 5.

5 | NUMERICAL RESULTS

In this section, the proposed optimisation algorithm is evaluated by numerical results for massive distributed antenna systems. There are three RRHs and six terminals in circular area with a radius of 10 meters, as shown in Figure 1. In order to interpret clearly numerical results, we assume that the Cartesian coordinates of three RRHs are \((0,10), (-10,-10)\) and \((10,-10)\), respectively, and six terminals are uniformly and randomly distributed over circular area. The large-scale fading is modelled as \(\beta_{l,k} = 10^{-3} d_{l,k}^{-3}\), where \(d_{l,k}\) represents the distance between terminal \(k\) and RRH \(l\). The other system parameters are listed in Table 2, unless otherwise stated.

As all terminals share the same channel estimation duration, maximising ergodic capacity of each terminal is equivalent to maximising ergodic sum capacity of the whole network. Figure 3 shows the impact of channel estimation duration on ergodic capacity of each terminal. Here, we only focus on the impact of \(\tau\) and thus the other parameters are arbitrary given within the range of feasible values. We assume that \(\tilde{\omega}_{l,k} = 1/3\), \(p_{l,k} = 0.5\) and \(\rho_{l,k}\) is obtained by (28), for \(\forall l,k\). It is observed from Figure 3 that the ergodic capacity of each terminal is monotone decreasing with respect to \(\tau\). Therefore, \(\tau\) should take the minimum value on the premise of no pilot contamination. According to the above assumed parameters, the optimal value \(\tau^+\) is 0.03, which is consistent with the result calculated from (27) and numerically verifies Lemma 1.

As the PS ratio of each terminal is independent of each other, the optimal PS ratio for maximising ergodic capacity of each terminal is also for maximising ergodic sum capacity of the whole network. Figure 4 investigates the impact of PS ratio on the ergodic capacity for each terminal. Here, we only focus on \(\rho_{l,k}\) and \(\tilde{\omega}_{l,k} = 1/3\), \(\tau^* = 0.03\) and \(p_{l,k} = 0.5\) are given, for \(\forall l,k\). It is observed from Figure 4 that the ergodic capacity of each terminal is a quasi-concave function with respect to its own PS ratio. Thus, there exists an optimal value to maximum ergodic capacity for each terminal. In addition, the optimal PS ratio is 2.27, 3.17, 2.82, 3.30, 3.02 and 2.73, respectively, in this simulation, which are identical to the results calculated from (28) and numerically verifies Lemma 2.

We show the impact of PA on ergodic sum capacity in Figure 5 and the other parameters are given, namely, \(\tilde{\omega}_{l,k} = 1/3\), \(\tau^* = 0.03\) and \(\rho_{l,k}\) is obtained by (28), for \(\forall l,k\). Since Algorithm 1 can obtain optimal PA for each terminal, we denote it by “OPA” in Figure 5. For comparison, we consider equal PA (EPA) scheme in which each RRH equally allocates its power to each terminal it serves without beamforming, namely, omnidirectional powering. Such EPA scheme serves widely as a benchmark for comparison in existing literatures such as [19] and [27].

| Parameters | Descriptions | Settings |
|------------|--------------|----------|
| \(p_{l,k}^{max} \) | Maximum power | 3 Watt |
| \(T_s\) | Frame length | 1 s |
| \(T_s\) | Sampling period | 0.005 s |
| \(\sigma_{dl}^2\) | Noise variance for uplink | \(-90\ dBm\) |
| \(\sigma_{dl}^2\) | Noise variance for downlink | \(-90\ dBm\) |
| \(\sigma_{dl}^2\) | Noise variance for baseband conversion | \(-50\ dBm\) |
| \(M_{l,k}\) | Number of antennas per RRH | 200 |
| \(\eta_{l,k}\) | Conversion efficiency | 0.8 |
| \(\varepsilon_{1,2,3}\) | Given tolerances | \(10^{-4}\) |

