Sensitivity and Asymptotic Analysis of Inter-cell Interference against Pricing for Multi-Antenna Base Stations

Ye Liu, *Member, IEEE*, Sangarapillai Lambotharan, *Senior Member, IEEE*, Mahsa Derakhshani, *Member, IEEE*, Arumugam Nallanathan, *Fellow, IEEE*, and Kai-Kit Wong, *Fellow, IEEE*

***Abstract*—We thoroughly investigate the downlink beamforming problem of a two-tier network in a reversed time-division duplex (R-TDD) system, where the interference leakage from a tier-2 base station (BS) towards nearby uplink tier-1 BSs is controlled through pricing. We show that soft interference control through the pricing mechanism does not undermine the ability to regulate interference leakage while giving flexibility to sharing the spectrum. We then analyze and demonstrate how the interference leakage is related to the variations of both the interference prices and the power budget. Moreover, we derive a closed-form expression for the interference leakage in an asymptotic case where both the charging BSs and the charged BS are equipped with large numbers of antennas, which provides further insights into the lowest possible interference leakage that can be achieved by the pricing mechanism.**

***Index Terms*—Beamforming, convex optimization, interference management, uplink-downlink duality, pricing, large-scale antenna array, heterogeneous networks (HetNets).**

I. Introduction

A. Motivation

With the increasing density and heterogeneity of today's cellular networks, interference remains as one of the key issues impacting the overall system performance [2, 3]. Equipped with multiple antennas, beamforming techniques allow base stations (BSs) to exploit the spatial domain and manage interference more effectively than single-antenna BSs [4]. More degrees of freedom (DoFs) become available when large-scale antenna arrays are installed on the BSs, which boost the network capacity [5] and the ability of interference management [6].

From the optimization point of view, the centralized beamforming designs would offer the best possible network performance in terms of interference management. However, as the network size grows large, it becomes increasingly complex to design the overall beamforming patterns of the BSs [7]. Also, obtaining the network-wide up-to-date channel state information, which is crucial for enabling coordinations among the BSs, becomes more difficult for large-scale networks [8]. The cost of installing the low-delay network-wide backhaul links is high and can be even infeasible [9], making the distributed designs an indispensable alternative. Pricing-based interference control arises as a natural distributed interference management scheme [10], where distributed beamforming designs based on pricing can achieve the same network utility compared to the centralized counterpart [11].

The effectiveness of beamforming techniques for interference mitigation relies on the accuracy of the channel state information (CSI). By utilizing the channel reciprocity, time division duplex (TDD) schemes are deemed as efficient protocols for the BSs to acquire CSI with minimum resources spent on signalling for channel measurement especially when the numbers of antennas equipped at the BSs are more than the number of users [12]. Among the family of TDD schemes, the reverse TDD (R-TDD) system, where the directions of the transmissions in two co-located tiers are reversed to each other, has the unique advantage that the interference subspace between a pair of BSs from different tiers can be accurately measured due to their fixed locations [13]. Also, because more antennas can be mounted to the BSs in tier-1 compared to the BS in tier-2, the capacity of the uplink of the tier-2 BSs in an R-TDD system can be significantly improved because tier-1 BSs have more DoFs to avoid interference leaking towards the tier-2 BSs [14, 15]. The R-TDD mode has also been shown to improve the sum DoF of a two-cell network as compared to the conventional TDD mode, especially in the situation where the two BSs are equipped with different numbers of antennas [16].

Motivated by the potential advantages of R-TDD systems, in this paper, we perform a detailed study on the interference leakage from a tier-2 BS towards nearby tier-1 BSs against pricing. Specifically,

1. We show that a pricing mechanism, in which the tier-2 BS is charged for causing interference to the tier-1 BSs, can limit the cross-tier interference leakage below any feasible targets. At the same time, the pricing mechanism provides more flexibility in spectrum sharing compared to the hard interference control counterpart.
2. We then analyze the behavior of the tier-2 BS under

Y. Liu, S. Lambotharan and M. Derakhshani are with the Signal Processing and Networks Research Group, Wolfson School, Loughborough University, Leicestershire, LE11 3TU, United Kingdom (email: y.liu6@lboro.ac.uk; s.lambotharan@lboro.ac.uk; m.derakhshani@lboro.ac.uk).

A. Nallanathan is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London, E1 4FZ, United Kingdom (email: a.nallanathan@qmul.ac.uk)).

K.-K. Wong is with the Department of Electronic and Electrical Engineering, University College London, London, WC1E 6BT, United Kingdom (email: kai-kit.wong@ucl.ac.uk).

This work has been supported by the Engineering and Physical Science Research Council of the UK, EPSRC, under the grants EP/M015475, EP/M016145, and EP/M016005.

Part of the work has been accepted by IEEE International Conference on Communications (ICC) in Kuala Lumpur, 2016 [1].

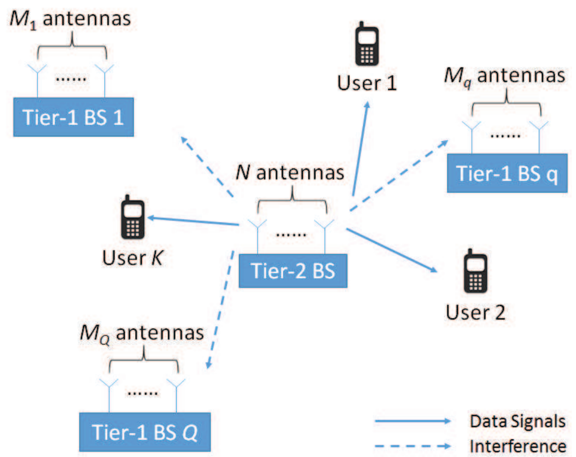


Fig. 1. The system model studied in this paper. The tier-2 BS serves K users in its downlink, while the tier-1 BSs charge the tier-2 BS for the interference caused by the latter.

the pricing mechanism. We prove that the interference leakage towards a tier-1 BS is a decreasing function of the price issued by the same tier-1 BS. Also, we show that the power budget of the tier-2 BS limits the latter's capability of avoiding interference leakage towards the tier-1 BSs. Simulation studies verify these proved behaviors of the tier-2 BS under the pricing mechanism.

3. To echo the trend of upgrading BSs with large scale antenna arrays, we derive a closed-form expression for the average interference leakage as a function of the prices and other system parameters when all BSs are equipped with large numbers of antennas. Verified by simulations, the derived expression allows us to accurately predict the interference leakage under such scenarios, which facilitates the implementation of the pricing-based interference control.

B. *Related work*

Pricing-based distributed coordination of power allocation and beamforming design has been studied in various network models, where the main focus has been on finding the best charges so that the performance of the distributed optimizations can be as good as the performance of the corresponding centralized designs. A detailed discussion on distributed pricing algorithms for power control and beamforming optimization can be found in [10]. Therein, the authors describe an iterative algorithm in a peer-to-peer network setting, where each transmitter first updates its power/beamformer to maximize its own payoff and then announces a price to nearby transmitters. The price is calculated as the marginal loss of its utility with respect to per unit of interference. It is also revealed in [10] that the iterative algorithm converges if the best response update of an underlying non-cooperative game converges, and the Nash equilibrium point must satisfy the Karush-Kuhn-Tucker (KKT) conditions of the corresponding centralized optimization problem.

The idea of [10] has been applied for distributed beamforming designs in several instances of multi-cellular networks. For example, [11] formulates the pricing-based power minimization problem as a non-cooperative game, where a Pareto-optimal solution can be found by the fixed-point iteration method [17]. [18] studies the maximization of a general concave network-wise utility function in a multi-carrier scenario. [19] focuses on the sum-rate maximization problem in a single-carrier multi-cell network. The problem of energy efficiency maximization based on distributed beamforming design is considered in [20]. [21] designs the precoding matrices of multiple-input multiple-output (MIMO) cognitive radio networks in a distributed manner where the noise leakage towards the primary network is maintained to be below some given thresholds.

In heterogeneous networks (HetNets) where more than one tier of transmitter-receiver pairs coexist, cross-tier interference control is an important issue for performance optimization. Pricing-based interference control mechanism, owning to its capability of influencing the behaviors of the transmitting nodes in the networks, has been considered in two-tier HetNets [22]. Also, Stackelberg game model [23] is applicable for interference control purposes in hierarchical HetNets, where an entity with high privilege (usually a macro BS) is able to issue prices against interference and the lower tier must pay for the caused interference accordingly [24, 25, 26]. The same pricing mechanism has also been applied in radar communication networks recently [27].

This paper differs from the previous studies in that we are interested in how a tier-2 BS behaves in the pricing-based interference control mechanism. Building on [1] in which we mainly analyze the effect of power budget of the tier-2 BS on the interference leakage towards a single tier-1 BS, in this paper, we conduct detailed analyses on a more general system setting with multiple tier-1 BSs. We show the equivalence between interference control by pricing and by issuing explicit interference constraints. Then, we show how the transmission power of the tier-2 BS and the interference leakage from the tier-2 BS changes with respect to the charge and the power budget at the tier-2 BS. Moreover, we analyze the interference leakage in the asymptotic scenario where BSs from both tiers are equipped with a large number of antennas, since it is expected that large-scale antenna arrays are to be incorporated with the existing HetNets to meet the huge demand of future data services [28].

C. *Organization*

The rest of this paper is organized as follows. Section II presents the system model of the two-tier R-TDD system. Section III formulates the downlink beamforming problems and analyzes the behavior of the tier-2 BS under the pricing mechanism. Section IV presents a closed-form expression for the interference leakage in the large-scale antenna array scenario. Section V presents simulation studies. Finally, Section VI concludes the findings of this study.

Notation: We use bold lower-case letters and bold upper-case letters to denote column vectors and matrices, respectively. Also, we use calligraphic upper-case letters to denote sets, e.g., set $\mathcal{A}$. The Hermitian of the vector $\mathbf{a}$ is denoted as $\mathbf{a}^\dagger$, and the transpose of $\mathbf{a}$ is denoted as $\mathbf{a}^T$. The positive semi-definiteness of matrix $\mathbf{A}$ is described as $\mathbf{A} \succeq 0$. The optimal value of a variable is marked by $(\cdot)^*$. The n-by-n identity matrix is denoted as $\mathbf{I}_n$. The nearest integer to a is denoted as $\lfloor a \rceil$.

II. System Model

Consider the system model in Fig. 1 where the tier-2 BS serves K single-antenna users in the downlink. Each of the Q tier-1 BSs, located within the transmission range of the tier-2 BS, receives signals in the uplink on the same frequency band used by the tier-2 BS, and therefore the tier-2 BS causes interference to the tier-1 BSs. The tier-2 BS is equipped with N antennas and the q-th tier-1 BS is equipped with M_q antennas. Denote $\mathcal{K} = \{1, 2, ..., K\}$ as the set of the downlink users connected to the tier-2 BS. Let $\mathbf{w}_k \in \mathbb{C}^{N\times 1}$ be the beamformer applied for user k at the tier-2 BS, where $\mathbb{C}$ is the set of complex numbers. The symbol vector transmitted by the tier-2 BS at a particular time instance n can be written as

$$\mathbf{x}(n) = \sum\nolimits_{\forall k \in \mathcal{K}} \mathbf{w}_k x_k(n), \tag{1}$$

where $x_k(n)$ is the data symbol of user k at time instance n. The power of $x_k(n)$ is normalized to one, and the squared Euclidean norm of $\mathbf{w}_k$ provides the actual transmission power for user k. Let $\mathbf{h}_k \in \mathbb{C}^{N\times 1}$ be the channel between the tier-2 BS and its k-th user, where elements in $\mathbf{h}_k$ are i.i.d. complex Gaussian random variables with zero mean and unit variance. The k-th user of the tier-2 BS has a minimum SINR requirement which is denoted as $\gamma_k > 0$. Also, the tier-2 BS has perfect knowledge of the channels between itself and the users it serves.

The received signal at the k-th downlink user of the tier-2 BS is given as

$$y_k(n) = (\mathbf{h}_k(n))^T \mathbf{x}(n) + z_k(n), \tag{2}$$

where $z_k(n)$ is the additive white Gaussian noise (AWGN) at user k, of which the variance is denoted by σ_k^2. Assume $x_{k_1}(n)$ and $z_{k_2}(n)$ are pairwise independent for all k_1 and k_2. Similarly, assume that $x_{k_1}(n)$ and $x_{k_2}(n)$ are also independent for all $k_1 \neq k_2$. The SINR at the k-th downlink user of the tier-2 BS is given as

$$\text{SINR}_k = \frac{|\mathbf{w}_k^\dagger \mathbf{h}_k|^2}{\sum_{j\in\mathcal{K}\setminus\{k\}} |\mathbf{w}_j^\dagger \mathbf{h}_k|^2 + \sigma_k^2}, \tag{3}$$

For notation simplicity, in the rest of this paper, the time index superscripts are omitted because we assume that all channels are quasi-static.

Denote $\mathbf{\Theta}_q \in \mathbb{C}^{N\times M_q}$ as the coefficient matrix of the channel between the tier-2 BS and the q-th tier-1 BS. Also, denote $\boldsymbol{\theta}_{q,m}$ as the channel vector between the tier-2 BS and the m-th antenna of the q-th tier-1 BS, where $m \in \mathcal{M}_q$ and $\mathcal{M}_q = \{1, 2, ..., M_q\}$. We assume that the tier-2 BS knows $\mathbf{\Theta}_q$ $\forall q$.