FIGURE 3 Ergodic capacity versus channel estimation duration

FIGURE 4 Ergodic capacity versus power-splitting ratio
We can observe from Figure 5 that ergodic sum capacity based on OPA outperforms that based on EPA with the same given transmit power from RRH and it verifies that Algorithm 1 can provide the optimal PA for maximising ergodic sum capacity. We also find that ergodic sum capacity benefits from increased transmit power for both OPA and EPA, however, the increased amplitude of ergodic sum capacity becomes smaller as higher transmit power leads to stronger interference among terminals. In Figure 6, we investigate the impact of RRH-terminal association on ergodic sum capacity and the other parameters are given, namely, $\tau^* = 0.03$, $\rho_{l,k} = 0.5$ and $P_{l,k}$ is obtained by (28), for $\forall l, k$. Since linear program can obtain the optimal RRH-terminal association, we denote it by “LPRTA” in Figure 6. For comparison, we consider minimum distance RRH-terminal association (MDRTA) scheme in which each terminal associates with its nearest RRH. Such MDRTA scheme is also used for comparison in [31]. In order to interpret the results easily, we randomly generate two sets of Cartesian coordinates representing location of 6 terminals, that is, $\text{Set I} = \{(0.8462, -3.6662), (4.4045, -5.5988), (3.620, -5.6968), (-2.3673, 2.2611), (5.0411, -7.5397), (7.7289, 4.9628)\}$ and $\text{Set II} = \{(-0.2378, -1.1545), (-0.9579, -3.4857), (0.1532, 0.1939), (5.7173, 5.3070), (2.5977, -21850), (5.6084, 0.5909)\}$. We can observe from Figure 6 that LPRTA achieves a higher ergodic sum capacity than MDRTA for both Set I and Set II with the same number of antennas, which verifies that linear program can provide optimal RRH-terminal association for maximising ergodic sum capacity. In addition, from Figures 5 and 6, we can find that a larger number of antennas always can improve system performance.

Finally, the convergence of Algorithm 2 is demonstrated in Figure 7. According to the analyses from subsection 4.5, we know that the convergence of Algorithm 2 mainly depends on three aspects: outer loop, inner loop and Algorithm 1. Thus, we show the convergence of each of them in Figure 7. In this experiment, the outer loop requires 7 iterations to converge. In the meantime, the inner loop performs 37 iterations and Algorithm 1 performs 217 iterations. They numerically verify the correctness of Theorem 2.

6 | CONCLUSIONS

This paper investigated the ergodic sum capacity for massive distributed antenna systems with SWIPT protocol. With the objective to maximise ergodic sum capacity of the whole system, an optimisation problem was formulated by jointly considering RRH-terminal association, channel estimation duration, PA and PS ratios. Due to the nonconvex and nonlinear properties of the formulated problem, we divided it into four subproblems and then was solved one by one. At last, an iterative optimisation algorithm was proposed for original problem and its convergence and complexity were also discussed in detail. Moreover, numerical results manifest the correctness and feasibility of the
proposed algorithm, and also show that it can provide the optimal ergodic sum capacity for massive distributed antenna systems with SWIPT protocol.

ACKNOWLEDGEMENTS

This work is supported by the Natural Science Foundation of China under Grant 61771257, the open research fund of National and Local Joint Engineering Laboratory of RF Integration and Micro-Assembly Technology in Nanjing University of Posts and Telecommunications (NUPT) under Grant KFJJ20170202, and the Start-up Fund of NUPT under Grant NY217127.