The tier-1 BSs would share their uplink spectrum with the tier-2 BS if the interference caused by the tier-2 BS towards the q-th tier-1 BS is kept below L_q. The tier-2 BS aims to minimize its transmission power (i.e., $\sum_{\forall k\in\mathcal{K}} ||\mathbf{w}_k||^2$), subject to SINR targets for its users, power budget constraint, and interference leakage requirements. To achieve these goals, the tier-2 BS needs to optimally choose its downlink beamformers solving the following optimization problem

TABLE I
Summary of Notation

Notation	Definition
c_q	Per-unit charge of the interference leakage issued by the q-th tier-1 BS
$\mathbf{h}_k$	Channel between user k and the tier-2 BS
K	Number of the downlink users served by the tier-2 BS
$\mathcal{K}$	The set of the downlink users served by the tier-2 BS
L_q	Maximum allowable interference leakage from the tier-2 BS to the q-th tier-1 BS
M_q	Number of antennas equipped on the q-th tier-1 BS
N	Number of antennas equipped on the tier-2 BS
$P_{\max}$	The power available at the tier-2 BS
Q	Number of tier-1 BSs
$\mathcal{Q}$	Index set of the tier-1 BSs
γ_k	SINR target of the k-th downlink user of the tier-2 BS
$\mathbf{\Theta}_q$	Channel between the tier-2 BS and the q-th tier-1 BS
σ_k^2	The AWGN power at the k-th user

$$(\mathbf{P1}):\ \underset{\{\mathbf{w}_k|k\in\mathcal{K}\}}{\text{minimize}} \sum_{\forall k\in\mathcal{K}} ||\mathbf{w}_k||^2, \tag{4a}$$

$$\text{subject to } \frac{|\mathbf{w}_k^\dagger \mathbf{h}_k|^2}{\sum_{j\in\mathcal{K}\setminus\{k\}} |\mathbf{w}_j^\dagger \mathbf{h}_k|^2 + \sigma_k^2} \geq \gamma_k, \forall k \in \mathcal{K}, \tag{4b}$$

$$\sum_{\forall k\in\mathcal{K}} ||\mathbf{w}_k||^2 \leq P_{\max}, \tag{4c}$$

$$\sum_{\forall k\in\mathcal{K}} ||\mathbf{w}_k^\dagger \mathbf{\Theta}_q||^2 \leq L_q,\ \forall q, \tag{4d}$$

where $P_{\max}$ denotes the power budget of the tier-2 BS and L_q is the maximum allowable interference power from the tier-2 BS to the q-th tier-1 BS. Table I summaries the definitions of the notation.

III. Interference Control with Pricing: Sensitivity Analysis

In this section, we first show that the interference control problem in P1 can be realized by a pricing mechanism. We then analyze the behavior of the tier-2 BS under the pricing mechanism for interference control.

A. Downlink Beamforming with Pricing

Based on the dual decomposition technique [29], the problem in P1 can be decomposed into two sub-problems at different levels. The sub-problem at the lower level concerns with the design of the downlink beamformers of the tier-2 BS, i.e.,

$$(\textbf{P1-1}): \ \underset{\{\mathbf{w}_k|k\in\mathcal{K}\}}{\text{minimize}} \sum_{\forall k\in\mathcal{K}} ||\mathbf{w}_k||^2 + \sum_{q\in\mathcal{Q}} c_q \sum_{\forall k\in\mathcal{K}} ||\mathbf{w}_k^\dagger \mathbf{\Theta}_q||^2, \tag{5a}$$

$$\text{subject to } \frac{|\mathbf{w}_k^\dagger \mathbf{h}_k|^2}{\sum_{j\in\mathcal{K}\setminus\{k\}} |\mathbf{w}_j^\dagger \mathbf{h}_k|^2 + \sigma_k^2} \ge \gamma_k, \forall k \in \mathcal{K}, \tag{5b}$$

$$\sum_{\forall k\in\mathcal{K}} ||\mathbf{w}_k||^2 \le P_{\max}, \tag{5c}$$

where c_q is the Lagrange multiplier associated with the q-th constraint in (4d). At the higher level, the other sub-problem of P1 deals with the optimization of the Lagrange multipliers c_q for the tier-1 BSs, i.e.,

$$(\textbf{P1-2}):$$

$$\underset{\{c_q|q\in\mathcal{Q}\}}{\text{minimize}} \sum_{\forall k\in\mathcal{K}} ||\mathbf{w}_k(\mathbf{c})||^2 + \sum_{q\in\mathcal{Q}} c_q \sum_{\forall k\in\mathcal{K}} ||\mathbf{w}_k(\mathbf{c})^\dagger \mathbf{\Theta}_q||^2 - \sum_{q\in\mathcal{Q}} c_q L_q, \tag{6a}$$

$$\text{subject to } c_q \ge 0, \ q \in \mathcal{Q}. \tag{6b}$$

In P1-2, $\mathbf{c}$ is the Q-by-1 vector containing $[c_1, c_2, \cdots, c_Q]$, and $\mathbf{w}_k(\mathbf{c})$ is the k-th downlink beamformer of the tier-2 BS which is a function of $\mathbf{c}$. We can interpret the formulations in P1-1 and P1-2 as a pricing mechanism for interference control, where the q-th tier-1 BS charges the tier-2 BS c_q for each unit of interference leaked to the former.

The equivalence between P1 and the problems P1-1 and P1-2 suggests that the tier-1 BSs can use the soft interference control mechanism in P1-1 to achieve the same interference control targets set out in P1. In other words, without compromising the capability on interference control, the pricing mechanism can serve as an alternative for interference control purposes when the hard interference leakage budgets in (4d) can be relaxed.

For any given price vector $\mathbf{c}$, the problem in P1-1 can be solved by the uplink-downlink duality between the downlink beamforming problem and its uplink counterpart [30]. Denote $\{\lambda_k \ge 0 | k \in \mathcal{K}\}$ and $\mu \ge 0$ as the Lagrange multipliers of (5b) and (5c), respectively. The Lagrangian dual problem of P1-1 is given as

$$(\textbf{P1-1-Dual}): \ \underset{\{\lambda_k|k\in\mathcal{K}\},\mu}{\text{maximize}} \sum_{\forall k\in\mathcal{K}} \lambda_k \sigma_k^2 - \mu P_{\max}, \tag{7a}$$

$$\text{subject to } (\mu+1)\mathbf{I}_N + \sum_{q\in\mathcal{Q}} c_q \cdot \mathbf{\Theta}_q \mathbf{\Theta}_q^\dagger + \sum_{j\in\mathcal{K}\setminus\{k\}} (\lambda_j \mathbf{h}_j \mathbf{h}_j^\dagger) - \frac{\lambda_k}{\gamma_k} \mathbf{h}_k \mathbf{h}_k^\dagger \succeq 0, \ \forall k \in \mathcal{K}. \tag{7b}$$

The constraint in (7b) implies that

$$\mathbf{v}_k^\dagger [(\mu+1)\mathbf{I}_N + \sum_{q\in\mathcal{Q}} c_q \cdot \mathbf{\Theta}_q \mathbf{\Theta}_q^\dagger + \sum\nolimits_{j\in\mathcal{K}\setminus\{k\}} (\lambda_j \mathbf{h}_j \mathbf{h}_j^\dagger) - \frac{\lambda_k}{\gamma_k} \mathbf{h}_k \mathbf{h}_k^\dagger] \mathbf{v}_k \ge 0$$

$$\Rightarrow \frac{\lambda_k \mathbf{v}_k^\dagger \mathbf{h}_k \mathbf{h}_k^\dagger \mathbf{v}_k}{\mathbf{v}_k^\dagger [\sum_{j\in\mathcal{K}\setminus\{k\}} (\lambda_j \mathbf{h}_j \mathbf{h}_j^\dagger) + \mathbf{\Sigma}(\mathbf{c},\mu)] \mathbf{v}_k} \le \gamma_k, \ k \in \mathcal{K}, \tag{8}$$

where $\mathbf{v}_k$ is the k-th uplink beamforming vector and

$$\mathbf{\Sigma}(\mathbf{c},\mu) \triangleq (\mu+1)\mathbf{I}_N + \sum_{q\in\mathcal{Q}} c_q \cdot \mathbf{\Theta}_q \mathbf{\Theta}_q^\dagger. \tag{9}$$

Then, P1-1-Dual is equivalent to the following problem

$$(\textbf{P1-1-Dual-Eqv}): \ \underset{\{\lambda_k,\mathbf{v}_k|k\in\mathcal{K}\},\mu}{\text{maximize}} \sum_{\forall k\in\mathcal{K}} \lambda_k \sigma_k^2 - \mu P_{\max}, \tag{10}$$

$$\text{subject to (8).}$$

Using similar techniques in [30], it can be shown that

$$\mathbf{v}_k^* = \frac{\mathbf{A}_k^{-1} \mathbf{h}_k}{\mathbf{h}_k^\dagger \mathbf{A}_k^{-1} \mathbf{h}_k}, \tag{11}$$

and the optimal downlink beamformer for user k in P1 can be calculated as

$$\mathbf{w}_k^* = \varepsilon_k \mathbf{v}_k^*, \tag{12}$$

where

$$\mathbf{A}_k \triangleq \sum_{j\in\mathcal{K}\setminus\{k\}} (\lambda_j^* \mathbf{h}_j \mathbf{h}_j^\dagger) + \mathbf{\Sigma}(\mathbf{c},\mu^*), \tag{13}$$

$$[\varepsilon_1^2 \ \varepsilon_2^2 \ \cdots \ \varepsilon_K^2]^T = \mathbf{\Omega}^{-1} [\sigma_1^2 \ \sigma_2^2 \ \cdots \ \sigma_K^2]^T, \tag{14}$$

$$\mathbf{\Omega}_{[k,k]} = \frac{|(\mathbf{v}_k^*)^\dagger \mathbf{h}_k|^2}{\gamma_k} = \gamma_k^{-1}, \tag{15}$$

$$\mathbf{\Omega}_{[j,k]} = -|(\mathbf{v}_j^*)^\dagger \mathbf{h}_k|^2, \ \forall j \ne k. \tag{16}$$

From (11), (12), and (13), it is clear that the tier-1 BSs can change the behavior of the tier-2 BS by adjusting $\mathbf{c}$, because $\mathbf{c}$ affects the set of optimal uplink beamformers $\{\mathbf{v}_k^* | k \in \mathcal{K}\}$, and $\{\mathbf{v}_k^* | k \in \mathcal{K}\}$ determines the directions of the optimal downlink beamformers of the tier-2 BS.

B. Behavior of the Tier-2 BS under the Pricing Mechanism

Before analyzing how the tier-2 BS responses to the pricing control, we present the following property of P1-1 that is useful for subsequent analyses.

Theorem 1. *The problem* P1-1 *is strictly convex and it satisfies strong duality.*

Proof. The proof is similar to Theorem 1 in [1] and is omitted. ∎

Suppose there is a feasible instance of P1-1 denoted as $\mathfrak{P}_1$, where $\mathbf{c} = \mathbf{c}_1$. Also, suppose that $\mathfrak{P}_2$ is another instance of P1-1 in which everything is the same as that in $\mathfrak{P}_1$ except $\mathbf{c} = \mathbf{c}_2$. Let $\{\mathbf{w}_k^*(\mathbf{c}_1) | k \in \mathcal{K}\}$ and $\{\mathbf{w}_k^*(\mathbf{c}_2) | k \in \mathcal{K}\}$ be the set of the optimal downlink beamformers of instances $\mathfrak{P}_1$ and $\mathfrak{P}_2$, respectively. Also, we define the following variables for notation simplicity,

$$\text{Tx}(\mathbf{c}) \triangleq \sum_{k\in\mathcal{K}} ||\mathbf{w}_k^*(\mathbf{c})||^2, \tag{17}$$

$$\text{IF}_q(\mathbf{c}) \triangleq \sum_{k\in\mathcal{K}} ||\left(\mathbf{w}_k^*(\mathbf{c})\right)^\dagger \mathbf{\Theta}_q||^2, \tag{18}$$

$$\text{Obj}_{-q}(\mathbf{c}) \triangleq \sum_{k\in\mathcal{K}} ||\mathbf{w}_k^*(\mathbf{c})||^2 + \sum_{q'\ne q} c_{q'} \sum_{k\in\mathcal{K}} ||\left(\mathbf{w}_k^*(\mathbf{c})\right)^\dagger \mathbf{\Theta}_{q'}||^2, \tag{19}$$

where $\text{Tx}(\mathbf{c})$ represents the optimal transmission power of the tier-2 BS, $\text{IF}_q(\mathbf{c})$ gives the interference leakage from the tier-2 BS to the q-th tier-1 BS when the former employs optimal downlink beamformers, and $\text{Obj}_{-q}(\mathbf{c})$ denotes the optimal objective function of P1-1 except the interference charge due to the q-th tier-1 BS. The next theorem shows the variation of $\text{IF}_q(\mathbf{c})$ and $\text{Obj}_{-q}(\mathbf{c})$ with respect to c_q.

Theorem 2. *For a feasible P1-1, $\text{Obj}_{-q}(\mathbf{c})$ is an increasing function of c_q, and $\text{IF}_q(\mathbf{c})$ is a decreasing function of c_q.*

Proof. See Appendix B. ■

Theorem 2 states that the interference leakage from the tier-2 BS towards the q-th tier-1 BS decreases as c_q increases while $\{c_{q'}|\forall q' \neq q\}$ remains fixed. From the equivalence between P1 and the problems P1-1 and P1-2, the fact that IF_q is a decreasing function of c_q implies that when the q-th tier-1 BS increases c_q in P1-1, it is equivalent to the situation that the same tier-1 BS reduces L_q in P1.