REFERENCES

1. Wang, D., et al.: Relay selection in full-duplex energy-harvesting two-way relay networks. IEEE Trans. Green Commun. Netw. 1(2), 182–191 (2017)
2. Wang, D., et al.: Capacity-enhancing full-duplex relay networks based on power-splitting (ps)-swipt. IEEE Trans. Veh. Technol. 66(6), 5445–5450 (2017)
3. Alsaba, Y., et al.: Beamforming in wireless energy harvesting communications systems: A survey. IEEE Commun. Sur. Tutor. 20(2), 1329–1360 (2018)
4. Ponnambaduge Perera, T.D., et al.: Simultaneous wireless information and power transfer (swipt): Recent advances and future challenges. IEEE Commun. Sur. Tutor. 20(1), 264–302 (2018)
5. Zhang, R., Ho, C.K.: MIMO broadcasting for simultaneous wireless information and power transfer. IEEE Trans. Wirel. Commun. 12(5), 1989–2001 (2013)
6. Zhou, X., et al.: Wireless information and power transfer: Architecture design and rate-energy tradeoff. IEEE Trans. Commun. 61(11), 4754–4767 (2013)
7. Ng, D.W.K., et al.: Robust beamforming for secure communication in systems with wireless information and power transfer. IEEE Trans. Wirel. Commun. 13(8), 4599–4615 (2014)
8. Ng, D.W.K., et al.: Wireless information and power transfer: Energy efficiency optimization in OFDMA systems. IEEE Trans. Wirel. Commun. 12(12), 6352–6370 (2013)
9. Chen, X., et al.: Wireless energy and information transfer tradeoff for limited-feedback multiantenna systems with energy beamforming. IEEE Trans. Veh. Technol. 63(1), 407–412 (2014)
10. Liu, L., et al.: Multi-antenna wireless powered communication with energy beamforming. IEEE Trans. Veh. Technol. 62(12), 4349–4361 (2014)
11. Zhou, X.: Training-based swipt: Optimal power splitting at the receiver. IEEE Trans. Veh. Technol. 64(9), 4377–4382 (2015)
12. Hu, Z., et al.: Weighted sum transmit power minimization for full-duplex system with swip and self-energy recycling. IEEE Access 4, 4874–4881 (2016)
13. Zhou, X., et al.: Wireless information and power transfer in multiuser OFDM systems. IEEE Trans. Wirel. Commun. 13(4), 2282–2294 (2014)
14. Chen, X., et al.: Enhancing wireless information and power transfer by exploiting multi-antenna techniques. IEEE Commun. Mag. 53(4), 133–141 (2015)
15. Rusek, F., et al.: Scaling up MIMO: Opportunities and challenges with very large arrays. IEEE Signal Process. Mag. 30(1), 40–60 (2013)
16. Kanag, G.N., et al.: Spectral-efficiency analysis of massive MIMO systems in centralized and distributed schemes. IEEE Trans. Commun. 64(5), 1930–1941 (2016)
17. Larsson, E.G., et al.: Massive MIMO for next generation wireless systems. IEEE Commun. Mag. 52(2), 186–195 (2014)
18. Ngo, H.Q., et al.: Energy and spectral efficiency of very large multiuser MIMO systems. IEEE Trans. Commun. 61(4), 1436–1449 (2013)
19. Wang, X., Zhai, C.: Simultaneous wireless information and power transfer for downlink multi-user massive antenna-array systems. IEEE Trans. Commun. 65(9), 4039–4048 (2017)
20. Fan, L., et al.: Exploiting BS antenna tilt for swipt in 3D massive mimo systems. IEEE Wirel. Commun. Lett. 6(5), 666–669 (2017)
21. Xu, K., et al.: Hybrid time-switching and power splitting swipt for full-duplex massive MIMO systems: A beam-domain approach. IEEE Trans. Veh. Technol. 67(8), 7257–7274 (2018)
22. Kudathanthirige, D., et al.: Max-min fairness optimal rate-energy tradeoff of SWIPT for massive MIMO downlink. IEEE Commun. Lett. 23(4), 688–691 (2019)
23. Dong, G., et al.: Achievable rate optimization for massive MIMO enabled SWIPT systems over downlink Rician channels. IEEE Access 6, 36810–36824 (2018)
24. Yuan, F., et al.: Joint wireless information and energy transfer in massive distributed antenna systems. IEEE Commun. Mag. 53(6), 109–116 (2015)
25. Rogalin, R., et al.: Scalable synchronization and reciprocity calibration for distributed multiuser MIMO. IEEE Trans. Wirel. Commun. 13(4), 1815–1831 (2014)
26. Van Chen, T., et al.: Joint power allocation and user association optimization for massive MIMO systems. IEEE Trans. Wirel. Commun. 15(9), 6384–6399 (2016)
27. Yang, G., et al.: Throughput optimization for massive MIMO systems powered by wireless energy transfer. IEEE J. Sel. Areas Commun. 33(8), 1640–1650 (2015)
28. Jose, J., et al.: Pilot contamination and precoding in multi-cell TDD systems. IEEE Trans. Wirel. Commun. 10(8), 2640–2651 (2011)
29. Weeraddana, P.C., et al.: Resource allocation for cross-layer utility maximization in wireless networks. IEEE Trans. Veh. Technol. 60(6), 2790–2809 (2011)
30. Zhou, T., et al.: Energy-efficient user association in downlink heterogeneous cellular networks. IET Commun. 10(13), 1553–1561 (2016)
31. Dong, G., et al.: Energy-efficiency-oriented joint user association and power allocation in distributed massive MIMO systems. IEEE Trans. Veh. Technology 68(6), 5794–5808 (2019)