Also, when $Q = 1$, Theorem 2 implies that $\text{Tx}(\mathbf{c})$ increases as the only tier-1 BS increases its price. To see this, observe that $\text{Obj}_{-q}(\mathbf{c})$ reduces to $\sum_{k\in\mathcal{K}} ||\mathbf{w}_k^*(\mathbf{c})||^2$ when $Q = 1$. The intuition behind this observation can be understood from the equivalence between P1 and the problems in P1-1 and P1-2, that increasing c_q when $Q = 1$ is equivalent to tightening the constraint in (4d). The same observation can be extended, as shown in the next corollary, to the case for $Q > 1$ when all tier-1 BSs increase their prices by the same factor.

Corollary 1. $\text{Tx}(\mathbf{c}_1) \leq \text{Tx}(\mathbf{c}_2)$ *if $\mathbf{c}_2 = \xi \cdot \mathbf{c}_1$, where $\xi \geq 1$.*

Proof. See Appendix B. ■

On the other hand, when $Q > 1$ increases, the fact that $\text{Obj}_{-q}(\mathbf{c})$ increases when c_q increases does not suggest a fixed pattern of change on the transmission power of the tier-2 BS. We will demonstrate in Fig. 3 of the simulation studies that the transmission power of tier-2 BS can either increase or decrease when c_q increases.

Given a feasible instance of P1-1, Theorem 2 and Corollary 1 indicate the possibility that the tier-2 BS would transmit at its maximum power for some charge $\mathbf{c}$. The next lemma and its corollary reveal the behavior of the tier-2 BS when it operates at its maximum power and the charge increases.

Lemma 1. *Suppose $\mathfrak{D}_1$ and $\mathfrak{D}_2$ are two feasible instances of* P1-1-Dual*, where the parameters in $\mathfrak{D}_1$ are the same as those in $\mathfrak{D}_2$ except $\mathbf{c} = \mathbf{c}_1$ in $\mathfrak{D}_1$ and $\mathbf{c} = \mathbf{c}_2 = \xi \cdot \mathbf{c}_1$ in $\mathfrak{D}_2$, where $\xi \geq 1$. Also, suppose the tier-2 BS transmits at its maximum power in $\mathfrak{D}_1$. Denote $\{\lambda_{1,\mathbf{c}_1}^*, \lambda_{2,\mathbf{c}_1}^*, ..., \lambda_{K,\mathbf{c}_1}^*, \mu_{\mathbf{c}_1}^*\}$ and $\{\lambda_{1,\mathbf{c}_2}^*, \lambda_{2,\mathbf{c}_2}^*, ..., \lambda_{K,\mathbf{c}_2}^*, \mu_{\mathbf{c}_2}^*\}$ as the optimal solutions of $\mathfrak{D}_1$ and $\mathfrak{D}_2$, respectively. Then,*

$$\frac{\lambda_{k,\mathbf{c}_2}^*}{\lambda_{k,\mathbf{c}_1}^*} = \frac{\mu_{\mathbf{c}_2}^* + 1}{\mu_{\mathbf{c}_1}^* + 1} = \xi. \tag{20}$$

Proof. See Appendix C. ■

Corollary 2. *When the tier-2 BS uses its maximum power at $\mathbf{c} = \mathbf{c}_1$, it will use its maximum power for any $\mathbf{c} = \xi \cdot \mathbf{c}_1$, where $\xi \geq 1$. Moreover, the optimal beamformers of the tier-2 BS when $\mathbf{c} = \xi \cdot \mathbf{c}_1$ are the same as those when $\mathbf{c} = \mathbf{c}_1$.*

Proof. The first part of the corollary is a direct consequence of Theorem 1. The second part of the corollary can be seen by plugging (20) into (11). ■

Corollary 2 suggests that the power budget of the tier-2 BS is related to the amount of interference leakage from the tier-2 BS towards a tier-1 BS, that the tier-2 BS becomes more able to reduce interference leakage if its power budget increases. We make this point clear in the following sensitivity analysis on the power budget of the tier-2 BS.

Theorem 3. *Suppose $\mathfrak{P}_3$ is a feasible instance of P1-1 where the tier-2 BS uses its maximum power, and denote the optimal objective function of $\mathfrak{P}_3$ as* Obj_3*. Let* Obj_4 *denote the optimal objective function of $\mathfrak{P}_4$, where $\mathfrak{P}_4$ is exactly the same as $\mathfrak{P}_3$ except P_{max} is reduced by $\delta > 0$. Then,* $\text{Obj}_4 > \text{Obj}_3$.

Proof. Because P1-1 satisfies strong duality, according to [31], we have

$$\text{Obj}_4 \geq \text{Obj}_3 + \mu^* \delta. \tag{21}$$

Then, because the tier-2 BS uses full power, $\mu^* > 0$ due to the complementary slackness, and the theorem follows. ■

We can make the following two observations from Theorem 3

1) When $Q = 1$, because the tier-2 BS transmits at a lower power level in $\mathfrak{P}_4$ than in $\mathfrak{P}_3$, (21) suggests that the interference leakage towards the single tier-1 BS increases when the saturated power budget of the tier-2 BS decreases.
2) When $Q > 1$ and $c_q = c \; \forall q \in \mathcal{Q}$, the objective function of P1-1 reduces to $\text{Tx}(c) + c \cdot \sum_{q=1}^{Q} \text{IF}_q(c)$. Because the transmission power of the tier-2 BS in $\mathfrak{P}_4$ is less than that in $\mathfrak{P}_3$, the total interference leakage towards all tier-1 BSs in $\mathfrak{P}_4$ is larger than that in $\mathfrak{P}_3$.

IV. Asymptotic Analysis of Interference under Pricing

In the previous section, we have analyzed the behavior of the tier-2 BS against pricing. To further enhance our understanding of the cross-tier interference against pricing, in this section, we attempt to derive a closed-form expression for the interference leakage. We show that such expressions can be obtained when both $\{M_q|q \in \mathcal{Q}\}$ and N grow large, where such scenarios are expected to be proliferated in the coming 5G era [14].

We start from the channel model of the system where all BSs are equipped with large scale antenna arrays. We assume that the distances between the antennas at each tier-1 BS are far enough, such that the small-scale fading experienced by the received signals at different antennas are independent of each other. Moreover, we assume that the distance between the tier-2 BS and each tier-1 BS is much larger than the antenna spacing at the tier-1 BS, such that the large-scale fading gain between an antenna at the tier-2 BS and an antenna at the q-th tier-1 BS is equal to a fixed value l_q[1]. We assume that

[1]The same assumption has been made in many works studying MIMO systems such as [5] and [12].

the tier-1 BSs are equipped with more antennas than the tier-2 BS, i.e., $M_q > N \ \forall q$.

Following the above assumptions, we can model the channel between the tier-2 BS and the q-th tier-1 BS as $\boldsymbol{\Theta}_q = \sqrt{l_q}\hat{\boldsymbol{\Theta}}_q$, where $\hat{\boldsymbol{\Theta}}_q = [\hat{\boldsymbol{\theta}}_{q,1} \ \hat{\boldsymbol{\theta}}_{q,2} \ ... \ \hat{\boldsymbol{\theta}}_{q,M_q}]$ consists of M_q zero-mean independent complex Gaussian vectors with covariance matrix $\mathbf{I}_N$. Note that $\hat{\boldsymbol{\Theta}}_q\hat{\boldsymbol{\Theta}}_q^\dagger$ is a central Wishart matrix with M_q DoFs and covariance matrix $\mathbf{I}_N$, i.e., $\hat{\boldsymbol{\Theta}}_q\hat{\boldsymbol{\Theta}}_q^\dagger \sim \mathcal{W}_N(M_q, \mathbf{I}_N)$ [32].

Similarly, we assume that $\mathbf{h}_j$ and $\mathbf{h}_k$ are independent for $j \neq k$, $\forall j, k \in \mathcal{K}$. Also, assume $\mathbf{h}_k = \sqrt{l_k} \cdot \hat{\mathbf{h}}_k$ for $k \in \mathcal{K}$, where l_k denotes the large-scale fading gain between the tier-2 BS and user k and $\hat{\mathbf{h}}_k$ is a zero-mean complex Gaussian vector with covariance matrix $\mathbf{I}_N$, such that $\hat{\mathbf{h}}_k\hat{\mathbf{h}}_k^\dagger \sim \mathcal{W}_N(1, \mathbf{I}_N)$. For tractability, we assume that the tier-2 BS always has sufficient transmission power budget such that $\mu^* = 0$.

Based on the above channel model, we now show that the interference leakage can be expressed in terms of system parameters in a more explicit way when the number of antennas at the tier-2 BS goes large. Denote $\text{IF}_{k,q}$ as the interference leakage from the k-th downlink beam of the tier-2 BS towards the q-th tier-1 BS. From (11) and (12), we have

$$\text{IF}_{k,q} \triangleq ||(\mathbf{w}_k^*)^\dagger \boldsymbol{\Theta}_q||^2 = \frac{\varepsilon_k^2}{l_k^2} \cdot \frac{||\hat{\mathbf{h}}_k^\dagger \mathbf{A}_k^{-1} \boldsymbol{\Theta}_q||^2}{||\hat{\mathbf{h}}_k^\dagger \mathbf{A}_k^{-1} \hat{\mathbf{h}}_k||^2}. \tag{22}$$

In the following lemma, we obtain the interference leakage when N goes large.

Lemma 2. *When $N \to \infty$,*

$$\text{IF}_{k,q} \xrightarrow[N\to\infty]{a.s.} \frac{\gamma_k l_q \sigma_k^2}{l_k} \cdot \frac{\sum_{m=1}^{M_q} \left[\frac{\text{tr}\{\mathbf{A}_{k,q,m}^{-2}\}}{(1+c_q l_q \text{tr}\{\mathbf{A}_{k,q,m}^{-1}\})^2} \right]}{\text{tr}\{\mathbf{A}_k^{-1}\}^2}, \tag{23}$$

where

$$\mathbf{A}_{k,q,m} \triangleq \mathbf{A}_k - c_q \cdot \boldsymbol{\theta}_{q,m} \boldsymbol{\theta}_{q,m}^\dagger. \tag{24}$$

Proof. See Appendix D. ■

The expression in (23) is a complicated function of $\{c_{q'} | q' \neq q\}$. To describe $\text{IF}_{k,q}$ as a closed-from expression for the prices, we need to simplify the traces of matrix inverses in (23). Assume that the tier-1 BSs set prices such that $c_q l_q \gg 1 \ \forall q$. Then, for sufficiently large $\{M_q | q \in \mathcal{Q}\}$, Appendix E shows that

$$\text{tr}\{\mathbf{A}_k^{-1}\} \approx \text{tr}\{\mathbf{T}_k^{-1}\}, \tag{25a}$$

$$\text{tr}\{\mathbf{A}_{k,q,m}^{-1}\} \approx \text{tr}\{\mathbf{T}_k^{-1}\}, \tag{25b}$$

$$\text{tr}\{\mathbf{A}_{k,q,m}^{-2}\} \approx \text{tr}\{\mathbf{T}_k^{-2}\}, \tag{25c}$$

where

$$\mathbf{T}_k \triangleq \sum_{j \in \mathcal{K}\setminus\{k\}} \lambda_k^* l_k \hat{\mathbf{h}}_j \hat{\mathbf{h}}_j^\dagger + \sum_{q\in\mathcal{Q}} c_q l_q \hat{\boldsymbol{\Theta}}_q \hat{\boldsymbol{\Theta}}_q^\dagger, \ \forall k \in \mathcal{K}. \tag{26}$$

Notice that $\mathbf{T}_k$ is a linear combination of Wishart matrices. According to [33], $\mathbf{T}_k$ can be approximated as

$$\mathbf{T}_k \approx \varpi_k d_k^{-1} \cdot \mathbf{S}_k, \tag{27}$$

where $\varpi_k \triangleq \sum_{j\neq k} \lambda_j^* l_j + \sum_q c_q l_q M_q$, $d_k \triangleq \lfloor \varpi_k^2 \varrho_k^{-1} \rceil$, $\varrho_k \triangleq \sum_{j\neq k} (\lambda_j^*)^2 l_j^2 + \sum_q c_q^2 l_q^2 M_q$, and $\mathbf{S}_k \sim \mathcal{W}_N(d_k, \mathbf{I}_N)$. To express $\text{tr}\{\mathbf{T}_k^{-1}\}$ and $\text{tr}\{\mathbf{T}_k^{-2}\}$ as continuous functions of the system parameters, Appendix F shows that

$$\lambda_k^* \approx \gamma_k l_k^{-1} N^{-1}, \ k \in \mathcal{K}. \tag{28}$$

Also, Appendix G demonstrates that

$$d_k \approx \varpi_k^2 \varrho_k^{-1} \tag{29}$$

when $c_q l_q \geq \lambda_k^* l_k \ \forall q, k$.

We are now ready to obtain a closed-form expression on the average interference leakage in terms of the system parameters $\{M_q, l_q, c_q, N, l_k, \gamma_k, \sigma_k | q \in \mathcal{Q}, \ k \in \mathcal{K}\}$. The expression for the average interference leakage is given in Theorem 4, and the properties of the average interference leakage with respect to c_q are shown in Theorem 5.

Theorem 4. *When $N \to \infty$ and $c_q l_q \gg 1 \ \forall q$, the average interference leakage from the k-th beam of the tier-2 BS towards the q-th tier-1 BS can be found as*

$$\mathbb{E}[\text{IF}_{k,q}] \approx \frac{\gamma_k \sigma_k^2 l_q M_q \cdot d_k}{l_k N \cdot (d_k - N) \cdot (1 + c_q l_q \frac{\varpi_k^{-1} d_k N}{d_k - N})^2}, \ k \in \mathcal{K}. \tag{30}$$

Proof. See Appendix H. ■

Theorem 5. *Assume* (28) *holds, then $\mathbb{E}[\text{IF}_{k,q}]$ is a decreasing function of c_q for all $k \in \mathcal{K}$. Consequently, $\mathbb{E}[\text{IF}_q]$ is a decreasing function of c_q. Moreover,*

$$\lim_{c_q\to\infty} \mathbb{E}[\text{IF}_{k,q}] = \frac{\gamma_k \sigma_k^2 l_q}{l_k} \cdot \left(\frac{M_q}{N} - 1 \right), \ k \in \mathcal{K}. \tag{31}$$

Proof. See Appendix I. ■

Theorem 5 shows that the average interference leakage from the tier-2 BS in the asymptotic region preserves its property shown in Theorem 2, that the interference leakage towards the q-th tier-1 BS reduces as c_q increases. Also, with the approximations in (28) and (29), we can express the average interference leakage from the tier-2 BS as a function of system parameters in closed-form.

More importantly, (31) gives the least interference leakage that a tier-1 BS can expect using the pricing scheme. This allows us to predict the feasibility of P1 for some given value of L_q. Observe that

1) (31) decreases as the number of antennas at the tier-2 BS increases. Also, (31) is linearly proportional to the number of antennas at the q-th tier-1 BS. Indeed, the tier-2 BS would have more DoFs to avoid causing interference towards the tier-1 BS when N increases and M_q decreases.
2) If the ratio between M_q and N is fixed, then (31) is deterministic when M_q and N both increase. In other words, it is the ratio between the number of antennas mounted on the BSs that determines the minimum interference leakage.

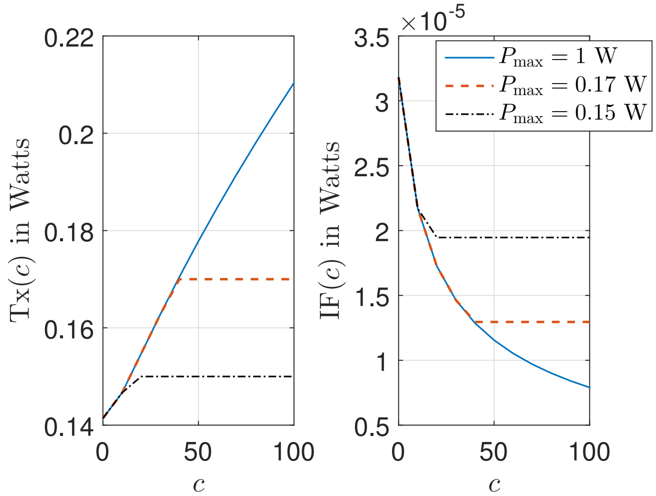


Fig. 2. A randomly generated network topology for $N = 5$, $K = 3$, and $M = 1$.

3) Moreover, (31) increases linearly with both γ_k and σ_k. This can be understood as higher SINR targets or higher noise power would impose more constraint on the tier-2 BS, which inevitably makes it harder for the tier-2 BS to avoid interference leakage.
4) Finally, (31) increases if l_q $\forall q \in \mathcal{Q}$ increases or if l_k $\forall k \in \mathcal{K}$ decreases. The former is intuitively clear, while the latter is true because the tier-2 BS would need to increase its transmission power if the large-scale fading between the tier-2 BS and the users become severe.

In the simulation studies, we will verify the accuracy of the expressions in (30) and (31).

V. Simulation Studies

In this section, we verify the analytical results in Sections III-B and IV based on the following simulation parameters. Specifically, let $\{\vartheta_i | i \in \mathcal{Q} \cup \mathcal{K}\}$ be the normalized distance between the tier-2 BS and the i-th tier-1 BS/tier-2 user. The path loss exponent is set to three, and we set $l_i = \vartheta_i^{-3}$ $\forall i \in \mathcal{Q} \cup \mathcal{K}$. Also, $\gamma_k = 1$ and $\sigma_k^2 = 0.1$ for all k.

Fig. 2 depicts the transmission power of the tier-2 BS and the interference leakage towards tier-1 for some fixed values of $\{\vartheta_i | i \in \mathcal{Q} \cup \mathcal{K}\}$ when $Q = 1$, $N = 5$, $M = 3$, and $K = 3$. We have the following observations:

1) When $P_{\max} = 1$ Watt, the transmission power of the tier-2 BS increases as c increases, and the tier-2 BS always has more power than it uses for all c values concerned in Fig. 2. Also, the tier-2 BS is forced to reduce its interference power towards the tier-1 BS as c increases. The tier-2 BS can reduce its interference towards the tier-1 BS because it has enough power for interference mitigation. This phenomenon is predicted by Theorems 2 and 1.
2) When $P_{\max} = 0.17$ Watt, the tier-2 BS has to use its maximum power when c takes the value close to 40. Also, the tier-2 BS causes the same interference towards the tier-1 BS for all c values that is larger than 40. This phenomenon reflects the validity of Corollary 2, i.e., the tier-2 BS will use the same set of downlink beamformers for all c values larger than 40.
3) When $P_{\max} = 0.15$ Watt and c is larger than or equal to 20, the interference from the tier-2 BS towards the tier-1 BS is larger than that when $P_{\max} = 0.17$ W for the same set of c values. Intuitively, the tier-2 BS would be less able to reduce interference leakage if it has less power budget; see Theorem 3 in [1].

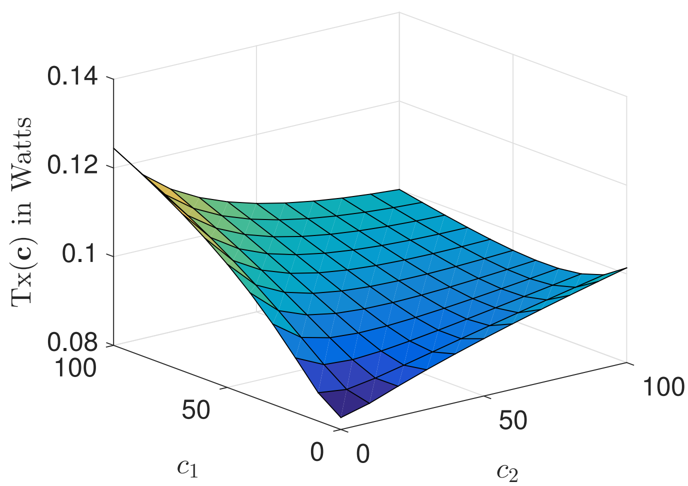
Fig. 3 further verifies the observations made in Sec. III-B when two tier-1 BSs charge the tier-2 BS for causing interference, where we set $K = 3$, $N = 4$, and $M_1 = M_2 = 3$. From Fig. 3, we can observe that the interference leakage towards the first (second) tier-1 BS decreases when c_1 (c_2) increases and c_2 (c_1) remains fixed, which matches the first part of Theorem 2. On the other hand, Fig. 3(a) shows that when c_2 increases, Tx($\mathbf{c}$) can either increase or decrease. For example, when $c_1 = 20$, Tx($\mathbf{c}$) decreases when c_2 increases from 0 to 10, and the same quantity increases when c_2 further increases. At the same time, IF$_1$($\mathbf{c}$) increases but at a decreasing rate of change. This observation can be explained by the equivalence between P1 and the formulations in P1-1 and P1-2. First, observe that increasing c_2 in the scenario of Fig. 3 has the effect of reducing L_2 and increasing L_1 in P1; this can be seen in Fig. 3(b) that IF$_1$($\mathbf{c}$) increases and IF$_2$($\mathbf{c}$) decreases as c_2 increases. Then, the feasible region of P1, due to the changes in L_1 and L_2, can be modified such that a better objective function (Tx($\mathbf{c}$)) becomes feasible.

Figures 4, 5, and 6 examine the accuracy of (30) when both the tier-2 BS and the tier-1 BSs are equipped with large number of antennas, where $Q = 3$. In all the three figures, c_q is set to be $0.1 \times l_q^{-1}$ for all q, except that c_1 in Fig. 4 is set equal to $0.1\xi l_q^{-1}$, where ξ changes between $[1, 51]$. The simulation result for $\mathbb{E}[\text{IF}_q]$ is obtained by averaging 1000 random realizations of $\hat{\mathbf{\Theta}}_q$ $\forall q$ and $\hat{\mathbf{h}}_k$ $\forall k$, where the values of ϑ_i $\forall i \in \mathcal{Q} \cup \mathcal{K}$ are fixed.

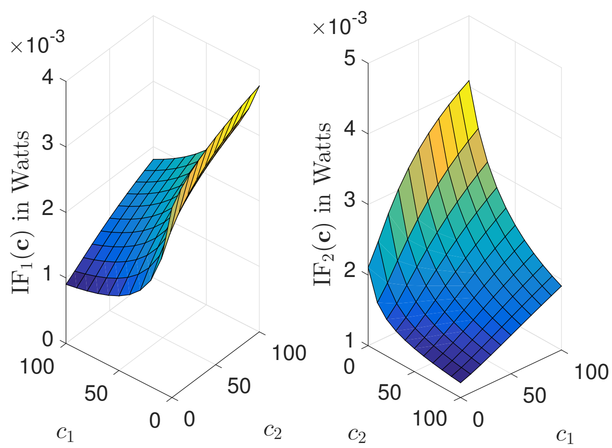
More specifically, Fig. 4 shows that the average interference leakage expression in (30) accurately predicts the trend that IF$_q$ reduces as c_q increases. Although we use the assumption that $c_q l_q \gg 1$ when deriving (30), the result from Fig. 4 indicates that (30) can accurately predict the average interference leakage for relatively small c_q for all $q \in \mathcal{Q}$. Moreover, as c_1 increases, $\mathbb{E}[\text{IF}_1]$ converges to a value that is accurately predicted by (31).

Fig. 5 shows that the prediction in (30) becomes more accurate when N increases. This is expected because (30) is obtained when $N \to \infty$. On the other hand, the results from Fig. 5 suggest that the prediction in (30) gives an error of only about 10% for $N = 20$, indicating that (30) can be used as an accurate estimate for moderate values of N. Also, we can see that the interference leakage increases as the ratio between M_q and N increases, which matches the observation from (31).

Fig. 6 depicts the average interference leakage against the number of users served by the tier-2 BS. Note that the large-scale fading gains l_k $\forall k \in \mathcal{K}$ are fixed in the simulation. Observe that the average interference leakage increases as K increases. Also, the prediction in (30) remains accurate for



(a)



(b)

Fig. 3. $\mathrm{Tx}(\mathbf{c})$ and $\mathrm{IF}_q(\mathbf{c})$ when $Q = 2$, $N = 4$, $M_1 = M_2 = 3$, and $K = 3$.

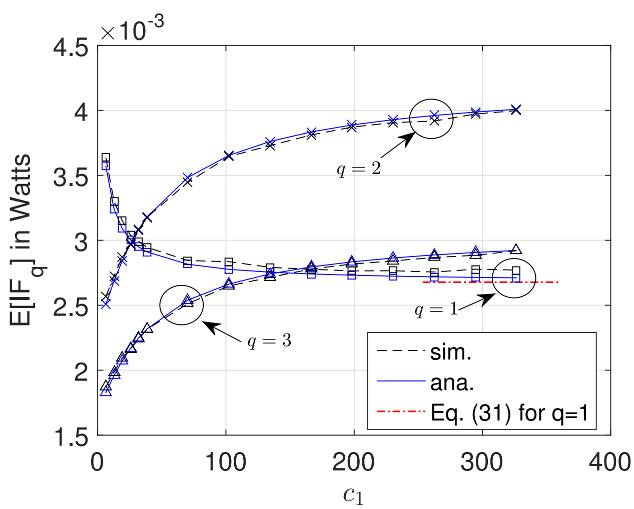


Fig. 4. $\mathbb{E}[\mathrm{IF}_q]$ against c_1, where $Q = 3$, $M_1 = M_2 = M_3 = 300$, $N = 100$, and $K = 3$. The dashed lines are results from simulation, the solid lines are results from (30), and the dash-dotted line is the lower bound of $\mathbb{E}[\mathrm{IF}_1]$ from (31).

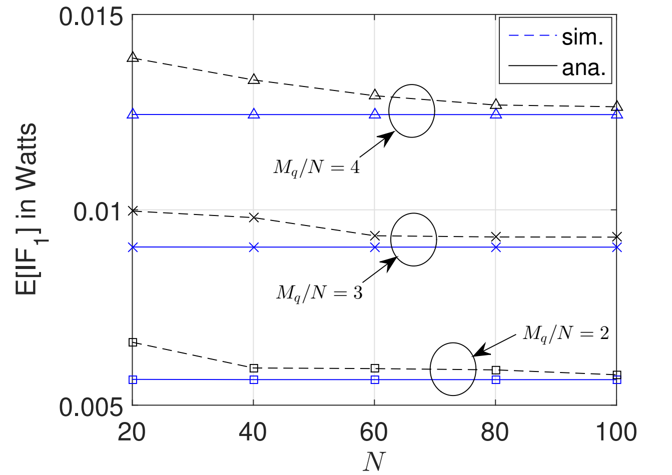


Fig. 5. $\mathbb{E}[\mathrm{IF}_1]$ for different ratios between M_q and N, where $Q = 3$, $M_1 = M_2 = M_3$, and $K = 3$.