APPENDIX A: PROOF OF THEOREM 1

By some operations, (13) can be rewritten as:

$$J_{l,k}^{ID} = \sqrt{P_{l,k}} \left( E \left\{ \sum_{t} \left| \hat{g}_{l,k}^{H} \sqrt{P_{l,k}} w_{l,k} \right| s_{l,k} \right\} + \left( \sum_{t} \left| \hat{g}_{l,k}^{H} \sqrt{P_{l,k}} w_{l,k} \right| s_{l,k} \right) \right)$$

$$+ \left( \sum_{t} \left| \hat{g}_{l,k}^{H} \sqrt{P_{l,k}} w_{l,k} \right| s_{l,k} \right)$$

$$+ \left( \sum_{t} \left| \hat{g}_{l,k}^{H} \sqrt{P_{l,k}} w_{l,k} \right| s_{l,k} \right)$$

As a result, the first term is considered as desired signal and the other terms are treated as uncorrelated Gaussian noise.

How to cite this article: Sun W, Liu C, Qian M, Xu S, Chen Y. Downlink ergodic sum capacity maximisation for massive distributed antenna systems with SWIPT protocol. IET Commun. 2021;15:464–475. https://doi.org/10.1049/cmu2.12079
The desired signal power is computed as

\[ \sum \mathbb{E} \left( \mathbf{g}_{j,t}^H \sqrt{p_j} \mathbf{w}_{j,t} \right)^2 = \mathbb{E} \left( \sqrt{p_j} \left| \mathbf{g}_{j,t}^H \right|^2 \right) \]

\[ = \mathbb{E} \left( \left| \mathbf{g}_{j,t}^H \right|^2 \right) \]

The second term and the third term represent beamforming gain uncertainty, and its power is computed as

\[ \mathbb{E} \left( \left| \mathbf{g}_{j,t}^H \sqrt{p_j} \mathbf{w}_{j,t} \right|^2 \right) = \mathbb{E} \left( \left| \mathbf{g}_{j,t}^H \right|^2 \right) \]

\[ = \mathbb{E} \left( \left| \mathbf{g}_{j,t}^H \sqrt{p_j} \mathbf{w}_{j,t} \right|^2 \right) \]

The fourth term and the fifth term represent the interference introduced by those terminals served by RRHs, and its power is computed as

\[ \sum_{j \neq k} \gamma_{j,k} p_{j,t} = \sum_{j \neq k} \gamma_{j,k} p_{j,t} \]

\[ = \sum_{j \neq k} \gamma_{j,k} p_{j,t} \]

The sixth term represents the channel estimation error, and its power is computed as

\[ \sum_{j \neq k} \gamma_{j,k} p_{j,t} \]

\[ = \sum_{j \neq k} \gamma_{j,k} p_{j,t} \]

With the preliminaries above, the SINR ratio is computed as

\[ \chi_{j,k} = \frac{\rho_{j,k} M \gamma_{j,k} p_{j,k}}{\rho_{j,k} \left( \sum_{j \neq k} \gamma_{j,k} p_{j,t} + \sigma_d^2 \right) + \sigma_p^2} \]

\[ \text{where} \]

\[ \xi_{\tau,1} = \rho_{j,k} \left( 1 - \rho_{j,k} \right) M \gamma_{j,k} \beta_{j,k}^2 \]

\[ \xi_{\tau,2} = \left[ \left( 1 - \rho_{j,k} \right) \eta_{j,k} \beta_{j,k} + \sigma_d^2 \right] + \sigma_p^2 \]

and

\[ \xi_{\tau,3} = \rho_{j,k} \sigma_d^2 \]

are all positive parameters irrelevant of \( \tau \). To obtain the optimal \( \tau \), we calculate the derivative of \( R_\tau \) as

\[ \frac{dR_\tau}{d\tau} = \frac{-L \sum_{k=1}^{K} \gamma_{j,k} p_{j,t}}{\left( M + 1 \right) \gamma_{j,k} p_{j,k} + \lambda_{j,k} p_{j,k}} \]

\[ \quad - \frac{L \sum_{k=1}^{K} \gamma_{j,k} p_{j,t}}{\left( M + 1 \right) \gamma_{j,k} p_{j,k} + \lambda_{j,k} p_{j,k}} \]

\[ \quad \times \frac{\xi_{\tau,1} - \xi_{\tau,2}}{\xi_{\tau,2}} \]