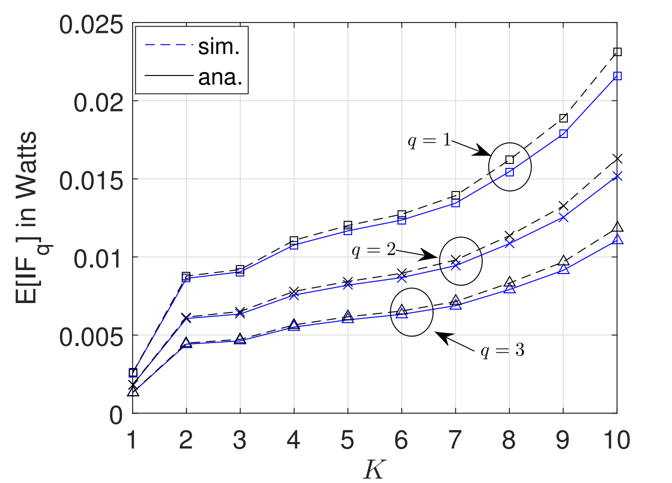


Fig. 6. $\mathbf{E}[\mathrm{IF}_q]$ against K, where $Q = 3$, $M_1 = M_2 = M_3 = 300$, and $N = 100$.

$K = 10$, where the difference between (30) and that obtained using simulation is only about 7%.

VI. Conclusion

We have analyzed a network model where multiple tier-1 BSs charge a near-by tier-2 BS for causing interference, while the tier-2 BS tries to minimize its transmission power plus the price payable to the tier-1 BSs subject to SINR targets of the tier-2 users and the power budget of the tier-2 BS. We have demonstrated that such a pricing mechanism could achieve the same interference leakage control targets, if feasible, as the mechanism where the tier-1 BSs impose explicit interference leakage constraints. Then, we have analyzed the effect of pricing and the effect of the power budget of the tier-2 BS on the pricing-based interference control scheme. Moreover, we have derived a closed-form expression for the interference power experienced by a tier-1 BS when both the tier-2 BS and the tier-1 BSs were equipped with large numbers of antennas, which demonstrated how the interference leakage was related to various system parameters such as the number of users and the number of antennas. Simulation studies have verified the accuracy of the analytical results, indicating that the derived expression can be used to predict the average interference leakage against pricing.

APPENDIX A
USEFUL LEMMAS

Lemma 3 ([34]). *Let $\mathbf{A} \in \mathbb{C}^{N\times N}$ be a Hermitian matrix that is invertible. Then, for any given vector $\mathbf{x} \in \mathbb{C}^N$ and any given scalar $\tau \in \mathbb{C}$ such that $\mathbf{A} + \tau\mathbf{x}\mathbf{x}^\dagger$ is invertible,*

$$\mathbf{x}^\dagger(\mathbf{A} + \tau\mathbf{x}\mathbf{x}^\dagger)^{-1} = \frac{\mathbf{x}^\dagger\mathbf{A}^{-1}}{1 + \tau\mathbf{x}^\dagger\mathbf{A}^{-1}\mathbf{x}}.$$

Lemma 4 ([34]). *Let $\mathbf{A} \in \mathbb{C}^{N\times N}$ and assume that $\mathbf{A}$ has uniformly bounded spectral norm. Also, let $\mathbf{x},\ \mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$). Assume that $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{A}$ are mutually independent. Then,*

1) $\mathbf{x}^\dagger\mathbf{A}\mathbf{x} - \mathrm{tr}\mathbf{A} \xrightarrow[N\to\infty]{a.s.} 0$, *and*
2) $\mathbf{x}^\dagger\mathbf{A}\mathbf{y} \xrightarrow[N\to\infty]{a.s.} 0$.

Lemma 5 ([32]). *For a central Wishart matrix $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{I}_n)$ with $m > n$,*

$$\mathbb{E}\left[\mathrm{tr}\{\mathbf{W}^{-1}\}\right] = \frac{n}{m-n}$$

while, for $m > n+1$,

$$\mathbb{E}\left[\mathrm{tr}\{\mathbf{W}^{-2}\}\right] = \frac{mn}{(m-n)^3 - (m-n)}.$$

APPENDIX B
PROOF OF THEOREM 2 AND COROLLARY 1

For the proof of Theorem 2, suppose $\mathbf{c}_2(q) > \mathbf{c}_1(q)$ and $\mathbf{c}_2(q') = \mathbf{c}_1(q')\ \forall q' \neq q$. That is to say that only the q-th tier-1 BS in $\mathfrak{P}_2$ sets its price larger than that in $\mathfrak{P}_1$, but any other tier-1 BSs sets the same price in both $\mathfrak{P}_1$ and $\mathfrak{P}_2$. The strict convexity of the problem (see Theorem 1) and the assumption that $\{\mathbf{w}_k^*(\mathbf{c}_1)|k \in \mathcal{K}\}$ is the optimal solution of $\mathfrak{P}_1$ suggest that

$$\mathrm{Obj}_{-q}(\mathbf{c}_1) + \mathbf{c}_1(q)\cdot \mathrm{IF}_q(\mathbf{c}_1) \leq \mathrm{Obj}_{-q}(\mathbf{c}_2) + \mathbf{c}_1(q)\cdot \mathrm{IF}_q(\mathbf{c}_2) \tag{32}$$

$$\Rightarrow \frac{1}{\mathbf{c}_1(q)}\mathrm{Obj}_{-q}(\mathbf{c}_1) + \mathrm{IF}_q(\mathbf{c}_1) \leq \frac{1}{\mathbf{c}_1(q)}\mathrm{Obj}_{-q}(\mathbf{c}_2) + \mathrm{IF}_q(\mathbf{c}_2). \tag{33}$$

Similarly, the assumption that $\{\mathbf{w}_k^*(\mathbf{c}_2)|k \in \mathcal{K}\}$ is the optimal solution of $\mathfrak{P}_2$ suggests that

$$\mathrm{Obj}_{-q}(\mathbf{c}_2) + \mathbf{c}_2(q)\cdot \mathrm{IF}_q(\mathbf{c}_2) \leq \mathrm{Obj}_{-q}(\mathbf{c}_1) + \mathbf{c}_2(q)\cdot \mathrm{IF}_q(\mathbf{c}_1) \tag{34}$$

$$\Rightarrow \frac{1}{\mathbf{c}_2(q)}\mathrm{Obj}_{-q}(\mathbf{c}_2) + \mathrm{IF}_q(\mathbf{c}_2) \leq \frac{1}{\mathbf{c}_2(q)}\mathrm{Obj}_{-q}(\mathbf{c}_1) + \mathrm{IF}_q(\mathbf{c}_1). \tag{35}$$

Adding (33) and (35) gives

$$\begin{aligned} &\frac{1}{\mathbf{c}_1(q)}\mathrm{Obj}_{-q}(\mathbf{c}_1) + \frac{1}{\mathbf{c}_2(q)}\mathrm{Obj}_{-q}(\mathbf{c}_2) \\ &\leq \frac{1}{\mathbf{c}_1(q)}\mathrm{Obj}_{-q}(\mathbf{c}_2) + \frac{1}{\mathbf{c}_2(q)}\mathrm{Obj}_{-q}(\mathbf{c}_1) \\ \Rightarrow &\mathbf{c}_2(q)\left(\mathrm{Obj}_{-q}(\mathbf{c}_1) - \mathrm{Obj}_{-q}(\mathbf{c}_2)\right) \\ &\leq \mathbf{c}_1(q)\left(\mathrm{Obj}_{-q}(\mathbf{c}_1) - \mathrm{Obj}_{-q}(\mathbf{c}_2)\right). \end{aligned} \tag{36}$$

Because $\mathbf{c}_2(q) > \mathbf{c}_1(q) \geq 0$, we must have

$$\mathrm{Obj}_{-q}(\mathbf{c}_1) - \mathrm{Obj}_{-q}(\mathbf{c}_2) \leq 0 \tag{37}$$

in order to have (36). Therefore, $\mathrm{Obj}_{-q}(\mathbf{c})$ is an increasing function of c_q. Adding (32) and (34) gives

$$\begin{aligned} &\mathbf{c}_1(q)\cdot \mathrm{IF}_q(\mathbf{c}_1) + \mathbf{c}_2(q)\cdot \mathrm{IF}_q(\mathbf{c}_2) \\ &\leq \mathbf{c}_1(q)\cdot \mathrm{IF}_q(\mathbf{c}_2) + \mathbf{c}_2(q)\cdot \mathrm{IF}_q(\mathbf{c}_1) \\ \Rightarrow &\mathbf{c}_1(q)\left[\mathrm{IF}_q(\mathbf{c}_1) - \mathrm{IF}_q(\mathbf{c}_2)\right] \leq \mathbf{c}_2(q)\left[\mathrm{IF}_q(\mathbf{c}_1) - \mathrm{IF}_q(\mathbf{c}_2)\right]. \end{aligned} \tag{38}$$

Again, because $\mathbf{c}_2(q) > \mathbf{c}_1(q) \geq 0$, we must have

$$\mathrm{IF}_q(\mathbf{c}_1) - \mathrm{IF}_q(\mathbf{c}_2) \geq 0. \tag{39}$$

Similar to the proof of Theorem 2, for the proof of Corollary 1, we have the following inequalities

$$\mathrm{Tx}(\mathbf{c}_1) + \sum_{q\in\mathcal{Q}} \mathbf{c}_1(q)\cdot \mathrm{IF}_q(\mathbf{c}_1) \leq \mathrm{Tx}(\mathbf{c}_2) + \sum_{q\in\mathcal{Q}} \mathbf{c}_1(q)\cdot \mathrm{IF}_q(\mathbf{c}_2) \tag{40}$$

$$\mathrm{Tx}(\mathbf{c}_2) + \sum_{q\in\mathcal{Q}} \mathbf{c}_2(q)\cdot \mathrm{IF}_q(\mathbf{c}_2) \leq \mathrm{Tx}(\mathbf{c}_1) + \sum_{q\in\mathcal{Q}} \mathbf{c}_2(q)\cdot \mathrm{IF}_q(\mathbf{c}_1) \tag{41}$$

If we multiply (40) by ξ and add the result to (41), we have

$$(\xi - 1)\mathrm{Tx}(\mathbf{c}_1) \leq (\xi - 1)\mathrm{Tx}(\mathbf{c}_2) \Rightarrow \mathrm{Tx}(\mathbf{c}_1) \leq \mathrm{Tx}(\mathbf{c}_2). \tag{42}$$

APPENDIX C
PROOF OF LEMMA 1

Let $\hat{\mu} \triangleq \mu + 1$, P1-1-Dual can be written as

$$\underset{\{\lambda_k|k\in\mathcal{K}\},\hat{\mu}}{\text{minimize}}\ \hat{\mu}P_{\max} - P_{\max} - \sum_{\forall k\in\mathcal{K}} \lambda_k\sigma_k^2, \tag{43a}$$

$$\text{subject to } \frac{\lambda_k}{\gamma_k}\mathbf{h}_k\mathbf{h}_k^\dagger - \hat{\mu}\mathbf{I}_N - \sum_{q\in\mathcal{Q}} c_q\cdot \mathbf{\Theta}_q\mathbf{\Theta}_q^\dagger - \sum_{j\in\mathcal{K}\setminus\{k\}} (\lambda_j\mathbf{h}_j\mathbf{h}_j^\dagger) \preceq 0,\ \forall k\in\mathcal{K}, \tag{43b}$$

$$-\lambda_k \leq 0,\ \forall k\in\mathcal{K}, \tag{43c}$$

$$1 - \hat{\mu} \leq 0. \tag{43d}$$

Let $\{\mathbf{Z}_k|\mathbf{Z}_k \succeq 0, \mathbf{Z}_k = \mathbf{Z}_k^\dagger, k\in\mathcal{K}\}$, $\{\alpha_k|\alpha_k \geq 0, k\in\mathcal{K}\}$, and $\beta \geq 0$ be the Lagrange multipliers of (43b), (43c), and (43d), respectively. The Lagrangian dual function of P1-1-Dual can be written as

$$\begin{aligned} g(\mathbf{Z}, \boldsymbol{\alpha}, \beta) = &\inf_{\lambda_k,\hat{\mu}} \Big\{\hat{\mu}P_{\max} - P_{\max} - \sum_{\forall k\in\mathcal{K}} \lambda_k\sigma_k^2 \\ &+ \sum_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_k\mathbf{M}_k) - \sum_{k\in\mathcal{K}} (\alpha_k\lambda_k) + \beta(1-\hat{\mu})\Big\}, \end{aligned} \tag{44}$$

where $\mathbf{Z}$ includes all elements in $\{\mathbf{Z}_k | k \in \mathcal{K}\}$, $\boldsymbol{\alpha}$ includes all elements in $\{\alpha_k | k \in \mathcal{K}\}$, and $\mathbf{M}_k \triangleq \frac{\lambda_k}{\gamma_k}\mathbf{h}_k\mathbf{h}_k^\dagger - \hat{\mu}\mathbf{I}_N - \sum_{q\in\mathcal{Q}} c_q \cdot \boldsymbol{\Theta}_q\boldsymbol{\Theta}_q^\dagger - \sum_{j\in\mathcal{K}\setminus\{k\}}(\lambda_j \mathbf{h}_j\mathbf{h}_j^\dagger)$ for $k \in \mathcal{K}$. Note that $g(\mathbf{Z}, \boldsymbol{\alpha}, \hat{\mu})$ gives a lower bound on the optimal value of P1-1-Dual because $\mathrm{tr}(\mathbf{Z}_k\mathbf{M}_k) \le 0$ for $k \in \mathcal{K}$ given that $\mathbf{Z}_k \succeq 0$, $\mathbf{Z}_k = \mathbf{Z}_k^\dagger$, $\mathbf{M}_k \preceq 0$, and $\mathbf{M}_k = \mathbf{M}_k^\dagger$ [31, pp. 52]. The Lagrangian dual problem of P1-1-Dual can thus be written as the following.

$$(\textbf{P1-1-Dual-Dual}):\ \underset{\mathbf{Z},\alpha,\beta}{\text{maximize}}\ g(\mathbf{Z}, \boldsymbol{\alpha}, \beta), \tag{45a}$$

$$\text{subject to } \mathbf{Z}_k \succeq 0, \mathbf{Z}_k = \mathbf{Z}_k^\dagger, k \in \mathcal{K}, \tag{45b}$$

$$\alpha_k \ge 0,\ \forall k \in \mathcal{K}, \tag{45c}$$

$$\beta \ge 0. \tag{45d}$$

After some manipulations, (44) can be rewritten as

$$\begin{aligned} g(\mathbf{Z}, \boldsymbol{\alpha}, \beta) = \inf_{\{\lambda_k|k\in\mathcal{K}\},\hat{\mu}} \Big\{ &\hat{\mu}[P_{\max} - \sum_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_k) - \beta] - P_{\max} \\ &+ \sum_{k\in\mathcal{K}} [\lambda_k \big(\frac{\mathrm{tr}(\mathbf{Z}_k\mathbf{h}_k\mathbf{h}_k^\dagger)}{\gamma_k} - \sum_{j\in\mathcal{K}\setminus\{k\}} \mathrm{tr}(\mathbf{Z}_j\mathbf{h}_k\mathbf{h}_k^\dagger) \\ &- \sigma_k^2 - \alpha_k\big)] - \sum_{q\in\mathcal{Q}} c_q \sum_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_k\boldsymbol{\Theta}_q\boldsymbol{\Theta}_q^\dagger) + \beta \Big\}. \end{aligned} \tag{46}$$

To avoid $g(\mathbf{Z}, \boldsymbol{\alpha}, \beta)$ from being minus infinity, we must have $P_{\max} - \sum_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_k) - \beta \ge 0$ and $\frac{\mathrm{tr}(\mathbf{Z}_k\mathbf{h}_k\mathbf{h}_k^\dagger)}{\gamma_k} - \sum_{j\in\mathcal{K}\setminus\{k\}} \mathrm{tr}(\mathbf{Z}_j\mathbf{h}_k\mathbf{h}_k^\dagger) - \sigma_k^2 - \alpha_k \ge 0$ for $k \in \mathcal{K}$. Therefore, P1-1-Dual-Dual is equivalent to the following problem

$$\underset{\mathbf{Z},\alpha,\beta}{\text{maximize}}\ -P_{\max} - \sum_{q\in\mathcal{Q}} c_q \sum_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_k\boldsymbol{\Theta}_q\boldsymbol{\Theta}_q^\dagger) + \beta, \tag{47a}$$

$$\begin{aligned}\text{subject to } &\frac{\mathrm{tr}(\mathbf{Z}_k\mathbf{h}_k\mathbf{h}_k^\dagger)}{\gamma_k} - \sum_{j\in\mathcal{K}\setminus\{k\}} \mathrm{tr}(\mathbf{Z}_j\mathbf{h}_k\mathbf{h}_k^\dagger) - \sigma_k^2 \\ &- \alpha_k \ge 0, k \in \mathcal{K}, \end{aligned}\tag{47b}$$

$$P_{\max} - \sum_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_k) - \beta \ge 0, \tag{47c}$$

$$\mathbf{Z}_k \succeq 0, \mathbf{Z}_k = \mathbf{Z}_k^\dagger, k \in \mathcal{K}, \tag{47d}$$

$$\alpha_k \ge 0,\ \forall k \in \mathcal{K}, \tag{47e}$$

$$\beta \ge 0. \tag{47f}$$

Notice that strong duality holds between P1-1-Dual and P1-1-Dual-Dual because we can always find a solution to P1-1-Dual such that all the constraints are not active [35].

Denote $\mathfrak{D}_1'$ and $\mathfrak{D}_2'$ as the corresponding P1-1-Dual-Dual instances of $\mathfrak{D}_1$ and $\mathfrak{D}_2$, respectively. Also, denote $\{\mathbf{Z}_{k,\mathbf{c}_1}^*, \alpha_{k,\mathbf{c}_1}^*, \beta_{\mathbf{c}_1}^* | k \in \mathcal{K}\}$ and $\{\mathbf{Z}_{k,\mathbf{c}_2}^*, \alpha_{k,\mathbf{c}_2}^*, \beta_{\mathbf{c}_2}^* | k \in \mathcal{K}\}$ as the optimal solutions to $\mathfrak{D}_1'$ and $\mathfrak{D}_2'$, respectively. Because we have assumed that the tier-2 BS uses all its power in $\mathfrak{D}_1$, $\hat{\mu}_{\mathbf{c}_1}^* > 1$ due to the complementary slackness. Also, from Theorem 1 and the complementary slackness, we have $\hat{\mu}_{\mathbf{c}_2}^* > 1$, Moreover, $\lambda_{k,\mathbf{c}_1}^* > 0$ and $\lambda_{k,\mathbf{c}_2}^* > 0$ for $k \in \mathcal{K}$ because the users clearly need to transmit at positive power levels in order to achieve positive SINR targets. Then, due to the complementary slackness, we have

$$\alpha_{k,\mathbf{c}_1}^* = \alpha_{k,\mathbf{c}_2}^* = 0, k \in \mathcal{K} \tag{48}$$

and

$$\beta_{\mathbf{c}_1}^* = \beta_{\mathbf{c}_2}^* = 0. \tag{49}$$

The problems $\mathfrak{D}_1'$ and $\mathfrak{D}_2'$ can then be written as

$$(\mathfrak{D}_1') : \underset{\mathbf{Z}}{\text{maximize}}\ -P_{\max} - \sum_{q\in\mathcal{Q}} \mathbf{c}_1(q) \sum_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_{k,\mathbf{c}_1}\boldsymbol{\Theta}_q\boldsymbol{\Theta}_q^\dagger), \tag{50a}$$

$$\begin{aligned}\text{subject to } &\frac{\mathrm{tr}(\mathbf{Z}_{k,\mathbf{c}_1}\mathbf{h}_k\mathbf{h}_k^\dagger)}{\gamma_k} - \sum_{j\in\mathcal{K}\setminus\{k\}} \mathrm{tr}(\mathbf{Z}_{j,\mathbf{c}_1}\mathbf{h}_k\mathbf{h}_k^\dagger) \\ &- \sigma_k^2 \ge 0, k \in \mathcal{K}, \end{aligned}\tag{50b}$$

$$P_{\max} - \sum_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_{k,\mathbf{c}_1}) \ge 0, \tag{50c}$$

$$\mathbf{Z}_{k,\mathbf{c}_1} \succeq 0, \mathbf{Z}_{k,\mathbf{c}_1} = \mathbf{Z}_{k,\mathbf{c}_1}^\dagger, k \in \mathcal{K}, \tag{50d}$$

and

$$(\mathfrak{D}_2') : \underset{\mathbf{Z}}{\text{maximize}}\ -P_{\max} - \xi \sum_{q\in\mathcal{Q}} \mathbf{c}_1(q) \sum_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_{k,\mathbf{c}_2}\boldsymbol{\Theta}_q\boldsymbol{\Theta}_q^\dagger), \tag{51a}$$

$$\begin{aligned}\text{subject to } &\frac{\mathrm{tr}(\mathbf{Z}_{k,\mathbf{c}_2}\mathbf{h}_k\mathbf{h}_k^\dagger)}{\gamma_k} - \sum_{j\in\mathcal{K}\setminus\{k\}} \mathrm{tr}(\mathbf{Z}_{j,\mathbf{c}_2}\mathbf{h}_k\mathbf{h}_k^\dagger) \\ &- \sigma_k^2 \ge 0, k \in \mathcal{K}, \end{aligned}\tag{51b}$$

$$P_{\max} - \sum_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_{k,\mathbf{c}_2}) \ge 0, \tag{51c}$$

$$\mathbf{Z}_{k,\mathbf{c}_2} \succeq 0, \mathbf{Z}_{k,\mathbf{c}_2} = \mathbf{Z}_{k,\mathbf{c}_2}^\dagger, k \in \mathcal{K}, \tag{51d}$$

respectively. Clearly, $\mathfrak{D}_1'$ and $\mathfrak{D}_2'$ are equivalent because they have the same constraints and equivalent objective functions. Therefore,

$$\begin{aligned} &\mathbf{Z}_{k,\mathbf{c}_1}^* = \mathbf{Z}_{k,\mathbf{c}_2}^* \\ \Rightarrow &\sum_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_{k,\mathbf{c}_1}^*\boldsymbol{\Theta}_q\boldsymbol{\Theta}_q^\dagger) = \sum_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_{k,\mathbf{c}_2}^*\boldsymbol{\Theta}_q\boldsymbol{\Theta}_q^\dagger),\ k \in \mathcal{K}. \end{aligned}\tag{52}$$

Now, from the strong duality between P1-1-Dual and P1-1-Dual-Dual and (49), we have

$$\begin{cases} \hat{\mu}_{\mathbf{c}_1}^* P_{\max} - \sum\limits_{k\in\mathcal{K}} (\lambda_{k,\mathbf{c}_1}^*\sigma_k^2) = -\sum\limits_{q=1}^{Q} \mathbf{c}_1(q) \sum\limits_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_{k,\mathbf{c}_1}^*\boldsymbol{\Theta}\boldsymbol{\Theta}^\dagger) \\ \hat{\mu}_{\mathbf{c}_2}^* P_{\max} - \sum\limits_{k\in\mathcal{K}} (\lambda_{k,\mathbf{c}_2}^*\sigma_k^2) = -\sum\limits_{q=1}^{Q} \mathbf{c}_2(q) \sum\limits_{k\in\mathcal{K}} \mathrm{tr}(\mathbf{Z}_{k,\mathbf{c}_2}^*\boldsymbol{\Theta}\boldsymbol{\Theta}^\dagger) \end{cases}. \tag{53}$$

From (52) and (53), we have

$$\frac{\sum_{k\in\mathcal{K}}(\lambda_{k,\mathbf{c}_2}^*\sigma_k^2) - \hat{\mu}_{\mathbf{c}_2}^* P_{\max}}{\sum_{k\in\mathcal{K}}(\lambda_{k,\mathbf{c}_1}^*\sigma_k^2) - \hat{\mu}_{\mathbf{c}_1}^* P_{\max}} = \xi. \tag{54}$$

Given $\{\lambda_{k,\mathbf{c}_1}^*, \hat{\mu}_{\mathbf{c}_1}^* | k \in \mathcal{K}\}$ as the optimal solution of $\mathfrak{D}_1$, we now propose $\{\lambda_{k,\mathbf{c}_2}^* = \xi\lambda_{k,\mathbf{c}_1}^*, \hat{\mu}_{\mathbf{c}_2}^* = \xi\hat{\mu}_{\mathbf{c}_1}^* | k \in \mathcal{K}\}$ as the optimal solution of $\mathfrak{D}_2$. We examine the validity of our proposal by checking its feasibility and the optimality. To check the feasibility, observe that because $\{\lambda_{k,\mathbf{c}_1}^*, \hat{\mu}_{\mathbf{c}_1}^* | k \in \mathcal{K}\}$ satisfies the constraints in P1-1-Dual when $\mathbf{c} = \mathbf{c}_1$, it follows that $\{\lambda_{k,\mathbf{c}_2}^* = \xi\lambda_{k,\mathbf{c}_1}^*, \hat{\mu}_{\mathbf{c}_2}^* = \xi\hat{\mu}_{\mathbf{c}_1}^* | k \in \mathcal{K}\}$ satisfies the constraints in P1-1-Dual when $\mathbf{c} = \mathbf{c}_2$. To check the optimality, observe that $\{\lambda_{k,\mathbf{c}_2}^* = \xi\lambda_{k,\mathbf{c}_1}^*, \hat{\mu}_{\mathbf{c}_2}^* = \xi\hat{\mu}_{\mathbf{c}_1}^* | k \in \mathcal{K}\}$ satisfies (54). Therefore, $\{\lambda_{k,\mathbf{c}_2}^* = \xi\lambda_{k,\mathbf{c}_1}^*, \hat{\mu}_{\mathbf{c}_2}^* = \xi\hat{\mu}_{\mathbf{c}_1}^* | k \in \mathcal{K}\}$ is indeed an optimal solution of $\mathfrak{D}_2$. The lemma then follows by recalling that $\hat{\mu} = \mu + 1$.