It is obvious that \( \frac{dR_\tau}{d\tau} < 0 \) due to \( \tau \in (0, 1) \). Consequently, \( R_\tau \) is monotonically decreasing with respect to \( \tau \) and so the minimum \( \tau \) can maximise \( R_\tau \). In addition, \( \tau_p \geq K \) should be satisfied to avoid pilot contamination. Thus, \( \tau_T = \tau_p / (\tau_p \geq K) \) holds and the minimum \( \tau \) can be easily obtained as

\[ \tau^* = \frac{T_p}{T} \]

\[ \text{APPENDIX C: PROOF OF LEMMA 2} \]

From (15) and (24), it is observed that the maximisation of \( R_\tau \) is equivalent to maximising \( \chi_{j,k} \) and consequently we only focus on \( \chi_{j,k} \) in this section, which is a function of \( \rho_{j,k} \). According to [19], we know that there exists at least one \( \rho_{j,k} \) to maximise \( \chi_{j,k} \). In order to obtain the optimal PS ratio, we calculate the derivative \( \frac{d\chi_{j,k}(\rho_{j,k})}{d\rho_{j,k}} \) and then obtain an equation as follows

\[ \left( \xi_{\rho,1} - \xi_{\rho,2} \right) \rho_{j,k} - 2 \xi_{\rho,1} \xi_{\rho,2} \rho_{j,k} + \left( \xi_{\rho,1} - \xi_{\rho,2} \right) \xi_{\rho,4} = 0. \]

\[ \text{(C.1)} \]

Next, we obtain the optimal PS ratio by solving (C.1).

When \( \xi_{\rho,1} \xi_{\rho,2} - \xi_{\rho,2} \xi_{\rho,4} = 0 \), the optimal PS ratio is

\[ \rho_{j,k}^* = \frac{\xi_{\rho,1} - \xi_{\rho,2}}{2 \xi_{\rho,1}}. \]

\[ \text{(C.2)} \]

When \( \xi_{\rho,1} \xi_{\rho,2} - \xi_{\rho,2} \xi_{\rho,4} \neq 0 \), the discriminant of the quadratic equation is

\[ \Delta = \left( -2 \xi_{\rho,1} \xi_{\rho,2} \rho_{j,k} \right)^2 - 4 \left( \xi_{\rho,1} \xi_{\rho,2} \xi_{\rho,4} \right) \left( \xi_{\rho,1} - \xi_{\rho,2} \right) \xi_{\rho,4} \]

\[ = 4 \xi_{\rho,2} \xi_{\rho,4} \left( \xi_{\rho,1} \xi_{\rho,4} + \xi_{\rho,1} \xi_{\rho,3} - \xi_{\rho,2} \xi_{\rho,3} \right). \]
Note that $\xi_{\rho,1} > \xi_{\rho,2}$ can be satisfied as the power of noise is relatively small [23]. Thus, the discriminant of the quadratic equation is positive and there are two solutions as follows

$$\rho_{l,k}^{1,*} = \frac{2\xi_{\rho,1}\xi_{\rho,4} - \sqrt{\Delta}}{2(\xi_{\rho,1}\xi_{\rho,4} - \xi_{\rho,2}\xi_{\rho,3})},$$ (C.4)

$$\rho_{l,k}^{2,*} = \frac{2\xi_{\rho,1}\xi_{\rho,4} + \sqrt{\Delta}}{2(\xi_{\rho,1}\xi_{\rho,4} - \xi_{\rho,2}\xi_{\rho,3})}. \quad \text{(C.5)}$$

It is obvious that $\rho_{l,k}^{2,*} > 1$ cannot satisfy (25c) and thus is not the solution we want. $\rho_{l,k}^{1,*}$ is the solution we want as it can satisfy (25c).

The proof of Lemma 2 is completed.