Appendix D
Proof of Lemma 2

To obtain the expression of $\text{IF}_{k,q}$ when N goes to infinity, we first rewrite the numerator of $\text{IF}_{k,q}$ as

$$\begin{aligned}
||\mathbf{h}_k^\dagger \mathbf{A}_k^{-1} \boldsymbol{\Theta}_q||^2 &= \sum_{m=1}^{M_q} ||\mathbf{h}_k^\dagger \mathbf{A}_k^{-1} \boldsymbol{\theta}_{q,m}||^2 \\
&= \sum_{m=1}^{M_q} ||\mathbf{h}_k^\dagger (\mathbf{A}_{k,q,m} + c_q \boldsymbol{\theta}_{q,m} \boldsymbol{\theta}_{q,m}^\dagger)^{-1} \boldsymbol{\theta}_{q,m}||^2 \\
&= \sum_{m=1}^{M_q} \left|\left| \frac{\sqrt{l_k l_q} \cdot \hat{\mathbf{h}}_k^\dagger \mathbf{A}_{k,q,m}^{-1} \hat{\boldsymbol{\theta}}_{q,m}}{(1 + c_q l_q \hat{\boldsymbol{\theta}}_{q,m}^\dagger \mathbf{A}_{k,q,m}^{-1} \hat{\boldsymbol{\theta}}_{q,m})} \right|\right|^2 && (55) \\
&= \sum_{m=1}^{M_q} \frac{l_k l_q \hat{\mathbf{h}}_k^\dagger \mathbf{A}_{k,q,m}^{-1} \hat{\boldsymbol{\theta}}_{q,m} \hat{\boldsymbol{\theta}}_{q,m}^\dagger \mathbf{A}_{k,q,m}^{-1} \hat{\mathbf{h}}_k}{(1 + c_q l_q \hat{\boldsymbol{\theta}}_{q,m}^\dagger \mathbf{A}_{k,q,m}^{-1} \hat{\boldsymbol{\theta}}_{q,m})^2} && (56)
\end{aligned}$$

where (56) is true due to Lemma 3. Then, by applying lemma 4 to (56) and the denominator of $\text{IF}_{k,q}$, we have

$$\begin{aligned}
||\mathbf{h}_k^\dagger \mathbf{A}_k^{-1} \boldsymbol{\Theta}_q||^2 &\xrightarrow[N\to\infty]{\text{a.s.}} \sum_{m=1}^{M_q} \left[\frac{l_k l_q \text{tr}\{\hat{\boldsymbol{\theta}}_{q,m}^\dagger \mathbf{A}_{k,q,m}^{-2} \hat{\boldsymbol{\theta}}_{q,m}\}}{(1 + c_q l_q \text{tr}\{\mathbf{A}_{k,q,m}^{-1}\})^2} \right] \\
&\xrightarrow[N\to\infty]{\text{a.s.}} \sum_{m=1}^{M_q} \left[\frac{l_k l_q \text{tr}\{\mathbf{A}_{k,q,m}^{-2}\}}{(1 + c_q l_q \text{tr}\{\mathbf{A}_{k,q,m}^{-1}\})^2} \right] && (57)
\end{aligned}$$

and

$$||\hat{\mathbf{h}}_k^\dagger \mathbf{A}_k^{-1} \hat{\mathbf{h}}_k||^2 \xrightarrow[N\to\infty]{\text{a.s.}} \text{tr}\{\mathbf{A}_k^{-1}\}^2 \tag{58}$$

respectively, where "$\xrightarrow[N\to\infty]{\text{a.s.}}$" denotes almost sure convergence. Moreover, to obtain the value of ε_k when N goes to infinity, observe that

$$\boldsymbol{\Omega}_{[j,k]} = -\frac{l_k |\hat{\mathbf{h}}_j^\dagger \mathbf{A}_j^{-1} \hat{\mathbf{h}}_k|^2}{l_j |\hat{\mathbf{h}}_j^\dagger \mathbf{A}_j^{-1} \hat{\mathbf{h}}_j|^2} \xrightarrow[N\to\infty]{\text{a.s.}} 0 \tag{59}$$

due to lemma 4. Then, $\boldsymbol{\Omega} \xrightarrow[N\to\infty]{\text{a.s.}} \text{diag}[\gamma_1^{-1} \; \gamma_2^{-1} \; \cdots \; \gamma_K^{-1}]$, and

$$\varepsilon_k^2 \xrightarrow[N\to\infty]{\text{a.s.}} \gamma_k \sigma_k^2. \tag{60}$$

Summarizing the results from (22), (57), (58), and (60), we obtain the asymptotic expression of $\text{IF}_{k,q}$ as shown in (23).

Appendix E
Justification on the Approximations in (25)

Here, we show that $\text{tr}\{\mathbf{A}_k^{-1}\} \approx \text{tr}\{\mathbf{T}^{-1}\}$ is valid when M is sufficiently large and $c_q l_q \gg 1 \; \forall q$.

Let η be one of the eigenvalues of $\hat{\boldsymbol{\Theta}}_q \hat{\boldsymbol{\Theta}}_q^\dagger$. The probability density function (PDF) of η is given as [36]

$$p_\eta = \frac{1}{N} \sum_{i=1}^{N} \varphi_i(\eta)^2 \eta^{M-N} e^{-\eta}, \tag{61}$$

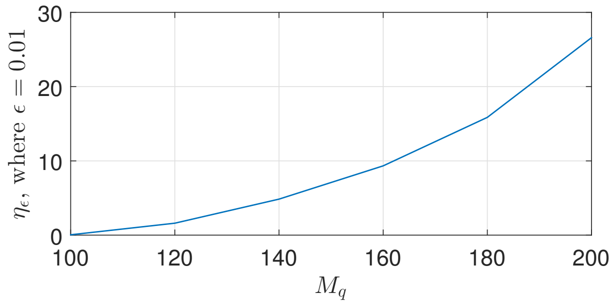


Fig. 7. η_{th} versus M, where $N = 100$ and $\epsilon = 0.01$.

where

$$\varphi_i \triangleq \left[\frac{(i-1)!}{(i-1+M-N)!} \right]^{\frac{1}{2}} L_{i-1}^{M-N}(\eta) \tag{62}$$

and $L_k^{M-N}(\eta) = \frac{1}{k!} e^\eta \eta^{N-M} \frac{d^k}{d\eta^k}(e^{-\eta}\eta^{M-N+k})$ is the Laguerre polynomial of order k. Based on the PDF of η, we numerically search for η_ϵ, where η_ϵ is the threshold value such that the probability of the event $\eta \le \eta_\epsilon$ is equal to ϵ. Fig. 7 shows the value of η_ϵ versus M_q, where $N = 100$ and $\epsilon = 0.01$. We can see that η_ϵ is much larger than one when M_q becomes large. This means that it is very unlikely that any eigenvalue of $\hat{\boldsymbol{\Theta}}_q \hat{\boldsymbol{\Theta}}_q^\dagger$ is close to one for sufficiently large M_q.

Given that η is an eigenvalue of $\hat{\boldsymbol{\Theta}}_q \hat{\boldsymbol{\Theta}}_q^\dagger$, we know that $c_q l_q \eta + 1$ is an eigenvalue of $c_q l_q \hat{\boldsymbol{\Theta}}_q \hat{\boldsymbol{\Theta}}_q^\dagger + \mathbf{I}_N$. Because $c_q l_q \gg 1$ and $\eta \gg 1$ with high probability, we have $c_q l_q \eta \approx c_q l_q \eta + 1$ with high probability. This means that the eigenvalues of $\mathbf{A}_k$ are approximately the same as those of $\mathbf{T}_k$. Consequently, we have $\text{tr}\{\mathbf{A}_k^{-1}\} \approx \text{tr}\{\mathbf{T}_k^{-1}\}$.

The validity of $\text{tr}\{\mathbf{A}_{k,q,m}^{-1}\} \approx \text{tr}\{\mathbf{T}_k^{-1}\}$ and $\text{tr}\{\mathbf{A}_{k,q,m}^{-2}\} \approx \text{tr}\{\mathbf{T}_k^{-2}\}$ can be realized by the fact that $\hat{\boldsymbol{\Theta}}_q \hat{\boldsymbol{\Theta}}_q^\dagger$ has similar eigenvalues as $\hat{\boldsymbol{\Theta}}_q \hat{\boldsymbol{\Theta}}_q^\dagger - \hat{\boldsymbol{\theta}}_{q,m} \hat{\boldsymbol{\theta}}_{q,m}^\dagger$ for sufficiently large M_q, such that the eigenvalues of $\mathbf{A}_{k,q,m}$ are similar to those of $\mathbf{T}_k$.

Appendix F
Approximation of λ_k^*

From (11), when $N \to \infty$, we have

$$\mathbf{v}_k^* = \frac{\mathbf{A}_k^{-1} \mathbf{h}_k}{\mathbf{h}_k^\dagger \mathbf{A}_k^{-1} \mathbf{h}_k} \xrightarrow[N\to\infty]{\text{a.s.}} \frac{\mathbf{A}_k^{-1} \mathbf{h}_k}{l_k \cdot \text{tr}\{\mathbf{A}_k^{-1}\}}. \tag{63}$$

From the argument in Appendix E, we know that the eigenvalues of $\mathbf{A}_k$ and $\mathbf{T}_k$ are approximately the same. Also, $\mathbf{A}_k = \mathbf{T}_k + \mathbf{I}_N$, indicating that the eigenvectors of $\mathbf{A}_k$ and $\mathbf{T}_k$ are the same. Therefore, we use the approximation $\mathbf{A}_k \approx \mathbf{T}_k$.

From (27), $\mathbf{T}_k \approx \varpi_k d_k^{-1} \cdot \mathbf{S}_k$, where $\mathbf{S}_k$ is a Wishart matrix with d_k DoFs and covariance matrix $\mathbf{I}_N$. Then, $\mathbf{T}_k^{-1} \approx \varpi_k^{-1} d_k \cdot \mathbf{S}_k^{-1}$, where $\mathbf{S}_k^{-1}$ is an inverse-Wishart matrix with d_k DoFs and covariance matrix $\mathbf{I}_N$ such that $\mathbb{E}[\text{tr}\{\mathbf{S}_k^{-1}\}] = \frac{N}{d_k - N}$ and $\mathbb{E}[\mathbf{S}_k^{-1}] = \frac{\mathbf{I}_N}{d_k - N - 1}$ [37]. Using respectively $\mathbb{E}[\mathbf{S}_k^{-1}]$

and $\mathbb{E}[\mathrm{tr}\{\mathbf{S}_k^{-1}\}]$ to approximate $\mathbf{S}_k^{-1}$ and $\mathrm{tr}\{\mathbf{S}_k^{-1}\}$, we can approximate $\mathbf{v}_k^*$ as

$$\mathbf{v}_k^* \approx \frac{\mathbb{E}[\mathbf{S}_k^{-1}]\mathbf{h}_k}{l_k \cdot \mathbb{E}[\mathrm{tr}\{\mathbf{S}_k^{-1}\}]} = \frac{(d_k - N)\mathbf{h}_k}{l_k N(d_k - N - 1)} \approx \frac{\hat{\mathbf{h}}_k}{\sqrt{l_k}N}, \quad (64)$$

where we have used the approximation $d_k - N - 1 \approx d_k - N$ for large N in (64).

From lemma 4, when $N \to \infty$ and $\mathbf{v}_k = \frac{\hat{\mathbf{h}}_k}{\sqrt{l_k}N}$ $\forall k$, the left hand side of (8) converges to

$$\frac{\lambda_k \hat{\mathbf{h}}_k^\dagger \mathbf{h}_k \mathbf{h}_k^\dagger \hat{\mathbf{h}}_k}{\hat{\mathbf{h}}_k^\dagger [\sum_{j\in\mathcal{K}\setminus\{k\}} \lambda_j \mathbf{h}_j \mathbf{h}_j^\dagger + \mathbf{\Sigma}(\mathbf{c},\mu)]\hat{\mathbf{h}}_k} \xrightarrow[N\to\infty]{\text{a.s.}} \lambda_k l_k N. \quad (65)$$

Then, the optimal value of λ_k for P1-1-Dual in this case is clearly $\gamma_k l_k^{-1} N^{-1}$.

Appendix G
Approximation of d_k

Define $\check{M} \triangleq \min\{M_q | q \in \mathcal{Q}\}$. When $c_q l_q \geq \lambda_k^* l_k$ $\forall q, k$, we have

$$\begin{aligned}\varpi_k^2 &> \sum_{j\neq k}\sum_q \lambda_j^* l_j c_q l_q M_q + \sum_q c_q^2 l_q^2 M_q^2 \\ &> \sum_{j\neq k} (\lambda_j^*)^2 l_j^2 \check{M} + \sum_q c_q^2 l_q^2 M_q \check{M}. \end{aligned} \quad (66)$$

Then, we can observe that

$$d_k > \left\lfloor \frac{\sum_{j\neq k}(\lambda_j^*)^2 l_j^2 \check{M} + \sum_q c_q^2 l_q^2 M_q \check{M}}{\sum_{j\neq k}(\lambda_j^*)^2 l_j^2 + \sum_q c_q^2 l_q^2 M_q} \right\rfloor = \check{M} > N. \quad (67)$$

Because N is assumed to be large, we can use the following approximation $d_k \approx \varpi_k^2 \varrho_k^{-1}$.

Appendix H
Proof of Theorem 4

From lemma 5 in Appendix A, we can evaluate the average traces of $\mathbf{T}_k^{-1}$ and $\mathbf{T}_k^{-2}$ as

$$\mathbb{E}[\mathrm{tr}\{\mathbf{T}_k^{-1}\}] \approx \varpi_k^{-1} d_k \cdot \mathbb{E}[\mathrm{tr}\{\mathbf{S}_k^{-1}\}] = \frac{\varpi_k^{-1} d_k N}{d_k - N} \quad (68)$$

and

$$\begin{aligned}\mathbb{E}[\mathrm{tr}\{\mathbf{T}_k^{-2}\}] \approx& \varpi_k^{-2} d_k^2 \cdot \mathbb{E}[\mathrm{tr}\{\mathbf{S}_k^{-2}\}] \\ =& \frac{\varpi_k^{-2} d_k^3 N}{(d_k - N)^3 - (d_k - N)}, \end{aligned} \quad (69)$$

respectively. We can then obtain the average interference leakage when $N \to \infty$ as

$$\begin{aligned}\mathbb{E}[\mathrm{IF}_{k,q}] \approx& \frac{\gamma_k l_q \sigma_k^2}{l_k} \cdot \frac{\sum_{m=1}^{M_q} \left[\frac{\mathbb{E}[\mathrm{tr}\{\mathbf{T}_k^{-2}\}]}{(1+c_q l_q \mathbb{E}[\mathrm{tr}\{\mathbf{T}_k^{-1}\}])^2}\right]}{\mathbb{E}[\mathrm{tr}\{\mathbf{T}_k^{-1}\}]^2} \\ =& \frac{\gamma_k l_q \sigma_k^2 M_q \cdot \mathbb{E}[\mathrm{tr}\{\mathbf{T}_k^{-2}\}]}{l_k \cdot \mathbb{E}[\mathrm{tr}\{\mathbf{T}_k^{-1}\}]^2 \cdot (1 + c_q l_q \mathbb{E}[\mathrm{tr}\{\mathbf{T}_k^{-1}\}])^2} \\ =& \frac{\gamma_k l_q \sigma_k^2 M_q \cdot d_k(d_k - N)}{l_k N \cdot [(d_k - N)^2 - 1] \cdot (1 + c_q l_q \frac{\varpi_k^{-1} d_k N}{d_k - N})^2} \\ \approx& \frac{\gamma_k l_q \sigma_k^2 M_q \cdot d_k}{l_k N \cdot (d_k - N) \cdot (1 + c_q l_q \frac{\varpi_k^{-1} d_k N}{d_k - N})^2}, \end{aligned} \quad (70)$$

where (70) is true because

$$\begin{aligned}(d_k - N)^2 - 1 =& d_k^2 - 2d_k N + N^2 - 1 \\ \approx& d_k^2 - 2d_k N + N^2 = (d_k - N)^2 \end{aligned} \quad (71)$$

for large N.

Appendix I
Proof of Theorem 5

Following (29) and (30), we can rewrite $\mathbb{E}[\mathrm{IF}_{k,q}]$ as

$$\mathbb{E}[\mathrm{IF}_{k,q}] = \frac{\gamma_k \sigma_k^2 l_q M_q \cdot \varpi_k^2}{l_k N \cdot (\varpi_k^2 - \varrho_k N) \cdot (1 + c_q l_q \frac{\varpi_k N}{\varpi_k^2 - \varrho_k N})^2}, \ k \in \mathcal{K}. \quad (72)$$

Let $\varpi_{k,q} \triangleq \sum_{j\neq k} \frac{\gamma_k}{N} + \sum_{q'\neq q} c_{q'} l_{q'} M_{q'}$ $\forall k \in \mathcal{K}$ and $\varrho_{k,q} \triangleq \sum_{j\neq k} \frac{\gamma_k^2}{N^2} + \sum_{q'\neq q} c_{q'}^2 l_{q'}^2 M_{q'}$ $\forall k \in \mathcal{K}$, such that $\varpi_k = c_q l_q M_q + \varpi_{k,q}$ and $\varrho_k = c_q^2 l_q^2 M_q + \varrho_{k,q}$ when the approximation in (28) is used. The derivative of $\mathbb{E}[\mathrm{IF}_{k,q}]$ with respect to c_q is

$$\begin{aligned}&\frac{\partial \mathbb{E}[\mathrm{IF}_{k,q}]}{\partial c_q} \\ &= -\frac{2l_q^2 \gamma_k \sigma_k^2 M_q \varpi_k \cdot \left(c_q^3 l_q^3 M_q^2 N \varpi_{k,q} + f_1 + f_2 + f_3\right)}{l_k \cdot \left[c_q^2 l_q^2 M_q^2 + c_q l_q \varpi_{k,q}\left(2M_q + N\right) + f_4\right]^3}, \end{aligned} \quad (73)$$

where

$$f_1 \triangleq c_q^2 l_q^2 M_q^2 \left[(\varpi_{k,q}^2 - N\varrho_{k,q}) + (M_q - N)\varrho_{k,q}\right], \quad (74a)$$
$$f_2 \triangleq c_q l_q \varpi_{k,q} \left[2M_q \varpi_{k,q}^2 + \left(2M_q - N\right)\varrho_{k,q}\right], \quad (74b)$$
$$f_3 \triangleq (\varpi_{k,q}^2 - N\varrho_{k,q}) \cdot (\varpi_{k,q}^2 + \varrho_{k,q} M_q), \quad (74c)$$
$$f_4 \triangleq \varpi_{k,q}^2 - N\varrho_{k,q}. \quad (74d)$$

Because $M_q > N$ $\forall q$, one can easily verify that $\varpi_{k,q}^2 - N\varrho_{k,q} > 0$ $\forall k \in \mathcal{K}$ for $N \to \infty$. Therefore, the functions f_1, f_2, f_3, and f_4 always take positive values. This means that $\frac{\partial \mathbb{E}[\mathrm{IF}_{k,q}]}{\partial c_q} < 0$ $\forall k \in \mathcal{K}$, and therefore $\mathbb{E}[\mathrm{IF}_{k,q}]$ is a decreasing function of c_q for all $k \in \mathcal{K}$.

Because $\mathbb{E}[\mathrm{IF}_q] = \mathbb{E}[\sum_{k=1}^K \mathrm{IF}_{k,q}] = \sum_{k=1}^K \mathbb{E}[\mathrm{IF}_{k,q}]$, where $\mathbb{E}[\mathrm{IF}_{k,q}]$ is a decreasing function of c_q for all k, we conclude that $\mathbb{E}[\mathrm{IF}_q]$ is a decreasing function of c_q.

To find the limit value of $\mathbb{E}[\mathrm{IF}_{k,q}]$ when c_q goes large, we can rewrite (72) as

$$\mathbb{E}[\mathrm{IF}_{k,q}] = \frac{\gamma_k \sigma_k^2 l_q M_q \varpi_k^2 (\varpi_k^2 - \varrho_k N)}{l_k N[(\varpi_k^2 - \varrho_k N) + c_q l_q \varpi_k N]^2} \quad (75)$$

where

$$\varpi_k^2 = (c_q l_q M_q + \varpi_{k,q})^2 = c_q^2 l_q^2 M_q^2 + \mathcal{O}(c_q), \quad (76)$$

$$\varpi_k^2 - \varrho_k N = c_q^2 l_q^2 (M_q^2 - M_q N) + \mathcal{O}(c_q), \quad (77)$$

$$[(\varpi_k^2 - \varrho_k N) + c_q l_q \varpi_k N]^2 = c_q^4 l_q^4 M_q^4 + \mathcal{O}(c_q^3). \quad (78)$$

Note that $\mathbb{E}[\mathrm{IF}_{k,q}]$ is a ratio between two polynomials of c_q, and we only need to know the term of the highest orders of c_q to find the limit, i.e.,

$$\lim_{c_q\to\infty}\mathbb{E}[\mathrm{IF}_{k,q}]=\frac{\gamma_k\sigma_k^2 l_q M_q}{l_k N}\cdot\frac{c_q^4 l_q^4 M_q^2(M_q^2-M_qN)}{c_q^4 l_q^4 M_q^4}$$
$$=\frac{\gamma_k\sigma_k^2 l_q(M_q-N)}{l_k N}. \tag{79}$$

References

[1] Y. Liu, S. Lambotharan, A. Nallanathan, and K. K. Wong, "Pricing based interference control in reversed time division duplex heterogeneous networks," in *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–6.

[2] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfröjd, and T. Svensson, "The role of small cells, coordinated multipoint, and massive MIMO in 5G," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 44–51, May 2014.

[3] W. Xu, L. Liang, H. Zhang, S. Jin, J. Li, and M. Lei, "Performance enhanced transmission in device-to-device communications: Beamforming or interference cancellation?" in *IEEE Global Communications Conference (GLOBECOM)*, Dec. 2012, pp. 4296–4301.

[4] G. Boudreau, J. Panicker, N. Guo, R. Chang, N. Wang, and S. Vrzic, "Interference coordination and cancellation for 4G networks," *IEEE Communications Magazine*, vol. 47, no. 4, pp. 74–81, Apr. 2009.

[5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, April 2013.

[6] J. Dommel, P. P. Knust, L. Thiele, and T. Haustein, "Massive MIMO for interference management in heterogeneous networks," in *IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, June 2014, pp. 289–292.

[7] Y. Shi, J. Zhang, and K. Letaief, "Group sparse beamforming for green cloud-RAN," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2809–2823, May 2014.

[8] Q. Ye, O. Y. Bursalioglu, H. C. Papadopoulos, C. Caramanis, and J. G. Andrews, "User association and interference management in massive MIMO hetnets," *IEEE Transactions on Communications*, vol. 64, no. 5, pp. 2049–2065, May 2016.

[9] C. Mikael, U. Engström, K. W. Helmersson, M. Hashemi, L. Manholm, and P. Wallentin, "Revisiting scheduling in heterogeneous networks when the backhaul is limited," *Ericsson Review*, vol. 91, Nov. 2014.

[10] D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig, and W. Utschick, "Distributed resource allocation schemes," *IEEE Signal Processing Magazine*, vol. 26, no. 5, pp. 53–63, Sep. 2009.

[11] D. Nguyen and T. Le-Ngoc, "Multiuser downlink beamforming in multicell wireless systems: A game theoretical approach," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3326–3338, July 2011.

[12] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, November 2010.

[13] K. Hosseini, J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO and small cells: How to densify heterogeneous networks," in *2013 IEEE International Conference on Communications (ICC)*, June 2013, pp. 5442–5447.

[14] J. Hoydis, K. Hosseini, S. T. Brink, and M. Debbah, "Making smart use of excess antennas: Massive MIMO, small cells, and TDD," *Bell Labs Technical Journal*, vol. 18, no. 2, pp. 5–21, Sept 2013.

[15] M. Kountouris and N. Pappas, "HetNets and massive MIMO: Modeling, potential gains, and performance analysis," in *2013 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)*, Sept 2013, pp. 1319–1322.

[16] S. W. Jeon and C. Suh, "Degrees of freedom of uplink-downlink multiantenna cellular networks," *IEEE Transactions on Information Theory*, vol. 62, no. 8, pp. 4589–4603, Aug 2016.

[17] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.

[18] W. Xu and X. Wang, "Pricing-based distributed downlink beamforming in multi-cell OFDMA networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 9, pp. 1605–1613, Oct. 2012.

[19] J. J. E. Garzás, M. Hong, A. Garcia, and A. García-Armada, "Interference pricing mechanism for downlink multicell coordinated beamforming," *IEEE Transactions on Communications*, vol. 62, no. 6, pp. 1871–1883, June 2014.

[20] C. Pan, W. Xu, J. Wang, H. Ren, W. Zhang, N. Huang, and M. Chen, "Pricing-based distributed energy-efficient beamforming for MISO interference channels," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 710–722, April 2016.

[21] D. N. Nguyen and M. Krunz, "Price-based joint beamforming and spectrum management in multi-antenna cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 11, pp. 2295–2305, Dec. 2012.

[22] T. M. Ho, N. H. Tran, C. T. Do, S. M. A. Kazmi, E. N. Huh, and C. S. Hong, "Power control for interference management and QoS guarantee in heterogeneous networks," *IEEE Communications Letters*, vol. 19, no. 8, pp. 1402–1405, Aug 2015.

[23] D. Fudenberg and J. Tirole, *Game Theory*. MIT Press, 1991.

[24] X. Kang, R. Zhang, and M. Motani, "Price-based resource allocation for spectrum-sharing femtocell networks: A Stackelberg game approach," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 538–549, April 2012.

[25] Y. Wang, X. Wang, and L. Wang, "Low-complexity Stackelberg game approach for energy-efficient resource allocation in heterogeneous networks," *IEEE Communications Letters*, vol. 18, no. 11, pp. 2011–2014, Nov 2014.

[26] N. D. Duong, A. S. Madhukumar, and D. Niyato, "Stackelberg Bayesian game for power allocation in two-tier networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2341–2354, April 2016.

[27] A. Deligiannis, S. Lambotharan, and J. A. Chambers, "Game theoretic analysis for MIMO radars with multiple targets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 6, pp. 2760–2774, December 2016.

[28] T. E. Bogale and L. B. Le, "Massive MIMO and mmwave for 5G wireless HetNet: Potential benefits and challenges," *IEEE Vehicular Technology Magazine*, vol. 11, no. 1, pp. 64–75, March 2016.

[29] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1439–1451, Aug 2006.

[30] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2646–2660, June 2007.

[31] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[32] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Commun. Inf. Theory*, vol. 1, no. 1, pp. 1–182, Jun. 2004.

[33] G. F. Pivaro, S. Kumar, G. Fraidenraich, and C. F. Dias, "On the exact and approximate eigenvalue distribution for sum of Wishart matrices," *IEEE Transactions on Vehicular Technology*, vol. PP, no. 99, pp. 1–1, 2017.

[34] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, February 2013.

[35] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, pp. 49–95, 1994.

[36] İ. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, 1999.

[37] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis (Probability and Mathematical Statistics)*. Academic Press, Jan. 1979.